On Cryptographic Properties of the Cosets of $R(1, m)$

Anne Canteaut, Claude Carlet, Pascale Charpin, and Caroline Fontaine

***Abstract*—We introduce a new approach for the study of weight distributions of cosets of the Reed–Muller code of order 1. Our approach is based on the method introduced by Kasami in [1], using Pless identities. By interpreting some equations, we obtain a necessary condition for a coset to have a "high" minimum weight. Most notably, we are able to distinguish such cosets which have three weights only. We then apply our results to the problem of the nonlinearity of Boolean functions. We particularly study the links between this criterion and the propagation characteristics of a function.**

***Index Terms*—Boolean function, derivation, nonlinearity, propagation criterion, Reed–Muller codes.**

Main Notation

- $\mathcal{B}_m$ is the set of Boolean functions of m variables;
- Ω_f, $f \in \mathcal{B}_m$ is the codeword of length 2^m equal to the ordered list of all values of f; $\mathbf{0}$ and $\mathbf{1}$ denote, respectively, the zero codeword and the all-one codeword;
- $x \cdot y$ denotes the usual dot product between two vectors x and y;
- $V^\perp$ denotes the dual of a subspace $V \subset \boldsymbol{F}_2^m$, i.e.,

$$V^\perp = \{x \in \boldsymbol{F}_2^m \,|\, \forall y \in V, x \cdot y = 0\};$$

- $\{\varphi_\alpha | \alpha \in \boldsymbol{F}_2^m\}$ is the subset of $\mathcal{B}_m$ consisting of all linear functions

$$\varphi_\alpha \colon x \mapsto \alpha \cdot x;$$

- H_α denotes the kernel of φ_α;
- $\mathcal{F}(f), \mathcal{L}(f), \mathcal{MD}(f), \mathcal{V}(f)$, and $\mathcal{N}(f)$ are, respectively, defined by (1), Definition II.1, (5), and (6);
- $\boldsymbol{F}_q$ is the finite field of order q;
- $\mathcal{A}$ is the group algebra $\boldsymbol{F}_2[\{\boldsymbol{F}_2^m, +\}]$;
- $\mathcal{W}$ is the set of two-dimensional affine subspaces of $\boldsymbol{F}_2^m$;
- $\mathcal{W}_0$ is the set of two-dimensional linear subspaces;
- $\langle e_1, \ldots, e_k \rangle$ is the linear space spanned by $e_1, \ldots, e_k$.

I. Introduction

THE general framework of this paper is double: coding theory (and in particular the class of Reed–Muller codes) on one hand and symmetric cryptography (block ciphers and stream ciphers) on the other hand. In both of these general domains, the Boolean functions defined on the set $\boldsymbol{F}_2^m$ of all binary words of length m play an important role. Some open problems on Boolean functions are of most interest in both fields. One of them is the determination of those functions which lie at large Hamming distance from the Reed–Muller code of order 1, $R(1, m)$. This code can be viewed as the set of all affine forms on the m-dimensional vector space $\boldsymbol{F}_2^m$ (an affine form is the sum of a linear form and of one of the constants 0 or 1). The Hamming distance between two Boolean functions is equal to the number of words of $\boldsymbol{F}_2^m$ at which they take different values. The maximum Hamming distance between a general Boolean function and $R(1, m)$ is the covering radius of this code. Its value is known only when m is even or when $m = 1, 3, 5, 7$.

The covering radius of a code is an important parameter, which can be used for analyzing and improving the decoding algorithms devoted to this code. The knowledge of the covering radius of $R(1, m)$ has therefore theoretical and practical importance for coders. It is also a serious challenge for cryptographers: the design of conventional cryptographic systems relies on two fundamental principles introduced by Shannon [2]: *confusion* and *diffusion*. The distance from a Boolean function to the set of all affine functions is called the *nonlinearity* of the function and it allows to quantify some kind of confusion. More precisely, the Boolean functions used in block ciphers must have a large nonlinearity to resist linear attacks [3]; in stream ciphers, the use of highly nonlinear Boolean functions prevents fast correlation attacks [4]. The knowledge of the maximum nonlinearity of Boolean functions is therefore necessary to appreciate (together with other criteria) the practical interest of a given Boolean function for cryptographic applications. Unfortunately, the covering radius of $R(1, m)$ for odd $m \geq 9$ is unknown. We know only that it lies between $2^{m-1} - 2^{\frac{m-1}{2}}$ and $2^{m-1} - 2^{\frac{m}{2}-1}$ (the lower bound can be slightly improved for $m \geq 15$). One aim of this paper is studying, for m odd, those functions whose nonlinearities lie between these two numbers.

For m even, the situation seems better since we know the exact value of the covering radius of $R(1, m)$: $2^{m-1} - 2^{\frac{m}{2}-1}$ (except that the *bent* functions, whose nonlinearity is maximum, are not all determined and that their determination is considered as a difficult open problem). However, from a cryptographic point of view, the case m even is in fact not better than the case m odd, since bent functions are not *balanced* (i.e., their values are not uniformly distributed); bent functions are then usually improper for use in cryptosystems. For this reason, it is also necessary to study those functions which have large but not optimal nonlinearity, say between $2^{m-1} - 2^{\frac{m}{2}}$ and $2^{m-1} - 2^{\frac{m}{2}-1}$. This is what we do also in this paper. Among these functions there are some balanced functions. The maximum nonlinearity of balanced functions is unknown for any $m \geq 8$.

Manuscript received March 8, 2000; revised November 28, 2000.

A. Canteaut and P. Charpin are with the INRIA, Projet CODES, Domaine de Voluceau, Rocquencourt, 78153 Le Chesnay Cedex, France (e-mail: Anne.Canteaut@inria.fr; Pascale.Charpin@inria.fr).

C. Carlet is with GREYC, University of Caen, 14032 Caen Cedex, France (e-mail: Claude.Carlet@inria.fr).

C. Fontaine is with LIFL, University of Sciences and Technology of Lille, 59655 Villeneuve d'Ascq Cedex, France (e-mail: Caroline.Fontaine@lifl.fr).

Communicated by I. F. Blake, Associate Editor for Coding Theory.

Publisher Item Identifier S 0018-9448(01)02714-6.

We study also other cryptographic criteria related to the notion of diffusion. The *strict avalanche criterion* (SAC) was introduced by Webster and Tavares [5] and this concept was generalized into the *propagation criterion* (PC) by Preneel [6] (see also [7]). The SAC, and its generalizations, are based on the properties of the derivatives of Boolean functions. These properties describe the behavior of a function whenever some input coordinates are complemented. We want to point out the relations between the propagation criterion and the nonlinearity. These two criteria are of most interest and form the subject of many current works. The general idea we develop, with these aims, is that the whole Fourier spectra of the functions have to be taken in account. This point of view leads us to consider both the Fourier spectrum of any given Boolean function and the coset of the Reed–Muller code of order 1 generated by the associated codeword. Therefore, several representations are proposed, in particular in the context of group codes, the aim being to have in hand all useful tools.

The paper is organized as follows. Section II is devoted to the presentation of the main tools. We first give basic properties on Boolean functions on $\boldsymbol{F}_2^m$ where the functions are implicitly represented by their *algebraic normal forms*. The study of algebraic properties of Boolean functions of m variables leads us to the study of binary codewords of length 2^m and of their relation with Reed–Muller codes. On the other hand, we need to use any basis in $\boldsymbol{F}_2^m$ and to treat some permutations on $\boldsymbol{F}_2^m$. So the codewords are viewed as formal sums in the binary group algebra $\mathcal{A}$ of the elementary 2-group $\{\boldsymbol{F}_{2^m}, +\}$. Section II is also devoted to the derivation and its significance considering the operations in $\mathcal{A}$.

These tools are applied in Section III, where we study the *maximal odd-weighting subspace* of a given Boolean function f. This concept was recently introduced in [8] and was shown to be linked with the nonlinearity of f. By replacing this concept in the ambient space of Reed–Muller codes, we prove the existence of maximal odd-weighting subspaces, for any f (Theorem III.1).

Section IV is devoted to the study of weight distributions of cosets of $R(1, m)$. By Theorem IV.1 we establish general results on the weight polynomial of any binary linear code of length 2^m and dimension $m+2$. We introduce *almost-optimal cosets* of $R(1, m)$ which correspond to functions with a high nonlinearity (see Definitions II.1 and IV.1). Considering the code $D \cup R(1, m)$, where D is any coset of $R(1, m)$, Corollary IV.1 is then deduced: we show that it is possible to distinguish among almost-optimal cosets those which have three weights only, the *three-valued almost-optimal cosets*. The next subsection is an extension of Corollary IV.1. We exhibit as an indicator of the nonlinearity, the number b_4 of codewords of weight 4 in the dual code. We are more explicit about the computation of b_4 for cosets which are contained in the third-order Reed–Muller code $R(3, m)$.

Note that, when m is odd, the main open problem is the determination of almost-optimal cosets of $R(1, m)$ with unknown weight distributions. But the context is similar for m even, if we consider the problem of the nonlinearity of balanced Boolean functions.

Section V deals with the propagation criterion and its relations with the nonlinearity. A function is said to be almost-optimal (resp., three-valued almost-optimal) if the associated coset of $R(1, m)$ satisfies this property.

In Section V-A, we study the *sum-of-squares indicator* $\mathcal{V}(f)$ of a Boolean function f, which measures the *global avalanche criterion* (GAC)—introduced in [9]. We first give an upper bound on $\mathcal{V}(f)$ in the case where f is almost-optimal (Proposition V.2). This result will have a lot of applications in the sequel of the paper. For instance, we show in this section that an almost-optimal function of degree 3 must have "many" balanced derivatives (Corollary V.1).

We next study the restrictions of a Boolean function f to each coset of any linear subspace of $\boldsymbol{F}_2^m$ (Section V-B). The main result is given by Theorem V.1, where we establish a relation between the Fourier spectrum of f and the Fourier spectra of its restrictions to these subspaces.

In Section V-C, we examine the cases where the derivatives $D_e f$ of a given function f are balanced for any $e \neq 0$ belonging to a subspace of codimension 1 or 2. These cases allow us to obtain some characterizations of bent functions and of three-valued almost-optimal functions. Theorem V.3 is most surprising since it provides a full explanation of links between bent functions and three-valued almost-optimal functions.

In the last section, we consider Boolean functions whose nonbalanced derivatives $D_a f$ exist when a belongs to a subset of rank $k < m$ only. In this case, we can be more precise, by applying the results of Section V-B. We notably characterize the almost-optimal functions which have a linear structure (Corollaries V.4 and V.5). By Theorem V.5, we show that the links between such functions and some of their decompositions are of most interest.

II. Definitions and Basic Properties

The *distance* between two codewords will always be the Hamming distance. The *weight* of any binary vector $\boldsymbol{a} = (a_1, \ldots, a_n)$ will be the Hamming weight

$$\mathrm{wt}(\boldsymbol{a}) = \sum_{i=1}^{n} a_i.$$

The support of $\boldsymbol{a}$, denoted by $\mathrm{supp}(\boldsymbol{a})$, is the set of all labels i such that $a_i \neq 0$.

A. Boolean Functions

We denote by $\mathcal{B}_m$ the set of Boolean functions of m variables. Let $f \in \mathcal{B}_m$; thus, f is a function from $\boldsymbol{F}_2^m$ to $\boldsymbol{F}_2$. The classical representation of f is its *algebraic normal form*

$$f(x_1, \ldots, x_m) = \sum_{u \in \boldsymbol{F}_2^m} \lambda_u \left(\prod_{i=1}^{m} x_i^{u_i} \right), \qquad \lambda_u \in \boldsymbol{F}_2.$$

The *degree* of f, denoted by $\deg(f)$, is the maximal value of $\mathrm{wt}(u)$ such that $\lambda_u \neq 0$. On the other hand, let us denote by Ω_f the codeword equal to the list of all values $f(x)$, $x \in \boldsymbol{F}_2^m$. Then we denote by $\mathcal{F}$ the mapping $\mathcal{F}(f) = \sum_{x \in \boldsymbol{F}_2^m} (-1)^{f(x)}$ related to the *Fourier transform* (see below). It is also related to the weight of the codeword Ω_f

$$\mathcal{F}(f) = \sum_{x \in \boldsymbol{F}_2^m} (-1)^{f(x)} = 2^m - 2\mathrm{wt}(\Omega_f). \tag{1}$$

We denote by φ_α, $\alpha \in \boldsymbol{F}_2^m$, the linear function $x \longmapsto \alpha \cdot x$. Note that the algebraic normal form of φ_α is

$$\varphi_\alpha(x_1, \ldots, x_m) = \sum_{i=1}^{m} \alpha_i x_i, \qquad \alpha_i \in \boldsymbol{F}_2.$$

We now give a list of basic definitions and properties; we keep the above stated notation.

Definition II.1: When $\mathcal{F}(f) = 0$, the function f is said to be *balanced*. The mapping $\alpha \in \boldsymbol{F}_2^m \mapsto \mathcal{F}(f + \varphi_\alpha)$ is called the *Fourier transform* of f. The multiset

$$\{\pm\mathcal{F}(f + \varphi_\alpha) | \alpha \in \boldsymbol{F}_2^m\}$$

is called the *Fourier spectrum* of f. The *nonlinearity* $\mathcal{N}(f)$ of f is the minimum Hamming distance between Ω_f and all codewords associated to the affine functions φ_α and $\varphi_\alpha + 1$. It is equal to $2^{m-1} - \frac{\mathcal{L}(f)}{2}$, where

$$\mathcal{L}(f) = \max_{\alpha \in \boldsymbol{F}_2^m} |\mathcal{F}(f + \varphi_\alpha)|.$$

Note that we are not only interested in the values appearing in the Fourier spectrum, but also in the number of times they occur. The multiset $\{\pm\mathcal{F}(f + \varphi_\alpha)\}$ is often called the *extended* Walsh spectrum (see, for instance, [10]).

The nonlinearity of f being the minimum Hamming weight of the coset $\Omega_f + R(1, m)$ we have $\mathcal{N}(f) \leq \rho(R(1, m))$ where $\rho(R(1, m))$ is the *covering radius* of $R(1, m)$:

$$\rho(R(1, m)) = \max_{\mathrm{y} \in \boldsymbol{F}_2^{2^m}} \left(\min_{\mathrm{x} \in \mathrm{y} + R(1, m)} \mathrm{wt}\,(\mathrm{x}) \right).$$

When m is even, it is known that $\rho(R(1, m)) = 2^{m-1} - 2^{m/2-1}$ and that the Fourier spectrum of functions of maximal nonlinearity is unique [11]. In particular, it does not contain 0 (hence those functions are not balanced).

Definition II.2: A Boolean function $f \in \mathcal{B}_m$, m even, is said to be bent when

$$\mathcal{N}(f) = \rho(R(1, m)) = 2^{m-1} - 2^{m/2-1}.$$

The Fourier spectrum of such a function is $\{\pm 2^{m/2}\}$.

The case where m is odd is completely different. A recent review is given in [12]. We have [13]

$$2^{m-1} - 2^{\frac{m-1}{2}} \leq \rho(R(1, m)) < 2^{m-1} - 2^{\frac{m}{2}-1}.$$

For $m = 3, 5, 7$, $\rho(R(1, m))$ is equal to $2^{m-1} - 2^{\frac{m-1}{2}}$. But the exact lower bound is not known for $m \geq 9$.

A function has a *good nonlinearity* if its nonlinearity is large, i.e., if $\mathcal{L}(f)$ is small. We say that $\mathcal{L}(f)$ is small when

$$2^{m/2} < \mathcal{L}(f) \leq 2^{(m+1)/2}.$$

This corresponds to the case where

$$2^{m-1} - 2^{\frac{m-1}{2}} \leq \mathcal{N}(f) < 2^{m-1} - 2^{\frac{m}{2}-1}.$$

The SAC was generalized into the *propagation criterion* (PC) by Preneel [6]. More recently, Zhang and Zheng introduced the *global avalanche criterion* (GAC), in order to measure the global avalanche characteristics of cryptographic functions [9]. These criteria are based on the properties of the functions $x \longmapsto f(x) + f(x + a)$, $a \in \boldsymbol{F}_2^m$.

Definition II.3: Let f be a Boolean function on $\boldsymbol{F}_2^m$ and $a \in \boldsymbol{F}_2^m$. We denote by $D_a f$ the derivative of f with respect to a

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

i) The *linear space* of f is the linear subspace of those a such that $D_a f$ is a constant function. Such a, $a \neq 0$, is said to be a linear structure of f [14].

ii) Let $E \subset \boldsymbol{F}_2^m$. The function f satisfies *PC with respect to* E if for all $e \in E$ the function $D_e f$ is balanced.

iii) The function f is said to have a good GAC if $|\mathcal{F}(D_a f)|$ is zero or is very close to zero for most nonzero a.

We now recall some fundamental formulas. Parseval's relation

$$\sum_{\alpha \in \boldsymbol{F}_2^m} \mathcal{F}^2(f + \varphi_\alpha) = 2^{2m} \tag{2}$$

and a formula which states the link between f and its derivatives

$$\begin{aligned} \mathcal{F}^2(f + \varphi_\alpha) &= \sum_{a \in \boldsymbol{F}_2^m} \mathcal{F}(D_a f + \varphi_\alpha(a)) \\ &= \sum_{a \in \boldsymbol{F}_2^m} (-1)^{\alpha \cdot a} \mathcal{F}(D_a f). \end{aligned} \tag{3}$$

This was proved by Carlet in [15] and [16], giving particularly

$$\mathcal{F}^2(f) = \sum_{a \in \boldsymbol{F}_2^m} \mathcal{F}(D_a f). \tag{4}$$

In [9], the authors propose two indicators related to the GAC: we shall denote by $\mathcal{MD}(f)$ the *absolute indicator*

$$\mathcal{MD}(f) = \max_{a \in \boldsymbol{F}_2^m, a \neq 0} |\mathcal{F}(D_a f)| \tag{5}$$

and by $\mathcal{V}(f)$ the second moment of the autocorrelation coefficients called the *sum-of-squares* indicator

$$\mathcal{V}(f) = \sum_{a \in \boldsymbol{F}_2^m} \mathcal{F}^2(D_a f) = \sum_{a,\, b \in \boldsymbol{F}_2^m} \mathcal{F}(D_a D_b f). \tag{6}$$

Note that obviously $\mathcal{V}(f) \geq 2^{2m}$, since $\mathcal{F}^2(D_0 f) = 2^{2m}$. The next formula provides a relation between $\mathcal{V}(f)$ and the Fourier spectrum of f, i.e., the values $|\mathcal{F}(f + \varphi_\alpha)|$, $\alpha \in \boldsymbol{F}_2^m$.

Proposition II.1: For any Boolean function $f \in \mathcal{B}_m$, we have

$$\begin{aligned} \forall\, e \in \boldsymbol{F}_2^m, \quad \sum_{\alpha \in \boldsymbol{F}_2^m} \mathcal{F}^2(f + \varphi_\alpha) \mathcal{F}^2(f + \varphi_{\alpha+e}) \\ = 2^m \sum_{a \in \boldsymbol{F}_2^m} \mathcal{F}^2(D_a f)\, (-1)^{e \cdot a} \end{aligned}$$

providing, for $e = 0$, a relation between the Fourier spectrum of f and the sum-of-squares indicator defined by (6)

$$\sum_{\alpha \in \boldsymbol{F}_2^m} \mathcal{F}^4(f + \varphi_\alpha) = 2^m \mathcal{V}(f). \tag{7}$$

Proof: Let

$$G = \sum_{\alpha \in \boldsymbol{F}_2^m} \mathcal{F}^2(f + \varphi_\alpha) \mathcal{F}^2(f + \varphi_{\alpha+e}).$$

According to (3), we have for all $e \in \boldsymbol{F}_2^m$

$$G = \sum_{\alpha \in \boldsymbol{F}_2^m} \left(\sum_{a \in \boldsymbol{F}_2^m} \mathcal{F}(D_a f)\,(-1)^{a \cdot \alpha} \right) \times \left(\sum_{b \in \boldsymbol{F}_2^m} \mathcal{F}(D_b f)\,(-1)^{b \cdot (\alpha + e)} \right)$$

$$= \sum_{\alpha \in \boldsymbol{F}_2^m} \sum_{a,\, b \in \boldsymbol{F}_2^m} \mathcal{F}(D_a f)\mathcal{F}(D_b f)\,(-1)^{b \cdot e + \alpha \cdot (a+b)}$$

$$= \sum_{a,\, b \in \boldsymbol{F}_2^m} \mathcal{F}(D_a f)\mathcal{F}(D_b f)\,(-1)^{b \cdot e} \sum_{\alpha \in \boldsymbol{F}_2^m} (-1)^{\alpha \cdot (a+b)},$$

where

$$\sum_{\alpha \in \boldsymbol{F}_2^m} (-1)^{\alpha \cdot (a+b)} = 0$$

unless $a = b$. Then we deduce, for any $e \in \boldsymbol{F}_2^m$

$$G = 2^m \sum_{a \in \boldsymbol{F}_2^m} \mathcal{F}^2(D_a f)\,(-1)^{e \cdot a}.$$

We complete the proof by using the definition of the sum-of-squares indicator given by (6). □

Our purpose is to point out that there are interesting connections between the GAC and the nonlinearity. Note, as a trivial example, that the bent functions—i.e., the functions which have the best nonlinearity $2^{m-1} - 2^{\frac{m}{2}-1}$ for m even—have a perfect GAC, since their derivatives are all balanced. *For such a function f, we have $\mathcal{MD}(f) = 0$ and $\mathcal{V}(f) = 2^{2m}$. Moreover these equalities hold for bent functions only.*

On the other hand, a function f which has a linear space V satisfies

$$|\mathcal{F}(D_a f)| = 2^m, \qquad \forall\, a \in V \tag{8}$$

(see Definition II.3). Hence, $\mathcal{MD}(f)$ takes the maximal value and one can say that f has not a good GAC. We obviously deduce a lower bound for $\mathcal{V}(f)$.

Lemma II.1: A function f, which has a linear space V of dimension k, $k \geq 1$, satisfies (8) and is such that $\mathcal{V}(f) \geq 2^{2m+k}$.

However, the nonlinearity of a function f which has a linear structure is not always so bad. We will show later that there exist such functions satisfying $\mathcal{L}(f) = 2^{(m+2)/2}$ for even m and $\mathcal{L}(f) = 2^{(m+1)/2}$ for odd m (see Corollaries V.4 and V.5).

For clarity, we notice that $\mathcal{MD}(f)$ and $\mathcal{V}(f)$ are invariant if we change f into $f + \varphi_\alpha$: since $D_a \varphi_\alpha$ is a constant function, we have

$$|\mathcal{F}(D_a f)| = |\mathcal{F}(D_a(f + \varphi_\alpha))|, \qquad \text{for any } \alpha$$

implying the next property.

Lemma II.2: For any $\alpha \in \boldsymbol{F}_2^m$, we have

$$\mathcal{MD}(f + \varphi_\alpha) = \mathcal{MD}(f) \quad \text{and} \quad \mathcal{V}(f + \varphi_\alpha) = \mathcal{V}(f).$$

We want to end this section with few elements on *resilient functions*. In this paper, we do not emphasize the criterion of correlation immunity. However, this concept is strongly related to the properties of balanced functions and thus with our next results (see [10, Sec. 7]).

Definition II.4: Let $\boldsymbol{e} = (e_1, \ldots, e_m)$ be any basis of $\boldsymbol{F}_2^m$. A function $f \in \mathcal{B}_m$ is said to be *correlation-immune of order k, $1 \leq k \leq m$, with respect to $\boldsymbol{e}$* if for any vector $\alpha = (\alpha_1, \ldots, \alpha_m)$ in $\boldsymbol{F}_2^m$ such that $0 < \mathrm{wt}(\alpha) \leq k$, the function

$$f + \varphi_\alpha, \qquad \alpha = \sum_{i=1}^{m} \alpha_i e_i$$

is balanced. The function f is said to be *resilient of order k* if it is additionally balanced.

We now recall the simplest link between nonlinearity and resiliency.

Proposition II.2: Let $f \in \mathcal{B}_m$. Let us denote by ν the number of 0's in the Fourier spectrum of f. Then we have

$$\nu \leq 2^m - \frac{2^{2m}}{\mathcal{L}(f)^2}$$

with equality if and only if the values occurring in the Fourier spectrum of f lie in $\{0, \pm\mathcal{L}(f)\}$.

Most notably, this implies:

- for m even, if $\mathcal{L}(f) \leq 2^{(m+2)/2}$ then $\nu \leq 2^{m-1} + 2^{m-2}$, with equality if and only if the values occurring in the Fourier spectrum of f lie in $\{0, \pm 2^{(m+2)/2}\}$;
- for m odd, if $\mathcal{L}(f) \leq 2^{(m+1)/2}$ then $\nu \leq 2^{m-1}$, with equality if and only if the values occurring in the Fourier spectrum of f lie in $\{0, \pm 2^{(m+1)/2}\}$.

Proof: We simply use Parseval's relation (see (2)). Let A be the set of all α such that $f + \varphi_\alpha$ is not balanced. Then we have

$$\sum_{\alpha \in A} \mathcal{F}^2(f + \varphi_\alpha) = 2^{2m}.$$

Since $|A| = 2^m - \nu$, we deduce that

$$\mathcal{L}(f)^2(2^m - \nu) \geq 2^{2m}$$

i.e.,

$$\nu \leq 2^m - \frac{2^{2m}}{\mathcal{L}(f)^2}.$$

Equality in the above formula holds if and only if all nonzero values of the Fourier spectrum are equal to $\pm\mathcal{L}(f)$. □

Remark II.1: By the previous property we give a significant upper bound on the number of balanced functions $f + \varphi_\alpha$, when f has a good nonlinearity. This contradicts a high order of resiliency.

B. Product and Derivation

The study of properties of Boolean functions of m variables leads us to the study of binary codewords of length 2^m. More generally, any set of Boolean functions provides a set of codewords and can be studied by means of tools of coding theory.

The main concern is with Reed–Muller codes as we first state in the next definition.

Definition II.5: The Reed–Muller code of length 2^m and order r, $1 \le r \le m$, denoted by $R(r, m)$, is the binary code of length 2^m composed of the codewords Ω_f where f is a Boolean function of m variables whose degree is less than or equal to r.

We described above some properties of $\mathcal{B}_m$ by taking the standard basis in $\boldsymbol{F}_2^m$. It is clear that any basis can be chosen. From now on, we will consider that $f \in \mathcal{B}_m$ is a function from $\boldsymbol{F}_2^m$ to $\boldsymbol{F}_2$ where $\boldsymbol{F}_2^m$ is viewed as an additive group. We will fix a basis in $\boldsymbol{F}_2^m$ when it will be necessary. However, we have to mention that generally, for cryptographic applications, the basis is fixed and the properties have to be considered relatively to the chosen basis.

The concept of "derivative" can be seen as a multiplication in a group algebra, the ambient space of binary codes of length 2^m. We begin by recalling some definitions and properties. An extensive study was made by Assmus and Key in [17] and Charpin in [18] and [19]; we only give basic elements for the use of the algebraic tools which are provided here.

Definition II.6: Let us denote by $\mathcal{A}$ the group algebra $\boldsymbol{F}_2[\boldsymbol{F}_2^m]$. The algebra $\mathcal{A}$ is the set of all binary words of length 2^m; such a word x is a formal polynomial

$$\mathrm{x} = \sum_{g \in \boldsymbol{F}_2^m} x_g X^g, \qquad x_g \in \boldsymbol{F}_2.$$

The operations are

$$\begin{aligned} a\mathrm{x} + b\mathrm{y} &= a \sum_{g \in \boldsymbol{F}_2^m} x_g X^g + b \sum_{g \in \boldsymbol{F}_2^m} y_g X^g \\ &= \sum_{g \in \boldsymbol{F}_2^m} (a x_g + b y_g) X^g, \\ \mathrm{xy} &= \sum_{g \in \boldsymbol{F}_2^m} x_g X^g \times \sum_{g \in \boldsymbol{F}_2^m} y_g X^g \\ &= \sum_{g \in \boldsymbol{F}_2^m} \left(\sum_{\substack{h,\, k \in \boldsymbol{F}_2^m \\ h+k=g}} x_h y_k \right) X^g. \end{aligned}$$

where $a \in \boldsymbol{F}_2$, $b \in \boldsymbol{F}_2$, $\mathrm{x} \in \mathcal{A}$, $\mathrm{y} \in \mathcal{A}$. Note that the multiplicative unit is X^0. The all-one vector and the null vector will be denoted by $\mathbf{1}$ and $\mathbf{0}$, respectively. By convention, X^0 is denoted 1. An ideal I of $\mathcal{A}$ is a subgroup (and, thus, a subspace) invariant under the multiplication by X^b, for some b. The algebra $\mathcal{A}$ has only one maximal ideal, called its *radical*, which is the set of all words of even weights

$$\mathcal{P} = \left\{ \sum_{g \in \boldsymbol{F}_2^m} x_g X^g \,\middle|\, \sum_{g \in \boldsymbol{F}_2^m} x_g = 0 \bmod 2 \right\}.$$

Thus, we can define the ideals $\mathcal{P}^j$, $1 \le j \le m+1$, generated by the products $\prod_{i=1}^{j} \mathrm{x}_i$, $\mathrm{x}_i \in \mathcal{P}$, providing the decreasing sequence

$$\mathcal{A} = \mathcal{P}^0 \supset \mathcal{P} \supset \cdots \supset \mathcal{P}^{m-1} \supset \mathcal{P}^m = \{\mathbf{0}, \mathbf{1}\}$$

where $\mathcal{P}^i \mathcal{P}^j = \mathcal{P}^{i+j}$ and $\mathcal{P}^{m+1} = \{\mathbf{0}\}$. Recall the fundamental result, due to Berman [20] (see also [17, Theorem 4.2]).

Theorem II.1: The powers of the radical of the algebra $\mathcal{A}$ are the Reed–Muller codes. More precisely, for any r, $R(r, m) = \mathcal{P}^{m-r}$.

In the sequel, we will generally use the notation $\mathcal{P}^j$ when we have to handle some multiplications in $\mathcal{A}$. Recall that $\mathcal{P}^j$ *is the subspace generated by the codewords whose supports are the* j*-dimensional subspaces* of $\boldsymbol{F}_2^m$[17, Corollary 3.11]

$$\sum_{v \in V} X^v = \prod_{i=1}^{j} (X^{v_i} + 1),\ V = \langle v_1, \ldots, v_j \rangle. \tag{9}$$

The so-called *Jenning's Basis* provides a basis of $\mathcal{A}$ containing a basis of each $\mathcal{P}^j$ as we recall in the next proposition—a proof, for any characteristic, can be found in [17, p. 1299].

Proposition II.3: Let $(e_1, \ldots, e_m)$ be a basis of $\boldsymbol{F}_2^m$. Then the set

$$\left\{ \prod_{i=1}^{m} (X^{e_i} + 1)^{k_i} \,\middle|\, (k_1, \ldots, k_m) \in \{0, 1\}^m \right\}$$

is a basis of $\mathcal{A}$. Moreover, for each j, $1 \le j \le m$, the set

$$\left\{ \prod_{i=1}^{m} (X^{e_i} + 1)^{k_i} \,\middle|\, \sum_{i=1}^{m} k_i \ge j \right\}$$

is a basis of $\mathcal{P}^j$, the Reed–Muller code of order $m - j$.

Let $f \in \mathcal{B}_m$. The associated codeword of f is written as follows in $\mathcal{A}$:

$$\Omega_f = \sum_{g \in \boldsymbol{F}_2^m} f(g) X^g.$$

So we clearly have

$$\mathcal{N}(f) = \min_{\alpha \in \boldsymbol{F}_2^m} \left\{ \mathrm{wt}\,(\mathrm{x}) \,\middle|\, \mathrm{x} = \sum_{g \in \boldsymbol{F}_2^m} (f(g) + \varphi_\alpha(g)) X^g \right\}. \tag{10}$$

On the other hand, for any $a \in \boldsymbol{F}_2^m$, we have

$$X^a \Omega_f = \sum_{g \in \boldsymbol{F}_2^m} f(g) X^{g+a} = \sum_{g \in \boldsymbol{F}_2^m} f(g+a) X^g, \tag{11}$$

showing that $(X^a + 1)\Omega_f$ is the associated codeword of $D_a f$.

More generally, the concept of kth-*derivative*, given in the next definition, is actually a multiplication in the algebra $\mathcal{A}$.

Definition II.7: Let V be a k-dimensional subspace of $\boldsymbol{F}_2^m$. The kth-*derivative* of $f \in B_m$ *with respect to* V is the function

$$D_{a_1, \ldots, a_k} f = D_{a_1} D_{a_2} \cdots D_{a_k} f$$

where $(a_1, \ldots, a_k)$ is any basis of V.

Proposition II.4: Let V be a k-dimensional subspace of $\boldsymbol{F}_2^m$; $(a_1, \ldots, a_k)$ denotes any basis of V. Let $f \in \mathcal{B}_m$ be any function of degree r. Set $h = D_{a_1, \ldots, a_k} f$. Then

$$\Omega_h = \sum_{g \in \boldsymbol{F}_2^m} \left(\sum_{a \in V} f(g+a) \right) X^g = \left(\sum_{v \in V} X^v \right) \Omega_f.$$

The degree of h is less than or equal to $r-k$. When $r<k$, h is the zero function. In particular, the derivative of f with respect to a has degree at most $r-1$ and corresponds to the product by X^a+1 in $\mathcal{A}$

$$\Omega_{D_a f} = (X^a+1)\Omega_f.$$

Proof: We deduce from (11)

$$(X^a+1)\Omega_f = \sum_{g\in \boldsymbol{F}_2^m} (f(g)+f(g+a))X^g = \Omega_{D_a f}.$$

Set $\mathrm{y} = \sum_{v\in V} X^v$, the codeword of support V. The general formula is easily obtained by expanding the product $(\mathrm{y}\Omega_f)$. For instance,

$$D_{a_1}D_{a_2}f(g) = f(g)+f(g+a_1)+f(g+a_2)+f(g+a_1+a_2).$$

The codeword y is in $\mathcal{P}^k$, by definition (see (9)). Assume that f has degree r—this means that the codeword Ω_f is in the Reed–Muller code of order r. So, from Theorem II.1, $\Omega_f \in \mathcal{P}^{m-r}$ implying that the product $\mathrm{y}\Omega_f$ is in $\mathcal{P}^k\mathcal{P}^{m-r} = \mathcal{P}^{m+k-r}$, which is the Reed–Muller code of order $r-k$. So the degree of h is less than or equal to $r-k$. □

In the next section, we will develop a concept directly stemming from the concept of derivation. To end this section we give some obvious properties and mention an important class of functions. Note that $f=1$ (resp., $f=0$) means that the function f is constant, with associated codeword $\mathbf{1}$ (resp., $\mathbf{0}$).

Proposition II.5: Let $f\in\mathcal{B}_m$. Then we have the following.

1) If there exists $a\in \boldsymbol{F}_2^m$ such that $D_a f = 1$ then f is balanced.
2) When $\deg(f)\le 2$, f is balanced if and only if there exists $a\in \boldsymbol{F}_2^m$ such that $D_a f = 1$.
3) When $\deg(f)\le 3$, $D_a f$ is balanced if and only if there exists $b\in \boldsymbol{F}_2^m$ such that $D_a D_b f = 1$.

Proof: For proving the first property, it is sufficient to notice that $\Omega_f + X^a\Omega_f = \mathbf{1}$ implies that $2\mathrm{wt}(\Omega_f) = 2^m$. We recall the proof of the second property in Appendix I. The third property is then deduced, since $D_a f$ has degree at most 2 when $\deg(f)\le 3$. □

Example II.1: The above property allows us to characterize a large class of balanced functions by means of their associated codewords. Let H be any subspace of codimension 1 in $\boldsymbol{F}_2^m$. The weight of the following codewords x is 2^{m-1}:

$$\mathrm{x} = (X^e+1)\mathrm{y}+\mathrm{z}, \quad \mathrm{z} = \sum_{h\in H} X^h, \quad \text{and} \quad e\notin H.$$

Indeed, x is balanced for any y, since

$$(X^e+1)\mathrm{x} = (X^e+1)\mathrm{z} = \mathbf{1}.$$

The corresponding functions have a linear structure.

The *partially bent functions* were introduced by Carlet in [16]. These functions are quadratic-like functions, in the sense that the dimension of their linear space is sufficient for determining their Fourier spectra. With our terminology we obtain directly, from [16, p. 137], the form of the codewords corresponding to partially bent functions.

Proposition II.6: A Boolean function f of m variables is said to be *partially bent* if there exists a basis $(e_1, \ldots, e_m)$ of $\boldsymbol{F}_2^m$ such that $f = g+\varphi_\alpha$, where g is a bent function on the $(m-k)$-dimensional space $\langle e_{k+1}, \ldots, e_m\rangle$ for some $k<m$ such that $m-k$ is even, and φ_α is a linear function.

The codewords corresponding to partially bent functions have the following form:

$$\Omega_f = \prod_{i=1}^{k}(X^{e_i}+1)\left(\sum_{a\in\langle e_{k+1},\ldots,e_m\rangle} g(a)X^a\right) + \Omega_{\varphi_\alpha}.$$

Note that $\langle e_1, \ldots, e_k\rangle$ is the linear space of f and that Ω_f lies in $\mathcal{P}^{\frac{m-k}{2}}$. Moreover, *$f$ is a balanced partially bent function* if and only if there is $c\in\boldsymbol{F}_2^m$ such that $D_c f = 1$ (see [16, Proposition 2] and Appendix I).

Open Problem II.1: Since any quadratic function is partially bent, the derivatives of any function of degree 3 are partially bent. Characterize a class of functions of degree r, $r>3$, whose derivatives are all partially bent.

Notice that there exist bent functions whose derivatives are not all partially bent. Consider, for instance, Maiorana–McFarland functions: we identify the elements of $\boldsymbol{F}_2^m$, $m=2t$, with the pairs (x, y) where $x=(x_1, \ldots, x_t)$ and $y=(y_1, \ldots, y_t)$ and we define

$$f(x, y) = x\cdot\pi(y) + g(y)$$

where π is some bijection from $\boldsymbol{F}_2^t$ to $\boldsymbol{F}_2^t$ (with the usual dot product "$\cdot$") and g is some function in $\mathcal{B}_t$. The derivative $D_{e_i}f$ of f with respect to the ith word of weight 1, e_i, for $1\le i\le t$, is equal to the ith coordinate function π_i of π. Since the m-variable function $D_{e_i}f$ only depends on t variables, its linear space has dimension at least $m-t=t$. Recall that the degree of a partially bent function is at most the half of the codimension of its linear space [16]. The derivative $D_{e_i}f$ cannot be partially bent if the degree of π_i is greater than $t/2$. This situation occurs, for example, if $\pi(y)=y^s$ where $\boldsymbol{F}_2^t$ is identified with the finite field with 2^t elements, and where s is such that $\gcd(s, 2^t-1)=1$ and the binary expansion of s contains more than $t/2$ 1's. An example of such π is $\pi(y)=y^{2^{t-1}-1}$ for $t\ge 3$.

III. Maximal Odd-Weighting Subspaces of Boolean Functions

Zheng, Zhang, and Imai introduced in [8] the *maximal odd-weighting subspace* of a given Boolean function f. They indicated the link between this concept and the nonlinearity of f. Replacing their concept in the ambient space $\mathcal{A}$ of Reed–Muller codes, we deduce additional properties.

Lemma III.1: Let V be a k-dimensional subspace of $\boldsymbol{F}_2^m$. Set $\mathrm{y} = \sum_{v\in V} X^v$ and $\lambda = 2^{m-k}$. We denote by $V_1, \ldots, V_\lambda$ the

λ cosets of V where $V_1 = V$. Let $\mathrm{x} \in \mathcal{A}$ and, for each i, denote by x_i the restriction of x to V_i. Then the product xy satisfies

$$\mathrm{xy} = \sum_{\substack{1 \le i \le \lambda \\ \mathrm{wt}\,(\mathrm{x}_i) \text{ is odd}}} \sum_{g \in V_i} X^g.$$

Furthermore,

i) $\mathrm{xy} = \mathbf{0}$ (resp., $= \mathbf{1}$) if and only if the weight of x_i is even (resp., odd) for all i, $1 \le i \le \lambda$;

ii) $\mathrm{wt}\,(\mathrm{xy}) = \lambda_o \times 2^k$, where λ_o is the number of x_i which have odd weights.

Proof: We have

$$\mathrm{x} = \sum_{g \in \boldsymbol{F}_2^m} x_g X^g$$

where $\boldsymbol{F}_2^m = \bigcup_{1 \le i \le \lambda} V_i$. Setting $V_i = a_i + V$, for each i, with $a_1 = 0$, we obtain

$$\mathrm{x_i} = \sum_{g \in V_i} x_g X^g = \sum_{u \in V} x_{a_i+u} X^{a_i+u} = X^{a_i} \sum_{u \in V} x_{a_i+u} X^u.$$

Now

$$\begin{aligned}
\mathrm{xy} &= \sum_{i=1}^{\lambda} \sum_{g \in V_i} x_g X^g \sum_{v \in V} X^v \\
&= \sum_{i=1}^{\lambda} X^{a_i} \sum_{u \in V} x_{a_i+u} X^u \sum_{v \in V} X^v \\
&= \sum_{i=1}^{\lambda} X^{a_i} \left(\sum_{u \in V} x_{a_i+u} \right) \left(\sum_{v \in V} X^v \right) \\
&= \sum_{i=1}^{\lambda} \left(\sum_{g \in V_i} X^g \right) \times (\mathrm{wt}\,(\mathrm{x}_i) \bmod 2)
\end{aligned}$$

giving the main formula. Note that $X^u \mathrm{y} = \mathrm{y}$, for any $u \in V$.

Since $\sum_{g \in V_i} X^g$ is the all-one vector of length 2^k and support V_i, i) and ii) are immediately deduced. □

Proposition III.1: Let $\mathrm{x} \in \mathcal{A}$ and $1 \le j \le m$. Then

i) x lies in $R(m-j, m)$ if and only if for any subspace W of $\boldsymbol{F}_2^m$ of dimension $m-j+1$, we have:

$$\mathrm{x} \left(\sum_{v \in W} X^v \right) = \mathbf{0}$$

— i.e., the restriction of x to each coset of W has an even weight.

ii) x lies in $R(m-j, m) \setminus R(m-j-1, m)$ if and only if $\mathrm{x} \in R(m-j, m)$ and there is a subspace V of $\boldsymbol{F}_2^m$ of dimension $m-j$ such that

$$\mathrm{x} \left(\sum_{v \in V} X^v \right) = \mathbf{1}$$

— i.e., the restriction of x to each coset of V has an odd weight.

Proof: Recall that $R(m-j, m) = \mathcal{P}^j$, implying

$$(\mathcal{P}^j)^\perp = R(m-j, m)^\perp = R(j-1, m) = \mathcal{P}^{m-j+1}.$$

Remember that any element can be represented with respect to a *Jenning's Basis* (see Theorem II.1 and Proposition II.3).

The code $\mathcal{P}^j$ is generated by the codewords whose supports are the subspaces U of dimension j [17, Corollary 3.11]. The dual of $\mathcal{P}^j$ is the code $\mathcal{P}^{m-j+1}$, which is generated by the codewords whose supports are the subspaces W of dimension $m-j+1$. Since

$$\mathcal{P}^j \mathcal{P}^{m-j+1} = \mathcal{P}^{m+1} = \{\mathbf{0}\}$$

we obviously have $\mathrm{x} \in \mathcal{P}^j$ if and only if the product of x with any generator of $\mathcal{P}^{m-j+1}$ is $\mathbf{0}$, completing the proof of i).

Assume that $\mathrm{x} \in \mathcal{P}^j$. The dual of $\mathcal{P}^{j+1}$ being $\mathcal{P}^{m-j}$, we have $\mathrm{x} \notin \mathcal{P}^{j+1}$ if and only if at least one generator with support V, say $\mathrm{y} = \sum_{v \in V} X^v$, where $\dim V = m-j$, satisfies $\mathrm{xy} \neq \mathbf{0}$. Since

$$\mathcal{P}^j \mathcal{P}^{m-j} = \mathcal{P}^m = \{\mathbf{0}, \mathbf{1}\}$$

we can conclude that $\mathrm{xy} = \mathbf{1}$, completing the proof of ii). □

Now we give the definition of Zheng *et al.* [8].

Definition III.1: Let f be a Boolean function on $\boldsymbol{F}_2^m$. Let U be some k-dimensional subspace of $\boldsymbol{F}_2^m$. Denote by f_U the restriction of f to U, i.e., the function on U defined by $f_U(x) = f(x)$.

Then U is said to be a *maximal odd-weighting subspace* of f if the weight of the codeword corresponding to f_U is odd and the weight of the codeword corresponding to $f_{U'}$ is even for all subspace U' which strictly contains U.

Using Proposition III.1 we are able to complete this definition.

Theorem III.1: Let f be a Boolean function of degree r. Recall that Ω_f denotes the corresponding codeword of f. Let U be a k-dimensional subspace of $\boldsymbol{F}_2^m$ and set $\mathrm{y} = \sum_{u \in U} X^u$. Then we have

a) U is a maximal odd-weighting subspace of f if and only if the product $\mathrm{y}\Omega_f$ is equal to the all-one codeword; or, equivalently, if the kth-derivative of f with respect to U, say $D_{e_1} \cdots D_{e_k} f$ for some basis $(e_1, \ldots, e_k)$ of U, is equal to the constant function 1.

b) If U is a maximal odd-weighting subspace of f, then $k \le r$. Moreover, there exists at least one r-dimensional maximal odd-weighting subspace of f.

Proof: By definition, U is a maximal odd-weighting subspace of f if and only if the weight of the restriction of Ω_f to U and to any coset L of U is odd. This is because the set $L \cup U$ is a subspace containing U and any subspace containing U is a union of an even number of cosets of U. In accordance with Lemma III.1, we obtain: U is a maximal odd-weighting subspace of f if and only if $\mathrm{y}\Omega_f = \mathbf{1}$. Since $\mathrm{y} = \prod_{i=1}^{k} (X^{e_i} + 1)$, then $\mathrm{y}\Omega_f$ is the codeword corresponding to $D_{e_1} \cdots D_{e_k} f$ (see Proposition II.4), completing the proof of a).

Since f has degree r, Ω_f is in $R(r, m) \setminus R(r-1, m)$. From Proposition III.1 ii) and from a), there exists V of dimension r which is a maximal odd-weighting subspace of f. Moreover,

from Proposition III.1 i), it is not the case for any W of dimension $k > r$. □

Remark III.1: In their paper, Zheng *et al.* noticed that if U is a maximal odd-weighting subspace of f of dimension k, $k > 2$, then $\mathcal{N}(f) \geq 2^{m-k}$. Note that it is simply because (with the above notation)

$$2^m = \mathrm{wt}\,(\mathrm{y}\Omega_f) \leq \mathrm{wt}\,(\mathrm{y})\mathrm{wt}\,(\Omega_f) = 2^k \mathrm{wt}\,(\Omega_f).$$

Moreover, when $k > 2$, this inequality holds for $f + \varphi_\alpha$, for any α, since any second derivative of φ_α is 0.

Note that, according to Proposition II.5, a Boolean function of degree 3 has a maximal odd-weighting subspace of dimension 2 as soon as it has a balanced derivative.

IV. The Weights of Cosets of the Reed–Muller Code of Order 1

In this section, we study the nonlinearity through the properties of weight polynomials of cosets of $R(1, m)$. To be more precise, we establish a necessary condition for such a coset to have a high minimum weight.

A. An Extension of the Results of Kasami

The major result of this section is presented in Theorem IV.1, providing a new point of view on the characterization of the weight distributions of the cosets $\mathrm{x} + R(1, m)$ for any m. This result is based on Pless identities, introduced by Pless in [21], and which are obtained from MacWilliams identities (see also [22, Ch. 5]).

Let C denote an $[n, k, \delta]$ binary linear code, and $C^\perp$ its dual, which has δ' as minimum distance. Let us denote by a_w (resp., b_w), $w \in [0, n]$, the number of codewords of C (resp., $C^\perp$) whose Hamming weight is w. If $\delta' \geq 4$, then we have the following Pless identities (see [22, p. 130]):

$$\begin{aligned}
&\sum_{w=0}^{n} a_w = 2^k \\
&\sum_{w=0}^{n} w\, a_w = 2^{k-1}\, n \\
&\sum_{w=0}^{n} w^2\, a_w = 2^{k-2}\, n(n+1) \\
&\sum_{w=0}^{n} w^3\, a_w = 2^{k-3}\, (n^2(n+3)) \\
&\sum_{w=0}^{n} w^4\, a_w = 2^{k-4}\, (n(n+1)\,(n^2+5n-2) + 4!\, b_4). \qquad (12)
\end{aligned}$$

In the next theorem, we treat linear binary codes C of length 2^m and dimension $m+2$. Note that we will focus later on the linear codes $(\mathrm{x} + R(1, m)) \cup R(1, m)$, for any $\mathrm{x} \notin R(1, m)$.

Theorem IV.1: Let m be a positive integer, $m \geq 3$. Consider any binary linear code C of length $n = 2^m$, dimension $k = m + 2$, and minimum distance δ. Let us denote by a_w (resp., b_w) the number of codewords of weight w in C (resp., $C^\perp$) and by $\mathcal{I}(\lambda)$ the number

$$\mathcal{I}(\lambda) = \sum_{w=1}^{2^{m-1}-1} (w - 2^{m-1})^2 \left((w - 2^{m-1})^2 - \lambda^2\right) a_w. \qquad (13)$$

Assume that C contains the all-one vector $\mathbf{1}$ and that $C^\perp$ is such that $b_1 = b_2 = b_3 = 0$. Then, for any positive integer $\lambda \leq 2^{m-1}$, we have

$$\mathcal{I}(\lambda) = 2^m \left(3b_4 - 2^{m-2}\left((2^{m-1}-1)^2 + (\lambda^2 - 2^{m-1})\right)\right). \qquad (14)$$

If $\delta \geq 2^{m-1} - \lambda$ then $\mathcal{I}(\lambda) \leq 0$ which can be expressed as

$$b_4 \leq \tfrac{1}{3}\, 2^{m-2} \left((2^{m-1}-1)^2 - 2^{m-1} + \lambda^2\right). \qquad (15)$$

Equality holds in (15) if and only if $\delta = 2^{m-1} - \lambda$ and if the weight distribution of C is: $a_0 = a_{2^m} = 1$ and we get the expression shown at the bottom of the page for the other nonzero a_w's. Since $b_4 \neq 0$, the minimum distance of $C^\perp$ is exactly 4.

Proof: The proof is based on the study of the numbers

$$I_\ell = \sum_{w=1}^{n-1} (w - 2^{m-1})^\ell a_w.$$

We are particularly interested in I_2 and I_4; according to (12), we have

$$\begin{aligned}
I_2 &= \sum_{w=0}^{n} (w - 2^{m-1})^2 a_w - 2^{2m-1} \\
&= 2^{k-2} n(n+1) - 2^m 2^{k-1} n + 2^{2m-2} 2^k - 2^{2m-1}
\end{aligned}$$

which gives, replacing k by $m+2$ and n by 2^m

$$I_2 = 2^{2m}(2^m+1) - 2^m 2^{2m+1} + 2^{2m-2}(2^{m+2}-2) = 2^{2m-1}. \qquad (16)$$

In the same way, we obtain

$$\begin{aligned}
I_4 &= \sum_{w=0}^{n} (w - 2^{m-1})^4 a_w - 2^{4m-3} \\
&= 2^{k-4}(n(n+1)(n^2+5n-2) + 4! b_4) \\
&\quad - 2^{m+k-2}(n^2(n+3)) + 3.2^{2m+k-3} n(n+1) \\
&\quad - 2^{3m+k-2} n + 2^{4m-4} 2^k - 2^{4m-3}
\end{aligned}$$

w	δ	2^{m-1}	$2^m - \delta$
a_w	$\dfrac{2^{2m-2}}{(\delta - 2^{m-1})^2}$	$2^{m+2} - \dfrac{2^{2m-1}}{(\delta - 2^{m-1})^2} - 2$	$\dfrac{2^{2m-2}}{(\delta - 2^{m-1})^2}$

which finally gives, replacing k by $m+2$ and n by 2^m

$$I_4 = 2^{m-2}\left(3.2^{2m} - 2^{m+1} - 2^{3m-1} + 4!\, b_4\right). \tag{17}$$

Since the codeword $\mathbf{1}$ belongs to C, we have $a_w = a_{n-w}$, for all $0 \le w \le n$, and thus,

$$I_\ell = \sum_{w=1}^{2^{m-1}-1} \left((w-2^{m-1})^\ell + (-1)^\ell (w-2^{m-1})^\ell\right) a_w.$$

Then

$$I_\ell = \begin{cases} 0, & \text{for odd } \ell \\ 2\sum\limits_{w=1}^{2^{m-1}-1} (w-2^{m-1})^\ell a_w, & \text{for even } \ell. \end{cases} \tag{18}$$

Thus, we have $I_4 \ge 0$, and from (17) we deduce

$$\begin{aligned} 0 &\le 3.2^{2m-3} - 2^{m-2} - 2^{3m-4} + 3.b_4 \\ &\le 2^{m-2}(3.2^{m-1} - 1 - 2^{2m-2}) + 3.b_4. \end{aligned}$$

Thus, b_4 must satisfy

$$b_4 \ge \frac{1}{3}\, 2^{m-2}\left((2^{m-1}-1)^2 - 2^{m-1}\right)$$

which implies $b_4 > 0$—i.e., the minimum distance of $C^\perp$ is 4.

Now we compute $I_4 - \lambda^2 I_2$. On one hand, we have by (18)

$$I_4 - \lambda^2 I_2 = 2\mathcal{I}(\lambda). \tag{19}$$

On the other hand, we express $I_4 - \lambda^2 I_2$ by means of (16) and (17). Therefore, we deduce from (19)

$$\begin{aligned} \mathcal{I}(\lambda) &= 3.2^{3m-3} - 2^{2m-2} - 2^{4m-4} + 3.b_4 2^m - 2^{2m-2}\lambda^2 \\ &= 2^m\left(2^{m-2}(3.2^{m-1} - 1 - 2^{2m-2} - \lambda^2) + 3.b_4\right) \\ &= 2^m\left(3.b_4 - 2^{m-2}((2^{m-1}-1)^2 - 2^{m-1} + \lambda^2)\right). \end{aligned}$$

Equation (13) implies that the quantity $\mathcal{I}(\lambda)$ consists of a sum of terms T_w, $1 \le w \le 2^{m-1}-1$, with $T_w \le 0$ for every w such that $(w-2^{m-1})^2 \le \lambda^2$. If $\lambda \ge 2^{m-1} - \delta$ then $T_w \le 0$ for every nonzero weight w, since $2^{m-1} - w \le 2^{m-1} - \delta$. Thus, if $\lambda \ge 2^{m-1} - \delta$ then $\mathcal{I}(\lambda) \le 0$, which exactly corresponds to the inequality (15).

Moreover, equality holds in (15) if and only if the values of w such that $a_w \ne 0$ lie in $\{2^{m-1}, 2^{m-1} \pm \lambda\}$ (i.e., $\mathcal{I}(\lambda) = 0$). Then $\lambda = 2^{m-1} - \delta$. We obtain the values a_δ by computing I_2 by means of (16) and (18). □

We now come back to the code $C_{\mathrm{x}} = (\mathrm{x} + R(1, m)) \cup R(1, m)$, $\mathrm{x} \notin R(1, m)$. Note that such a code satisfies the hypothesis of the previous theorem. Indeed, the code contains the all-one vector and, denoting by a_w (resp., b_w), $w \in [0, n]$, the number of codewords of C_{x} (resp., $C_{\mathrm{x}}^\perp$) of weight w, we have the following proposition.

Proposition IV.1: The code $C_{\mathrm{x}}^\perp$ is contained in $R(m-2, m)$; thus, we have $b_1 = b_2 = b_3 = 0$. The codewords of C_{x} which have weight 4 are the indicators of two-dimensional affine subspaces of $\boldsymbol{F}_2^m$.

Proof: This result comes from well-known properties of Reed–Muller codes: $R(1, m)^\perp = R(m-2, m)$ is the extended Hamming code and has minimum weight 4. The codewords of weight 4 have the form

$$\mathrm{y} = X^a + X^{a+b} + X^{a+c} + X^{a+b+c}, \quad \text{with } a, b, c \in \boldsymbol{F}_2^m,\ b \ne c \ne 0. \tag{20}$$

Their supports are two-dimensional affine subspaces (see [22, Ch. 13, Theorems 4 and 5]). Since $R(1, m) \subset C_{\mathrm{x}}$, then $C_{\mathrm{x}}^\perp \subset R(m-2, m)$, completing the proof. □

We focus here on the weight enumerators of cosets of $R(1, m)$ whose minimum weights are near the optimal value. Two values of λ are of most interest: $2^{m/2}$ for m even and $2^{(m-1)/2}$ for m odd, corresponding to the following kinds of cosets.

Definition IV.1: A coset of $R(1, m)$ is said to be *almost-optimal* if its minimum weight is greater than or equal to w_0, where $w_0 = 2^{m-1} - 2^{(m-1)/2}$ for odd m, and $w_0 = 2^{m-1} - 2^{m/2}$ for even m. It is said to be *three-valued* when it has exactly three nonzero weights.

Proposition IV.2: A coset of $R(1, m)$ is three-valued almost-optimal if and only if its weight distribution is

w	$2^{m-1} - 2^{(m-1)/2}$	2^{m-1}	$2^{m-1} + 2^{(m-1)/2}$
a_w	2^{m-1}	2^m	2^{m-1}

for odd m and

w	$2^{m-1} - 2^{m/2}$	2^{m-1}	$2^{m-1} + 2^{m/2}$
a_w	2^{m-2}	3.2^{m-1}	2^{m-2}

for even m.

Proof: Suppose that a coset $\mathrm{x}+R(1, m)$ has three weights only. Clearly, these weights lie in $\{2^{m-1}, \delta, 2^m - \delta\}$. Combining (18) and (16), we obtain

$$\sum_{w=\delta}^{2^{m-1}-1} (w-2^{m-1})^2 a_w = (2^{m-1} - \delta)^2 a_\delta = 2^{2m-2}.$$

Thus, $2^{m-1} - \delta$ is a power of 2. Assume that $\mathrm{x} + R(1, m)$ is almost-optimal. Then the only possibility for m odd is $\delta = 2^{m-1} - 2^{(m-1)/2}$. When m is even, the only possibility for the coset to have exactly three weights is $\delta = 2^{m-1} - 2^{m/2}$. □

Consider the notation of Theorem IV.1. By replacing C by C_{x} we obtain the following necessary condition on three-valued almost-optimal cosets.

Corollary IV.1: If the coset $\mathrm{x}+R(1, m)$ is almost-optimal, then we have

- if m is odd, then $b_4 \le \frac{1}{3}\, 2^{m-2}\,(2^{m-1}-1)^2$;
- if m is even, then $b_4 \le \frac{1}{3}\,(2^{m-2}\,(2^{m-1}-1)^2 + 2^{2m-3})$.

In both cases, equality holds if and only if $\mathrm{x}+R(1, m)$ is three-valued almost-optimal.

Proof: We simply apply Theorem IV.1.

- If m is odd, we set $\lambda = 2^{(m-1)/2}$. As the coset is almost-optimal, $\delta \ge 2^{m-1} - 2^{(m-1)/2}$.

- If m is even, we set $\lambda = 2^{m/2}$. As the coset is almost-optimal, $\delta \geq 2^{m-1} - 2^{m/2}$(where δ is the minimum distance of C_{x}—i.e., the minimum weight of the coset). □

Remark IV.1:

1) By taking $\lambda = 2^{(m-2)/2}$ with m even, we obtain cosets whose minimum weight is $\rho(R(1, m))$ only. These cosets have two weights, $2^{m-1} \pm 2^{(m-2)/2}$, and correspond to the bent functions. Moreover,
$$b_4 = \tfrac{1}{3}\, 2^{m-2} \left((2^{m-1} - 1)^2 - 2^{m-2} \right).$$
2) It is quite easy to construct three-valued almost-optimal cosets. Although these cosets are not yet classified, they are completely known when $\mathrm{x} \in R(2, m)$ (see [22, Ch. 14] and a short presentation in Appendix I). We will give other examples in the next section.
3) Assume that $\mathrm{x} + R(1, m)$ is almost-optimal and that the upper bound on b_4 is not reached. Very little is known about these cosets and several hard open problems are involved, as the covering radius of $R(1, m)$ for m odd, or the covering radius of $R(1, m)$ restricted to codewords of weight 2^{m-1} for any m. Examples of such cosets can be found in [23] for $m = 5$ and in [12] for $m < 11$.

Open Problem IV.1: The dual of the code C_{x} is a subspace of $R(m-2, m)$ of codimension 1. How can a subspace containing *few codewords of weight* 4 be constructed?

B. Computing b_4

Let us denote by $\mathcal{W}$ the set of all affine subspaces of $\boldsymbol{F}_2^m$ of dimension 2 and by $\mathcal{W}_0$ the subset of $\mathcal{W}$ of all *linear subspaces* of $\boldsymbol{F}_2^m$. Recall that $C_{\mathrm{x}} = (\mathrm{x} + R(1, m)) \cup R(1, m)$, $\mathrm{x} \notin R(1, m)$, and that b_4 is the number of codewords of weight 4 in $C_{\mathrm{x}}^{\perp}$. In this subsection, we want to be more explicit about the computation of b_4. We later apply our results to the cosets which are contained in $R(3, m)$.

The codewords of weight 4 in $C_{\mathrm{x}}^{\perp}$ are of type (20). These codewords have as support an element V of $\mathcal{W}$; they belong to $R(m-2, m)$ (i.e., $\mathcal{P}^2$). In this section, we will denote by y^V such a codeword. Let $a, b, c \in \boldsymbol{F}_2^m$, $b \neq c \neq 0$, and $V = \{a, a+b, a+c, a+b+c\}$. Then
$$\mathrm{y}^V = \sum_{v \in V} X^v = X^a (X^b + 1)(X^c + 1). \tag{21}$$
The next result is a direct application of Lemma III.1.

Proposition IV.3: For any $V \in \mathcal{W}_0$ we denote by x_i, $1 \leq i \leq 2^{m-2}$, the restrictions of x to the cosets of V. Then the number b_4 of codewords of weight 4 in $C_{\mathrm{x}}^{\perp}$ can be expressed as follows:
$$b_4 = \sum_{V \in \mathcal{W}_0} \frac{2^m - \mathrm{wt}\,(\mathrm{y}^V \mathrm{x})}{4}$$
where the codeword y^V is defined by (21). Moreover,
$$\mathrm{wt}\,(\mathrm{y}^V \mathrm{x}) = 4 \times \#\{i \,|\, \mathrm{wt}\,(\mathrm{x}_i) \text{ is odd}\}.$$

Proof: Let $V \in \mathcal{W}_0$. Since $\mathrm{y}^V \in R(m-2, m)$, $\mathrm{y}^V \in C_{\mathrm{x}}^{\perp}$ if and only if y^V is orthogonal to x—i.e., the weight of the restriction x_1 of x to V is even.

In accordance with Lemma III.1, we have
$$\mathrm{wt}\,(\mathrm{y}^V \mathrm{x}) = \#\{i \,|\, \mathrm{wt}\,(\mathrm{x}_i) \text{ is odd}\} \times 4$$
where the x_i's are the restrictions of x to the cosets of V. This implies that the number of cosets $h + V$ such that $\mathrm{y}^{h+V} \in C_{\mathrm{x}}^{\perp}$ is equal to $(2^m - \mathrm{wt}\,(\mathrm{y}^V \mathrm{x}))/4$. The value of b_4 is obtained by considering all $V \in \mathcal{W}_0$. □

When there are few possible values for $\mathrm{wt}\,(\mathrm{y}^V \mathrm{x})$, the expression of b_4 becomes simpler. It is especially the case when x is in $R(3, m)$.

Corollary IV.2: Let x be in $R(3, m) \backslash R(1, m)$. Let us define

- $N_0 = \#\{V \in \mathcal{W}_0 | \mathrm{y}^V \mathrm{x} = \mathbf{0}\}$
- $N_b = \#\{V \in \mathcal{W}_0 | \mathrm{wt}\,(\mathrm{y}^V \mathrm{x}) = 2^{m-1}\}$.

Then $b_4 = 2^{m-2} N_0 + 2^{m-3} N_b$. We have $N_b = 0$ when $\mathrm{x} \in R(2, m)$.

Proof: As $\mathrm{x} \in R(3, m)$ and $\mathrm{y}^V \in R(m-2, m)$, $\mathrm{y}^V \mathrm{x}$ is in $R(1, m)$—since
$$\mathcal{P}^2 \mathcal{P}^{m-3} = \mathcal{P}^{m-1}.$$
So $\mathrm{wt}\,(\mathrm{y}^V \mathrm{x})$ belongs to $\{0, 2^{m-1}, 2^m\}$. When $\mathrm{x} \in R(2, m)$, we have $\mathrm{y}^V \mathrm{x} \in \{\mathbf{0}, \mathbf{1}\}$ implying $N_b = 0$.

In accordance with Proposition IV.3, we obtain
$$b_4 = \sum_{\substack{V \in \mathcal{W}_0 \\ \mathrm{y}^V \mathrm{x} = 0}} 2^{m-2} + \sum_{\substack{V \in \mathcal{W}_0 \\ \mathrm{wt}\,(\mathrm{y}^V \mathrm{x}) = 2^{m-1}}} 2^{m-3} = N_0 2^{m-2} + N_b 2^{m-3}.$$
□

Remark IV.2: Note that the weight enumerators of the cosets $\mathrm{x} + R(1, m)$ with $\mathrm{x} \in R(2, m)$ are known. For cosets which are not contained in $R(2, m)$, the weight enumerators are generally not known. The study of such cosets contained in $R(3, m)$ is the *first* open problem. In this paper, we point out that these cosets have specific properties. However, it seems difficult to strengthen any conjecture.

Corollary IV.3: Let $R_{\mathrm{x}} = \mathrm{x} + R(1, m)$ with $\mathrm{x} \in R(3, m) \setminus R(1, m)$. Let
$$N_1 = \#\{V \in \mathcal{W}_0 \,|\, \mathrm{y}^V \mathrm{x} = \mathbf{1}\}$$
and N_0 has been defined in the previous corollary. If R_{x} is almost-optimal, we have

- if m is odd, then $N_1 - N_0 \geq \frac{2^{m-1}-1}{3}$.
- if m is even, then $N_0 - N_1 \leq \frac{2^{m-1}+1}{3}$.

In both cases, equality holds if and only if R_{x} is three-valued almost-optimal.

Proof: Recall that
$$\#\mathcal{W}_0 = \tfrac{1}{3}(2^m - 1)(2^{m-1} - 1) = N_0 + N_1 + N_b \tag{22}$$
implying
$$\begin{aligned} b_4 &= 2^{m-2} N_0 + 2^{m-3} N_b \\ &= 2^{m-3} \left(N_0 - N_1 + \frac{(2^m - 1)(2^{m-1} - 1)}{3} \right). \end{aligned}$$
Suppose that R_{x} is almost-optimal. According to Corollary IV.1 we obtain the expected bounds. □

For m even, it is easy to find cosets, defined as above, satisfying $N_0 - N_1 < (2^{m-1}+1)/3$ (see the next example). In the case where m is odd it is not so easy to find cosets satisfying $N_1 - N_0 > (2^{m-1}-1)/3$. Actually, the existence of such cosets is just proved by Canteaut in [24]; she exhibits almost-optimal cosets with five weights which are contained in $R(3, 9)$. These weights are $2^8 \pm 2^4$, $2^8 \pm 2^3$ and 2^8. However, the determination of the minimum weights of such cosets remains an open problem for $m \leq 13$ (see the end of Section V-A for more explanations).

Example IV.1: Let $m = 6$ and

$$f(x_1, \ldots, x_6) = x_1x_2x_3 + x_1x_2x_4 + x_1x_2 + x_3x_4 + x_5x_6.$$

The weight distribution of the coset $\Omega_f + R(1, 6)$ is

$$a_{24} = a_{40} = 8, \quad a_{28} = a_{36} = 32, \quad \text{and} \quad a_{32} = 48.$$

This coset is almost-optimal with five weights.

V. The Propagation Criterion and the Nonlinearity

We come back to the terminology of Boolean functions but we will always consider together a given function f of m variables and its associated binary codeword Ω_f. So we first fix the terminology for functions which generate a coset $\Omega_f + R(1, m)$ with a high minimum weight (see Definition IV.1 and the following proposition).

Definition V.1: The Boolean function f is said to be almost-optimal if its associated coset $\Omega_f + R(1, m)$ is almost-optimal or equivalently if

- $\mathcal{L}(f) \leq 2^{(m+2)/2}$, when m is even;
- $\mathcal{L}(f) \leq 2^{(m+1)/2}$, when m is odd.

The function f is said to be three-valued almost-optimal if its associated coset is three-valued almost-optimal—i.e., its Fourier spectrum is $\{0, \pm 2^{(m+2)/2}\}$ when m is even and $\{0, \pm 2^{(m+1)/2}\}$ when m is odd.

Recall the definition of the so-called $PC(\ell)$ property.

Definition V.2: Let $\boldsymbol{e} = (e_1, \ldots, e_m)$ be a basis of $\boldsymbol{F}_2^m$. Then f satisfies the *propagation criterion of order* ℓ (PC (ℓ)), *with respect to* $\boldsymbol{e}$ if, for any vector $a = (a_1, \ldots, a_m)$ in $\boldsymbol{F}_2^m$ such that $0 < \mathrm{wt}\,(a) \leq k$

$$D_a f, \; a = \sum_{i=1}^{m} a_i e_i$$

is balanced.

A. Bounds on the Sum-of-Squares Indicator

From now on, we focus on almost-optimal functions $f \in \mathcal{B}_m$, $f \notin R(1, m)$, $m \geq 3$. Notation is the same as in Theorem IV.1 and its proof: we consider the code

$$C_{\mathrm{x}} = (\mathrm{x} + R(1, m)) \cup R(1, m)$$

with $\mathrm{x} = \Omega_f$; a_w denotes the number of codewords of weight w in C_{x}, and $\mathcal{I}(\lambda)$ is defined by (13). Recall that the sum-of-squares indicator $\mathcal{V}(f)$ allows to measure the global avalanche criterion of f (see Section II-A, (6)). The next propositions are in fact corollaries of Theorem IV.1; our aim is to make explicit the link between two points of view (in terms of codewords and in terms of functions).

Lemma V.1: Let $\lambda \leq 2^{m-1}$ be any positive integer. Then

$$\mathcal{I}(\lambda) = 2^{m-4}(\mathcal{V}(f) - \lambda^2 2^{m+2}).$$

Thus, $\mathcal{I}(\lambda) \leq 0$ if and only if $\mathcal{V}(f) \leq \lambda^2 2^{m+2}$.

Proof: From (19), we have

$$\mathcal{I}(\lambda) = \sum_{w=\delta}^{2^{m-1}-1} (w - 2^{m-1})^2[(w - 2^{m-1})^2 - \lambda^2]a_w$$
$$= \frac{1}{2}\left[I_4 - \lambda^2 I_2\right].$$

According to (18) we obtain

$$I_4 = 2\sum_{w=1}^{2^{m-1}-1} (w - 2^{m-1})^4 a_w = \frac{1}{8}\sum_{\alpha \in \boldsymbol{F}_2^m} \mathcal{F}^4(f + \varphi_\alpha) \quad (23)$$

since $\mathcal{F}^4(f + \varphi_\alpha) = (2^m - 2w)^4$ where w is the weight of $\Omega_{f+\varphi_\alpha}$, and

$$a_w = |\{\alpha \in \boldsymbol{F}_2^m \mid \mathrm{wt}\,(\Omega_{f+\varphi_\alpha}) = w \text{ or } 2^m - w\}|.$$

Moreover, $I_2 = 2^{2m-1}$.

It follows that

$$\mathcal{I}(\lambda) = \frac{1}{16}\left[\sum_{\alpha \in \boldsymbol{F}_2^m} \mathcal{F}^4(f + \varphi_\alpha) - \lambda^2 2^{2m+2}\right].$$

Using (7), we deduce that

$$\mathcal{I}(\lambda) = 2^{m-4}\left[\mathcal{V}(f) - \lambda^2 2^{m+2}\right]$$

completing the proof. □

Proposition V.1: Let $f \in \mathcal{B}_m$ and $\mathrm{x} = \Omega_f$. Let b_w denote the number of codewords of weight w in $C_{\mathrm{x}}^{\perp}$. Then

$$b_4 = \frac{1}{48}\left(\mathcal{V}(f) + 2^{m+2}\left((2^{m-1}-1)^2 - 2^{m-1}\right)\right).$$

Proof: We simply consider together the formula given in the previous lemma and (14). So

$$\mathcal{V}(f) - \lambda^2 2^{m+2}$$
$$= 2^4\left(3b_4 - 2^{m-2}\left((2^{m-1}-1)^2 + (\lambda^2 - 2^{m-1})\right)\right).$$

Hence,

$$\mathcal{V}(f) = 2^4\left(3b_4 - 2^{m-2}\left((2^{m-1}-1)^2 - 2^{m-1}\right)\right)$$

completing the proof. □

Proposition V.2: Let m be a positive integer, $m \geq 3$, and $f \in \mathcal{B}_m$. Assume that f is almost-optimal. Then

- if m is odd then $\mathcal{V}(f) \leq 2^{2m+1}$ with equality if and only if f is three-valued almost-optimal;
- if m is even then $\mathcal{V}(f) \leq 2^{2m+2}$ with equality if and only if f is three-valued almost-optimal.

Proof: Since f is almost-optimal, the minimum weight δ of the coset $\Omega_f + R(1, m)$ satisfies $\delta \geq 2^{m-1} - 2^{(m-1)/2}$ for odd m and $\delta \geq 2^{m-1} - 2^{m/2}$ for even m. According to

Theorem IV.1, this implies $\mathcal{I}(2^{(m-1)/2}) \leq 0$ for odd m and $\mathcal{I}(2^{m/2}) \leq 0$ for even m.

From Lemma V.1, replacing λ by either $2^{(m-1)/2}$ or $2^{m/2}$ (depending on whether m is odd or even), we immediately deduce the expected bounds on $\mathcal{V}(f)$. □

Example V.1: There are many three-valued almost-optimal functions. The *almost-bent* functions provide such functions (see, for instance, [25]–[28]). Any three-valued almost-optimal partially bent function is linearly equivalent to (see Proposition II.6)

$$f(x_1, \ldots, x_m) = g(x_1, \ldots, x_{m-\ell}) + \varphi_\alpha(x_1, \ldots, x_m)$$

where g is bent and $\ell = 1$ for odd m and $\ell = 2$ for even m. For these functions, $\mathcal{L}(f) = 2^{(m+\ell)/2}$ and $\mathcal{V}(f) = 2^{2m+\ell}$.

It is easy to find almost-optimal functions such that $\mathcal{L}(f) = 2^{(m+\ell)/2}$ and $\mathcal{V}(f) < 2^{2m+\ell}$, where ℓ is defined as above. These functions have a good (generally not the best) nonlinearity but are not three-valued (see Example IV.1, a number of numerical results in [29], [12] and Proposition V.5).

It is not so easy to obtain almost-optimal functions such that $\mathcal{L}(f) < 2^{(m+\ell)/2}$ (implying $\mathcal{V}(f) < 2^{2m+\ell}$ according to Proposition V.2). The class of bent functions seems to be the only known large class. Numerical results are easily obtained for m even (see [30], [12]). When m is odd, the only known such functions are obtained from those given in [31] for $m = 15$.

Note that there exist non-almost-optimal functions f such that $\mathcal{V}(f) \leq 2^{2m+\ell}$.

Example V.2: For $m = 5$ one finds in [23] the function

$$f(x_1 \ldots, x_5) = x_1x_2x_3x_4x_5 + x_1x_2x_3 + x_1x_4x_5 + x_4x_5 + x_3x_5 + x_2x_4 + x_2x_3.$$

It generates a coset of $R(1, m)$ with weight distribution

$$a_{11} = a_{21} = 4, \quad a_{13} = a_{19} = 16, \quad \text{and} \quad a_{15} = a_{17} = 12$$

$a_w = 0$ otherwise. Thus, $\mathcal{L}(f) = 10$. Using (7), we obtain $\mathcal{V}(f) = 1904$ which is strictly less than $2^{11} = 2048$.

Let $f \in \mathcal{B}_m$ be a function of degree d. Set the notation

$$E_f = \{e \in \boldsymbol{F}_2^m \mid D_e f \text{ is balanced}\} \quad \text{and} \quad \overline{E_f} = \boldsymbol{F}_2^m \setminus E_f.$$

In [10, Proposition 14], we have stated the following relation between the cardinality of $\overline{E_f}$, denoted by $\#\overline{E_f}$, and the value of $\mathcal{V}(f)$.

Proposition V.3: Let $f \in \mathcal{B}_m$ be a function of degree d. Then

$$\mathcal{V}(f) \geq 2^{2m} + \left(\#\overline{E_f} - 1\right) 2^{2\lfloor \frac{m-2}{d-1} \rfloor + 4}.$$

This is of most interest for functions of degree 3. We obviously obtain from the previous result and from Proposition V.2 the following corollary.

Corollary V.1: Let $f \in \mathcal{B}_m$ a function of degree 3. So the following properties hold.

i) When m is even then $\mathcal{V}(f) \geq 2^{2m} + (\#\overline{E_f} - 1)2^{m+2}$. Thus, if $\mathcal{L}(f) \leq 2^{(m+2)/2}$ then $\#E_f \geq 2^{m-2} - 1$.

ii) When m is odd then $\mathcal{V}(f) \geq 2^{2m} + (\#\overline{E_f} - 1)2^{m+1}$. Thus, if $\mathcal{L}(f) \leq 2^{(m+1)/2}$ then $\#E_f \geq 2^{m-1} - 1$.

Therefore, we point out that for almost-optimal functions of degree 3, the rank of E_f must be *high*. Note that we call *rank* of E_f the dimension of the subspace generated by the elements of E_f (remark that $E_f \cup \{0\}$ is not, in general, a subspace).

Corollary V.2: An almost-optimal function of degree 3 is such that the rank of E_f is at least $m - 2$ for even m and at least $m - 1$ for odd m.

When m is odd, such a function is PC(1), unless $E_f \cup \{0\}$ is a subspace of codimension 1. In this case, f is three-valued almost-optimal.

Proof: If E_f is a set of rank k then its cardinality is at most $2^k - 1$ (E_f does not contain 0). Assume that f is almost-optimal. Clearly, Corollary V.1 provides the lower bounds $m-2$ (m even) and $m - 1$ (m odd) for k. When m is odd, k is either $m - 1$ or m.

Assume that m is odd. If $k = m - 1$ then $\#E_f = 2^{m-1} - 1$ (since $E_f \cup \{0\}$ is a subspace of codimension 1), implying $\mathcal{V}(f) = 2^{2m+1}$ thanks to Corollary V.1 and Proposition V.2. In accordance with Proposition V.2, f is three-valued almost-optimal; note that it can be proved by another way, using Theorem V.2 of Section V-C.

When $k = m$, it means that there exists a basis of $\boldsymbol{F}_2^m$, say $\boldsymbol{e} = (e_1, \ldots, e_m)$, such that $e_i \in E_f$ and $D_{e_i} f$ is balanced, for all i; so f is PC(1), with respect to $\boldsymbol{e}$. □

Note that it is very easy to construct almost-optimal functions of degree 3, which are three-valued. It is more difficult to construct such functions which are almost-optimal and not three-valued, especially when m is odd—as we indicated in other terms at the end of Section IV-B. Moreover, the general problem of the maximal nonlinearity of functions of degree 3 remains open for odd m.

It is known that, for any odd $m \leq 13$, all almost-optimal functions f of degree 3 satisfy $\mathcal{L}(f) = 2^{(m+1)/2}$ [32]. It has been recently proved by Canteaut that, for any odd $m \leq 7$, all almost-optimal functions of degree 3 are three-valued. For $m = 9$, she has proved that there is only one weight polynomial for almost-optimal non-three-valued cosets of $R(1, m)$ which are contained in $R(3, 9)$; moreover, she proves that such cosets exist [24].

Open Problem V.1: For odd m, $m > 13$, does there exist $f \in \mathcal{B}_m$ of degree 3 such that f is almost-optimal and $\mathcal{L}(f) < 2^{(m+1)/2}$?

B. Decompositions on Affine Subspaces of $\boldsymbol{F}_2^m$

We are going to study the restrictions of $f \in \mathcal{B}_m$ to any subspace W of $\boldsymbol{F}_2^m$. Lemma V.2 is derived from well-known properties of the Fourier transform.

Lemma V.2: Let f be a Boolean function of m variables and let V be a subspace of $\boldsymbol{F}_2^m$ of dimension k. Then we have, for any $\beta \in \boldsymbol{F}_2^m$

$$\sum_{\alpha \in V} \mathcal{F}^2(f + \varphi_{\alpha+\beta}) = 2^k \sum_{e \in V^\perp} (-1)^{\beta \cdot e} \mathcal{F}(D_e f).$$

Proof: According to (3), we have for any $\alpha \in \boldsymbol{F}_2^m$

$$\mathcal{F}^2(f+\varphi_\alpha) = \sum_{e\in\boldsymbol{F}_2^m} (-1)^{\alpha\cdot e}\mathcal{F}(D_e f).$$

We deduce that, for any $\beta \in \boldsymbol{F}_2^m$

$$\begin{aligned}\sum_{\alpha\in V} \mathcal{F}^2(f+\varphi_{\alpha+\beta}) &= \sum_{\alpha\in V}\sum_{e\in\boldsymbol{F}_2^m} (-1)^{(\alpha+\beta)\cdot e}\mathcal{F}(D_e f)\\ &= \sum_{e\in\boldsymbol{F}_2^m} (-1)^{\beta\cdot e}\mathcal{F}(D_e f)\sum_{\alpha\in V}(-1)^{\alpha\cdot e}\\ &= 2^k \sum_{e\in V^\perp} (-1)^{\beta\cdot e}\mathcal{F}(D_e f). \qquad \square\end{aligned}$$

Remark V.1: Note that for $V = \boldsymbol{F}_2^m$, Lemma V.2 provides the well-known formula of Parseval. When $V^\perp = \{0, a\}$, we get the following relation:

$$\sum_{\alpha\in V} \mathcal{F}^2(f+\varphi_{\alpha+\beta}) = 2^{m-1}\left(2^m + (-1)^{a\cdot\beta}\mathcal{F}(D_a f)\right),$$

for all $\beta \in \boldsymbol{F}_2^m$.

We need to define precisely the restrictions of any $f \in \mathcal{B}_m$ to a subspace W, of dimension k, and to the cosets of W. Let such a coset $W' = a + W$, $a \notin W$. The restriction of f to W' can be identified with $h \in \mathcal{B}_k$ such that $h(x) = f(a+x)$. This representation depends, in fact, on the choice of $a \in W'$ since for $b = a + u$, $u \in W$, we have $h'(x) = f(b+x) = h(u+x)$ (h' is a *translation* of h). However, in the context of our study, h and h' have the same properties. So when we say *the decomposition of* f (as defined below) we mean that, for a fixed W, the restrictions are chosen up to translations.

Definition V.3: Let W be a subspace of $\boldsymbol{F}_2^m$ of dimension k. *The decomposition of* f *with respect to* W is the sequence $\{h_a | a \in V\}$ where V is such that $\boldsymbol{F}_2^m$ is the direct sum of W and V and h_a is the Boolean function of k variables, from W to $\boldsymbol{F}_2$, defined by $h_a(x) = f(a+x)$ for any $x \in W$.

Theorem V.1: Let W be a subspace of $\boldsymbol{F}_2^m$ of dimension k and let $(h_a | a \in V)$ be the decomposition of f with respect to W. Then

$$\sum_{\alpha\in W^\perp} \mathcal{F}^2(f+\varphi_\alpha) = 2^{m-k}\sum_{a\in V}\mathcal{F}^2(h_a).$$

Proof: Consider Ω_f the associated codeword of f. We have

$$\Omega_f = \sum_{a\in V} X^a \Omega_{h_a}. \tag{24}$$

We obviously deduce

$$\mathrm{wt}(\Omega_f) = \sum_{a\in V}\mathrm{wt}(\Omega_{h_a}).$$

Note the extension of this property to $D_\beta f$, for any $\beta \in W$. Indeed, we have for such a β

$$\Omega_{D_\beta f} = \sum_{a\in V} X^a \Omega_{D_\beta h_a}.$$

Thus,

$$\begin{aligned}\mathcal{F}(D_\beta f) = 2^m - 2\mathrm{wt}(\Omega_{D_\beta f}) &= \sum_{a\in V}\left(2^{m-k} - 2\mathrm{wt}(\Omega_{D_\beta h_a})\right)\\ &= \sum_{a\in V}\mathcal{F}(D_\beta h_a).\end{aligned}$$

Set $L = \sum_{\alpha\in W^\perp}\mathcal{F}^2(f+\varphi_\alpha)$. According to Lemma V.2 and to the above formula, we have

$$\begin{aligned}L &= 2^{m-k}\sum_{\beta\in W}\mathcal{F}(D_\beta f) = 2^{m-k}\sum_{\beta\in W}\left(\sum_{a\in V}\mathcal{F}(D_\beta h_a)\right)\\ &= 2^{m-k}\sum_{a\in V}\left(\sum_{\beta\in W}\mathcal{F}(D_\beta h_a)\right) = 2^{m-k}\sum_{a\in V}\mathcal{F}^2(h_a)\end{aligned}$$

according to (4). $\square$

Corollary V.3: Let $(h_a | a \in V)$ be the decomposition of f with respect to the k-dimensional subspace W. Then

$$\sum_{a\in V}\mathcal{F}^2(h_a) \le \mathcal{L}^2(f).$$

Moreover, $\mathcal{L}(h_a) \le \mathcal{L}(f)$, for all $a \in V$.

Proof: According to Theorem V.1 and since $|W^\perp| = 2^{m-k}$, we obviously deduce

$$2^{m-k}\sum_{a\in V}\mathcal{F}^2(h_a) \le 2^{m-k}\mathcal{L}^2(f)$$

implying $\mathcal{F}^2(h_a) \le \mathcal{L}^2(f)$ for every a. Moreover, this property holds if we replace h_a by $h_a + \ell$, where ℓ is any linear function of $\mathcal{B}_k$—considering the decomposition of $f + \varphi_\beta$, for some $\beta \in \boldsymbol{F}_2^m$, instead of f. Hence, $\mathcal{L}(h_a) \le \mathcal{L}(f)$ for all a, completing the proof. $\square$

Remark V.2: We have

$$\begin{aligned}&\mathcal{N}(f) - \mathcal{N}(h_a)\\ &\quad = 2^{m-1} - 2^{k-1} + (\mathcal{L}(h_a) - \mathcal{L}(f))/2 \le 2^{m-1} - 2^{k-1}\end{aligned}$$

since $\mathcal{L}(h_a) - \mathcal{L}(f) \le 0$. This upper bound on $\mathcal{N}(f) - \mathcal{N}(h_a)$ was already proved by Zheng *et al.* in [8]. The authors noticed that when m is odd, $k = (m+1)/2$ and h_a is an affine function, then $\mathcal{N}(f) \le 2^{m-1} - 2^{(m-1)/2}$.

Notice that, when m is even, $k = m/2$ and h_a affine, we find again the covering radius of $R(1, m)$.

The previous results provide the exact connection between the nonlinearity of f and the nonlinearity of each element of any decomposition of f—"any" means "with respect to W, for any W." The well-known conjecture of Dobbertin has to be placed in this context. In [30], he introduced the notion of *normal* function for even m. A function $f \in \mathcal{B}_m$ is said to be normal if it is constant on at least one $m/2$-dimensional flat. He proposed the next conjecture.

Conjecture. *Any bent function is normal.*

The link between the nonlinearity of a function and the nonlinearity of each element of its decomposition has several consequences. For instance, when f is almost-optimal, any function h of any decomposition of f is such that $\mathcal{L}(h) \le 2^{(m+1)/2}$ for odd m, and $\mathcal{L}(h) \le 2^{(m+2)/2}$ for even m. This notably leads to the following property.

Proposition V.4: Assume that m is odd. Suppose that W has codimension 1. For simplicity, we denote by (h_1, h_2) the decomposition of f with respect to W.

If h_1 (or h_2) is partially bent and not bent then $\mathcal{L}(f) \geq 2^{(m+1)/2}$.

Proof: If h_1 is partially bent then $\mathcal{L}(h_1) = 2^\rho$, where ρ is an integer such that $\rho \geq (m-1)/2$, with equality if and only if h_1 is bent.

If h_1 is not bent we have, in accordance with Corollary V.3

$$\mathcal{L}(f) \geq \mathcal{L}(h_1) \geq 2^{(m+1)/2}$$

completing the proof. □

Remark V.3: The function with five variables given in Example V.6

$$g(x_1, \ldots, x_5) = x_3x_5 + x_2x_4 + x_1x_2x_3 + x_2x_3x_4x_5$$

is almost-optimal (not three-valued) of degree 4. It satisfies the hypothesis of Proposition V.4, since one element of its decomposition with respect to the hyperplane

$$\{(x_1, \ldots, x_5) \in \boldsymbol{F}_2^m, x_2 = 0\}$$

is quadratic

$$g(x_1, \ldots, x_5) = (1 + x_2)x_3x_5 + x_2(x_4 + x_3x_5 + x_1x_3 + x_3x_4x_5).$$

This proves that the class of such functions is interesting.

Example V.3: It is very easy to construct a function f satisfying the hypothesis of Proposition V.4, with algebraic normal form equal, up to equivalence, to

$$f(x_1, \ldots, x_m) = g(x_1, \ldots, x_{m-1}) + x_m h(x_1, \ldots, x_{m-1}) \tag{25}$$

where g has degree 2.

Let $m = 7$. The functions

$$f(x_1, \ldots, x_7) = x_1x_2 + x_3x_4 + x_7h(x_1, \ldots, x_6)$$

where h is any function in $\mathcal{B}_6$, satisfy $\mathcal{L}(f) \geq 2^{(7+1)/2}$. Indeed, it is well known that $\mathcal{L}(x_1x_2 + x_3x_4) = 2^4$ (see Appendix I and [22, Ch. 15, Sec. 2]).

C. Derivatives on Subspaces of Large Dimensions

Now we are considering the cases where the derivatives D_ef of a given function f are balanced for any $e \neq 0$ belonging to a subspace of codimension 1 or 2. This allows us to obtain a new characterization of bent functions and of some three-valued almost-optimal functions. We first fix notation.

Recall that φ_α, $\alpha \in \boldsymbol{F}_2^m$, denotes the linear function $x \longmapsto \alpha \cdot x$. We denote by H_α the kernel of φ_α

$$H_\alpha = \{x \in \boldsymbol{F}_2^m, \ \varphi_\alpha(x) = 0\}.$$

We denote by $\overline{H}_\alpha$ the affine subspace $\boldsymbol{F}_2^m \setminus H_\alpha$. Clearly, φ_α is the characteristic function of $\overline{H}_\alpha$.

Lemma V.3: Let $\alpha \in \boldsymbol{F}_2^m$ and φ_α the associated linear function with kernel H_α. We have

$$\mathcal{F}^2(f + \varphi_\alpha) = \sum_{e \in H_\alpha} \mathcal{F}(D_ef) - \sum_{e \in \overline{H}_\alpha} \mathcal{F}(D_ef) \tag{26}$$

$$\mathcal{F}^2(f) + \mathcal{F}^2(f + \varphi_\alpha) = 2 \sum_{e \in H_\alpha} \mathcal{F}(D_ef) \tag{27}$$

$$\mathcal{F}^2(f) - \mathcal{F}^2(f + \varphi_\alpha) = 2 \sum_{e \in \overline{H}_\alpha} \mathcal{F}(D_ef) \tag{28}$$

$$\mathcal{F}^2(f) = \mathcal{F}^2(f + \varphi_\alpha) \Leftrightarrow \sum_{e \in \overline{H}_\alpha} \mathcal{F}(D_ef) = 0 \tag{29}$$

$$\mathcal{F}^2(f)\mathcal{F}^2(f + \varphi_\alpha) = \left(\sum_{e \in H_\alpha} \mathcal{F}(D_ef) \right)^2 - \left(\sum_{e \in \overline{H}_\alpha} \mathcal{F}(D_ef) \right)^2. \tag{30}$$

Proof: Relations (3) and (4) can be rewritten

$$\mathcal{F}^2(f + \varphi_\alpha) = \sum_{e \in H_\alpha} \mathcal{F}(D_ef) - \sum_{e \in \overline{H}_\alpha} \mathcal{F}(D_ef)$$

which is exactly (26), and

$$\mathcal{F}^2(f) = \sum_{e \in H_\alpha} \mathcal{F}(D_ef) + \sum_{e \in \overline{H}_\alpha} \mathcal{F}(D_ef).$$

Formulas (27), (28), and (30) are obtained by combining the above relations. Formula (28) obviously implies (29). Note that (27) can be directly obtained from Lemma V.2 ($k = 1$). □

Lemma V.4: Let m be a positive integer, $m \geq 3$, and $f \in \mathcal{B}_m$. Define, for any $\alpha \in \boldsymbol{F}_2^m$, the property $(\mathcal{H}_\alpha)$: the function D_ef is balanced for every nonzero element e of H_α. If f satisfies $(\mathcal{H}_\alpha)$ for some α, then

$$\mathcal{F}^2(f + \varphi_\beta) + \mathcal{F}^2(f + \varphi_{\beta+\alpha}) = 2^{m+1}$$

for all $\beta \in \boldsymbol{F}_2^m$.

Proof: Since D_ef is balanced if and only if $\mathcal{F}(D_ef) = 0$, $(\mathcal{H}_\alpha)$ implies, in accordance with (27)

$$\mathcal{F}^2(f) + \mathcal{F}^2(f + \varphi_\alpha) = 2\mathcal{F}(D_0f) = 2^{m+1}.$$

Moreover, this property holds for any $f + \varphi_\beta$, since any function $D_e\varphi_\beta$ is constant, implying that $D_e(f + \varphi_\beta)$ is balanced as soon as D_ef is balanced. □

Theorem V.2: Let m be an odd integer, $m \geq 3$, $\alpha \in \boldsymbol{F}_2^m$, $\alpha \neq 0$, and $f \in \mathcal{B}_m$. Then the following properties are equivalent:

i) f satisfies $(\mathcal{H}_\alpha)$;

ii) f is three-valued almost-optimal and
$$\mathcal{F}^2(f + \varphi_\beta) \neq \mathcal{F}^2(f + \varphi_{\beta+\alpha})$$
for all $\beta \in \boldsymbol{F}_2^m$;

iii) both restrictions of f to H_α and $\overline{H}_\alpha$ are bent.

Proof: i) ⇒ ii). Lemma V.4 implies that

$$\mathcal{F}^2(f + \varphi_\beta) + \mathcal{F}^2(f + \varphi_{\beta+\alpha}) = 2^{m+1}$$

for all $\beta \in \boldsymbol{F}_2^m$. If there exists $\beta \in \boldsymbol{F}_2^m$ such that

$$\mathcal{F}^2(f + \varphi_\beta) = \mathcal{F}^2(f + \varphi_{\alpha+\beta})$$

then we obtain $\mathcal{F}^2(f + \varphi_\beta) = 2^m$ where 2^m is not a square, a contradiction.

Moreover, applying Lemma B.1 (in Appendix II), we deduce that $\mathcal{F}^2(f + \varphi_\beta) \in \{0, 2^{m+1}\}$, for all β. So f is three-valued almost-optimal.

ii) ⇒ iii). Let us denote by (h_1, h_2) the decomposition of f with respect to H_α. From Theorem V.1, we have

$$\mathcal{F}^2(f) + \mathcal{F}^2(f + \varphi_\alpha) = 2\left(\mathcal{F}^2(h_1) + \mathcal{F}^2(h_2)\right). \tag{31}$$

Since the Fourier spectrum of f is $\{0, \pm 2^{(m+1)/2}\}$ and $\mathcal{F}^2(f) \neq \mathcal{F}^2(f+\varphi_\alpha)$, we obtain $\mathcal{F}^2(h_1)+\mathcal{F}^2(h_2) = 2^m$, implying (see Lemma B.1)

$$\mathcal{F}^2(h_1) = \mathcal{F}^2(h_2) = 2^{m-1}.$$

This property holds for $f+\varphi_\beta$, for any β. Note that the decomposition of $f+\varphi_\beta$ with respect to H_α, when β ranges over $\boldsymbol{F}_2^m$, is $(h_1+\ell_1,\ h_2+\ell_2)$ where ℓ_i is any affine or constant function and where $\ell_1+\ell_2$ is constant. This proves that the Fourier spectrum of each h_i is $\{\pm 2^{m/2}\}$; thus h_i is a bent function of $m-1$ variables.

iii) $\Rightarrow$ i). Since h_i is bent, then $D_e h_i$ is balanced for any nonzero $e \in H_\alpha$; but

$$\mathcal{F}(D_e f) = \mathcal{F}(D_e h_1) + \mathcal{F}(D_e h_2)$$

for any such e. So we obtain $\mathcal{F}(D_e f) = 0$ for all such e. Hence f satisfies $(\mathcal{H}_\alpha)$. □

Remark V.4: It is important to notice that when f is three-valued almost-optimal (m odd) we have for any α

$$\mathcal{F}^2(f)+\mathcal{F}^2(f+\varphi_\alpha) \in \{0,\ 2^{m+1},\ 2^{m+2}\}$$

according to Lemma B.1. Thus, the values occurring in the Fourier spectrum of h_1 (resp., h_2) are always contained in $\{0, \pm 2^{(m-1)/2}, \pm 2^{(m+1)/2}\}$. This means that h_1 and h_2 are both almost-optimal and this is true for any α—then for any corresponding decomposition of f.

Theorem V.3: Let m be an even integer, $m \geq 4$, and let $f \in \mathcal{B}_m$. Then the following properties are equivalent:

i) there is $\alpha \in \boldsymbol{F}_2^m$ such that f satisfies $(\mathcal{H}_\alpha)$;

ii) f is bent;

iii) f satisfies $(\mathcal{H}_\alpha)$ for all $\alpha \in \boldsymbol{F}_2^m \setminus \{0\}$;

iv) for any α, the decomposition $(h_1,\ h_2)$ of f with respect to H_α satisfies: h_1 and h_2 are three-valued almost-optimal and for any linear Boolean function ℓ of $\mathcal{B}_{m-1}$, we have

$$\mathcal{F}^2(h_1+\ell) \neq \mathcal{F}^2(h_2+\ell)$$

(i.e., $\mathcal{F}^2(h_1+\ell) = 2^m$ if and only if $\mathcal{F}^2(h_2+\ell) = 0$).

Proof: Recall that a Boolean function f is bent if and only if $D_e f$ is balanced for all $e \neq 0$. Hence: ii)$\Leftrightarrow$ iii) and ii) $\Rightarrow$ i).

Assume that f satisfies $(\mathcal{H}_\alpha)$ for some α. Then for all β we have from Lemma V.4

$$\mathcal{F}^2(f+\varphi_\beta)+\mathcal{F}^2(f+\varphi_{\beta+\alpha}) = 2^{m+1}.$$

This implies, from Lemma B.1

$$\mathcal{F}^2(f+\varphi_\beta) = \mathcal{F}^2(f+\varphi_{\beta+\alpha}) = 2^m$$

completing the proof of i) $\Leftrightarrow$ ii).

Assuming that f is bent, we fix $\alpha \in \boldsymbol{F}_2^m$ and we denote by (h_1, h_2) the decomposition of f with respect to H_α. As in the previous proof, we obtain (31) which implies here (by using Lemma B.1), $\mathcal{F}^2(h_i) \in \{0, 2^m\}$ and $\mathcal{F}^2(h_1) \neq \mathcal{F}^2(h_2)$. This property holds if we consider $f+\varphi_\beta$ instead of f in (31)

$$\mathcal{F}^2(f+\varphi_\beta)+\mathcal{F}^2(f+\varphi_{\alpha+\beta}) = 2(\mathcal{F}^2(h_1+\ell)+\mathcal{F}^2(h_2+\ell'))$$

where $\ell,\ \ell' \in \mathcal{B}_{m-1}$, ℓ is a linear function (which can be 0), and ℓ' is either ℓ or $1+\ell$. Thus, $\mathcal{F}^2(h_2+\ell') = \mathcal{F}^2(h_2+\ell)$ and then $\mathcal{F}^2(h_1+\ell) \neq \mathcal{F}^2(h_2+\ell)$ completing the proof of iv).

Conversely, if iv) is satisfied then

$$\mathcal{F}^2(f+\varphi_\beta)+\mathcal{F}^2(f+\varphi_{\alpha+\beta}) = 2^{m+1}$$

for all β, implying that f is bent, completing the proof of ii) $\Leftrightarrow$ iv). □

Remark V.5: Note that the previous theorem is of interest for effective purpose. For checking that a function f is bent it is sufficient to compute the $\mathcal{F}(D_e f)$ for e in some hyperplane.

Example V.4: On the other hand, Property iv) provides some constructions: for every bent function f and every $\alpha \in \boldsymbol{F}_2^m$, $\alpha \neq 0$, both restrictions of f to H_α and $\overline{H}_\alpha$ are three-valued almost-optimal. For instance, choose f in class $\mathcal{PS}_{ap}$ (cf. [33]): $\boldsymbol{F}_2^m$ is identified, as a vector space, with $\boldsymbol{F}_{2^{m/2}} \times \boldsymbol{F}_{2^{m/2}}$ (i.e., the elements of $\boldsymbol{F}_2^m$ are considered as ordered pairs (x, y) where x and y belong to the finite field $\boldsymbol{F}_{2^{m/2}}$) and f is defined as $f(x, y) = g(\frac{x}{y})$, with $\frac{x}{0} = 0$, where g is any balanced Boolean function on $\boldsymbol{F}_{2^{m/2}}$. We do not know how to prove directly (i.e., without using Theorem V.3) that the restrictions of such a function to any hyperplane are three-valued almost-optimal.

We study now the more general case where a function f has balanced derivatives $D_a f$ for all nonzero a of W, a subspace of $\boldsymbol{F}_2^m$ of codimension 2. First note that, with the notation of Section V-B, we obtain, by applying Lemma V.2 and Theorem V.1

$$\sum_{\alpha \in W^\perp} \mathcal{F}^2(f+\varphi_\alpha) = 4 \sum_{e \in W} \mathcal{F}(D_e f) = 4 \sum_{i=1}^{4} \mathcal{F}^2(h_i) \quad (32)$$

where $(h_1, \ldots, h_4)$ is the decomposition of f with respect to W as described at the beginning of Section V-B. These formulas hold when f is replaced by $f+\varphi_\beta$, for any $\beta \in \boldsymbol{F}_2^m$.

Theorem V.4: Let m be any positive integer, $m \geq 3$, and $f \in \mathcal{B}_m$. Assume that there exists a linear subspace $W \subset \boldsymbol{F}_2^m$ of codimension 2 such that $D_a f$ is balanced for any nonzero $a \in W$. Let $(h_1, \ldots, h_4)$ be the decomposition of f with respect to W.

- If m is odd then f is three-valued almost-optimal and every h_i is three-valued almost-optimal.
- If m is even, then either f is bent or $\mathcal{L}(f) = 2^{(m+2)/2}$ and the values occurring in the Fourier spectrum of f belong to $\{0, \pm 2^{m/2}, \pm 2^{(m+2)/2}\}$. Moreover, all the h_i have the same Fourier spectrum: either all the h_i are bent, either all the h_i are three-valued almost-optimal, or the h_i have the same Fourier spectrum with values $\{0, \pm 2^{(m-2)/2}, \pm 2^{m/2}\}$. If all the h_i are three-valued almost-optimal then f is bent.

Proof: Since $\mathcal{F}(D_a f) = 0$ for any nonzero a in W, we have from (32), for any β

$$\sum_{\alpha \in W^\perp} \mathcal{F}^2(f+\varphi_{\alpha+\beta}) = 4\mathcal{F}(D_0 f) = 2^{m+2}.$$

Since $W^\perp$ has cardinality 4, we deduce from Lemma B.2 (in Appendix II) that the Fourier spectrum of f is $\{0, \pm 2^{(m+1)/2}\}$ when m is odd; the values occurring in this Fourier spectrum belong to $\{0, \pm 2^{m/2}, \pm 2^{(m+2)/2}\}$ when m is even. Hence, f is either three-valued almost-optimal (m odd), either bent or such

that $\mathcal{L}(f) = 2^{(m+2)/2}$, and the values of its Fourier transform belong to $\{0, \pm 2^{m/2}, \pm 2^{(m+2)/2}\}$.

Consider now the decomposition of f, say $(h_1, \ldots, h_4)$, with respect to W. We have from (32) again

$$\sum_{i=1}^{4} \mathcal{F}^2(h_i) = 2^m \tag{33}$$

and this property holds for any $f + \varphi_\beta$ and its decomposition, which implies that the values occurring in the Fourier spectrum of each h_i are (by applying Lemma B.2)

- if m is odd, $\{0, \pm 2^{(m-1)/2}\}$—i.e., h_i is three-valued almost-optimal;
- if m is even, either $\{\pm 2^{(m-2)/2}\}$ (i.e., h_i is bent) or contained in $\{0, \pm 2^{(m-2)/2}, \pm 2^{m/2}\}$—with $\mathcal{L}(h_i) = 2^{m/2}$.

According to Lemma B.2, the sum in (33) for even m is either $2^{m-2} \times 4$ or $2^m + 0 \times 3$. If one h_i is bent this sum is always $2^{m-2} \times 4$ implying that all h_i are bent too.

Similarly, if one h_i is three-valued almost-optimal, the values of its Fourier transform are in $\{0, \pm 2^{m/2}\}$. Since the value 2^{m-2} never appears, the sum in (33) is always $2^m + 0 \times 3$ implying that this property holds for all h_i. Moreover, for any β

$$\mathcal{F}(f + \varphi_\beta) = \sum_{i=0}^{4} \mathcal{F}(h_i + \ell) = \pm 2^{m/2} + 0 \times 3 = \pm 2^{m/2}$$

for some ℓ; so f is bent. Now suppose that the h_i are neither bent nor three-valued almost-optimal; then the values appearing in their Fourier spectra are $0, \pm 2^{(m-2)/2}$, and $\pm 2^{m/2}$. We know that the number of times value $\pm 2^{(m-2)/2}$ occurs is the same for each h_i (by using (33) and Lemma B.2 as above), and Parseval's relation settles the case of the two other magnitudes. □

Note that there exist some functions f such that all h_i are three-valued almost-optimal and f is not.

Example V.5: For $m = 7$, we consider

$$\begin{aligned} f(x_1, \ldots, x_7) = {} & x_6 x_7 h_1(x_1, \ldots, x_5) \\ & + (1 + x_6) x_7 h_2(x_1, \ldots, x_5) \\ & + x_6 (1 + x_7) h_3(x_1, \ldots, x_5) \\ & + (1 + x_6)(1 + x_7) h_4(x_1, \ldots, x_5) \end{aligned}$$

where

$$\begin{aligned} h_1(x_1, \ldots, x_5) &= x_1 x_2 x_3 + x_1 x_4 + x_2 x_5 \\ h_2(x_1, \ldots, x_5) &= x_1 x_2 + x_3 x_4 \\ h_3(x_1, \ldots, x_5) &= x_2 x_3 + x_4 x_5 \\ h_4(x_1, \ldots, x_5) &= x_1 x_2 x_3 + x_1 x_4 x_5 + x_2 x_3 + x_2 x_4 + x_3 x_5. \end{aligned}$$

Although all functions h_i are three-valued almost-optimal, f is not: the coefficients $\mathcal{F}(f + \varphi_\alpha)$ belong to

$$\{0, \pm 8, \pm 16, \pm 24, \pm 32, \pm 40\}.$$

Moreover, $D_a f$ is balanced for 23 values of $a \in \boldsymbol{F}_2^7$.

Remark V.6: Take any bent function f on $\boldsymbol{F}_2^m$ (m even), any $(m-2)$-dimensional subspace W of $\boldsymbol{F}_2^m$ and any $a \in \boldsymbol{F}_2^m$. Then the Boolean function $g = f + 1_{a+W}$, where 1_{a+W} denotes the indicator of the flat $a + W$, satisfies the hypothesis of Theorem V.4. Indeed, we have for every $b \in W$: $D_b g = D_b f$ since $a + W$ is invariant under the translation by vector b.

It is clear that the set of functions which satisfy the hypothesis of Theorem V.4 (m even) contains all bent functions and also some functions whose Fourier spectrum is $\{0, \pm 2^{(m+2)/2}\}$. But we have also the following.

Proposition V.5: For every even $m \geq 6$, there exists $f \in \mathcal{B}_m$ satisfying the hypothesis of Theorem V.4, whose Fourier transform takes on exactly the three magnitudes 0, $2^{(m+2)/2}$, and $2^{m/2}$.

Proof: Let $m = 2t$; we identify the elements of $\boldsymbol{F}_2^m$ with the ordered pairs (x, y) where $x = (x_1, \ldots, x_t)$ and $y = (y_1, \ldots, y_t)$. Choose $g \in \mathcal{B}_m$ in Maiorana–McFarland class of bent functions in the form

$$g(x, y) = x \cdot y + k(y)$$

where k is some function in $\mathcal{B}_t$. Set $W = \{(x, y) \,|\, x_1 = x_2 = 0\}$ and

$$f(x, y) = x_1 x_2 + g(x, y) = x_1 x_2 + x \cdot y + k(y).$$

As remarked above, for any $e \in W$ we have $D_e f = D_e g$. Since g is bent then $D_e f$ is balanced. Now remark that

$$f(x, y) = (x_1 + y_2)(x_2 + y_1) + y_1 y_2 + \sum_{i=3}^{t} x_i y_i + k(y).$$

Thus, f is linearly equivalent to the function

$$x_1 x_2 + y_1 y_2 + \sum_{i=3}^{t} x_i y_i + k(y).$$

We see by exchanging x_2 and y_1 that f is linearly equivalent to the function

$$f'(x, y) = x \cdot y + k(x_2, y_2, y_3, \ldots, y_t).$$

It is a simple matter to check that, if $m \geq 6$, there exists a function $k(y)$ such that $\mathcal{F}^2(f' + \varphi_a(x) + \varphi_b(y))$ takes at least once each value of $\{0, 2^{m-2}, 2^m\}$—where $\varphi_a(x) = a \cdot x$ and $\varphi_b(x) = b \cdot y$ in $\mathcal{B}_t$. Take for instance $k(y) = y_1 y_2 y_3$. Then

$$\begin{aligned} & f'(x, y) + \varphi_a(x) + \varphi_b(y) \\ & \quad = \left(\sum_{i=1, 3, \ldots, t} x_i (y_i + a_i) \right) + x_2 y_2 (1 + y_3) + a_2 x_2 + b \cdot y. \end{aligned}$$

Since

$$\sum_{x_1, x_3, \ldots, x_t \in \boldsymbol{F}_2} (-1)^{\sum_{i=1, 3, \ldots, t} x_i (y_i + a_i)} \neq 0$$

if and only if $y_i = a_i$ for every $i = 1, 3, \ldots, t$, we deduce:

$$\begin{aligned} & \mathcal{F}(f' + \varphi_a(x) + \varphi_b(y)) \\ & \quad = \pm 2^{t-1} \sum_{x_2, y_2 \in \boldsymbol{F}_2} (-1)^{x_2 y_2 (1 + a_3) + a_2 x_2 + b_2 y_2}. \end{aligned}$$

If $a_3 = 0$, then we obtain $\pm 2^t$; if $a_3 = 1$, $a_2 = 0$, and $b_2 = 0$, we obtain $\pm 2^{t+1}$; and if $a_3 = 1$, $a_2 \neq 0$, or $b_2 \neq 0$ we obtain 0. The proof is complete. □

Remark V.7: There exist three-valued almost-optimal functions with m variables, m odd, which do not satisfy the hypotheses of Theorem V.4. For example, for $m = 7$, the function

$$f(x_1, \ldots, x_7) = x_2x_3 + x_4x_6 + x_5x_7 + x_1x_6x_7 + x_5x_6x_7 + x_2x_3x_6x_7 + x_4x_5x_6x_7$$

is three-valued almost-optimal. It has exactly 14 nonbalanced derivatives, which are all D_ef for

$$e \in \langle e_1, e_2, e_3\rangle \cup \langle e_1, e_4, e_5\rangle.$$

Open Problem V.2: Find some general property for a function f such that D_af is balanced when $a \in W$, $a \neq 0$, where W has codimension 3.

D. The Nonbalanced Derivatives

On the other hand, we consider the set of nonbalanced derivatives. Recall that, for $f \in \mathcal{B}_m$, E_f is the set $\{e \in \boldsymbol{F}_2^m | \mathcal{F}(D_ef) = 0\}$ and $\overline{E_f}$ is the complementary set $\boldsymbol{F}_2^m \setminus E_f$. In this section, we consider the rank of $\overline{E_f}$. For clarity, we first indicate an obvious property.

Lemma V.5: Let r be the rank of $\overline{E_f}$. Then $r < m$ means that there is a subspace V of dimension r such that $a + V$ is contained in E_f for all $a \notin V$.

It is natural to first consider the small values of r. As a direct application of our previous results, we are able to characterize the functions which correspond to the cases $r \leq 2$.

Corollary V.4: Let m be an odd integer, $m \geq 3$, $f \in \mathcal{B}_m$, and $e \in \boldsymbol{F}_2^m$. Then the following properties are equivalent:

- f is almost-optimal and e is a linear structure of f;
- f is three-valued almost-optimal and e is a linear structure of f;
- $\overline{E_f} = \{0, e\}$.

Proof: If f has a linear structure then $\mathcal{V}(f) \geq 2^{2m+1}$ (see Lemma II.1). Suppose that, moreover, f is almost-optimal. In accordance with Proposition V.2, the only possibility is $\mathcal{V}(f) = 2^{2m+1}$ which means (when f is almost-optimal) that f is three-valued almost-optimal. Since $\mathcal{F}^2(D_0f) = \mathcal{F}^2(D_ef) = 2^{2m}$, we deduce

$$\mathcal{V}(f) = \sum_{a \in \{0, e\}} \mathcal{F}^2(D_af)$$

providing $\mathcal{F}^2(D_af) = 0$ for $a \notin \{0, e\}$, according to (6).

Assume now that $\overline{E_f} = \{0, e\}$. Clearly, the set $E_f \cup \{0\}$ contains a subspace of codimension 1. So we apply Theorem V.2 and deduce that f is three-valued almost-optimal. Since

$$\mathcal{F}^2(f) = \mathcal{F}(D_0f) + \mathcal{F}(D_ef)$$

then $\mathcal{F}(D_ef) = \pm 2^m$, completing the proof. □

Corollary V.5: Let m be an even integer, $m \geq 4$, and $f \in \mathcal{B}_m$. Let V be some linear space of dimension 2. Then the following properties are equivalent:

- f is almost-optimal and any $e \in V$ is a linear structure of f;
- f is three-valued almost-optimal and any $e \in V$ is a linear structure of f;
- $\overline{E_f} = V$.

Proof: We proceed as in the previous proof. If V is a linear space for f, then $\mathcal{V}(f) \geq 2^{2m+2}$. If, moreover, f is almost-optimal then $\mathcal{V}(f) = 2^{2m+2}$ which means that f is three-valued almost-optimal. Now compute the sum-of-squares indicator

$$\mathcal{V}(f) = \sum_{a \in V} \mathcal{F}^2(D_af) + \sum_{a \notin V} \mathcal{F}^2(D_af) = 2^{2m+2} + \sum_{a \notin V} \mathcal{F}^2(D_af)$$

providing $\overline{E_f} = V$.

Conversely, assume that $\overline{E_f} = V$. Thus, $E_f \cup \{0\}$ contains a subspace of codimension 2, say W. In accordance with Theorem V.4, $\mathcal{L}(f) = 2^{(m+2)/2}$ and the values of the Fourier transform of f lie in $\{0, \pm 2^{m/2}, \pm 2^{(m+2)/2}\}$ (f cannot be bent). We can assume that $\mathcal{F}^2(f) = 2^{m+2}$. By using (27), taking $H_\alpha = W \cup (a + W)$ with $a \in V$, $a \neq 0$, we obtain

$$\begin{aligned} \mathcal{F}^2(f) + \mathcal{F}^2(f + \varphi_\alpha) &= 2(\mathcal{F}(D_0f) + \mathcal{F}(D_af)) \\ &= 2^{m+1} + 2\mathcal{F}(D_af) \end{aligned} \tag{34}$$

where $0 < |\mathcal{F}(D_af)| \leq 2^m$. Then $\mathcal{F}(D_af) = \pm 2^m$. Note that this property holds for any $a \in V$. So we have proved that f is almost-optimal and that it has any $a \in V$ as linear structure, completing the proof. □

Remark V.8: Note that we are not able to give the Fourier spectrum of any almost-optimal function which has a linear space of dimension 1 when m is even. Actually, this problem is equivalent to the determination of Fourier spectrum of almost-optimal functions of $\mathcal{B}_{m-1}$ (see the next example).

Example V.6: There exist almost-optimal (non-three-valued) functions of degree 4 for odd $m \geq 5$. For instance, for $m = 5$, the function

$$g(x_1, \ldots, x_5) = x_3x_5 + x_2x_4 + x_1x_2x_3 + x_2x_3x_4x_5$$

is given in [23]. Its Fourier transform takes all the values in $\{0, \pm 4, \pm 8\}$. Note that in the decomposition of g with respect to the subspace defined by $x_3 = 0$ is x_2x_4, a quadratic component of g (see Proposition V.4). Moreover, one can check that $\overline{E_g}$ is a subspace of dimension 3.

Now consider the function of six variables

$$f(x_1, \ldots, x_6) = g(x_1, \ldots, x_5) + x_6.$$

It is clear that $(0, \ldots, 0, 1)$ is a linear structure of f and it is easy to check that the set of values appearing in the Fourier spectrum of f is $\{0, \pm 8, \pm 16\}$.

The previous corollaries were partially proved in [34] where the authors study the cases $\#\overline{E_f} = 1, 2, \ldots, 6$. Generally, it seems difficult to characterize f such that $\overline{E_f}$ is a linear space of dimension k for some k (see [10]). When the rank of $\overline{E_f}$ is 3, we can give the next property but cannot describe the case $\#\overline{E_f} = 8$—examples are easily obtained (see Example V.6).

Corollary V.6: Let $f \in \mathcal{B}_m$ and assume that the rank of $\overline{E_f}$ is 3.

If $\#\overline{E_f} < 8$ then E_f contains all nonzero elements of some subspace of codimension 2. So Theorem V.4 can be applied.

Proof: Set $E'_f = E_f \cup \{0\}$. Assume that the cardinality of $\overline{E_f}$ is strictly less than 8. Let (e_1, e_2, e_3) be linearly independent in $\overline{E_f}$; by completing, we have a basis $(e_1, \ldots, e_m)$ of $\boldsymbol{F}_2^m$ such that $W = \langle e_4, \ldots, e_m \rangle$ is contained in E'_f. But there is some a, a linear combination of (e_1, e_2, e_3), which is in E'_f. So the subspace $W \cup (a + W)$, of codimension 2 is contained in E'_f. □

Note that, for any t, there exist some functions f such that $\overline{E_f}$ has rank t (such functions can be constructed recursively, by taking partially bent functions). Hence, this property does not seem to be significant. It nevertheless induces some simplifications on the decompositions of the function, as shown in the next theorem.

Theorem V.5: Suppose that $\overline{E_f}$ is contained in W, a subspace of dimension t, $1 < t < m$. Considering notation of Theorem V.1, let $\{h_a | a \in V\}$ be the decomposition of f with respect to W, then we have

i) $\sum_{a \in V} \mathcal{F}^2(h_a) = \mathcal{F}^2(f)$. Moreover,

$$\mathcal{L}^2(f) \leq \sum_{a \in V} \mathcal{L}^2(h_a);$$

ii) $\mathcal{F}^2(f + \varphi_\alpha) = \mathcal{F}^2(f)$ for any $\alpha \in W^\perp$ and the Fourier spectrum of f cannot have more than 2^t magnitudes.

Proof: Let $\alpha \in W^\perp$. Any $e \in \overline{H}_\alpha$ does not belong to W, since $e \cdot \alpha = 1$. It follows that $D_e f$ is balanced for all $e \in \overline{H}_\alpha$. We deduce from Lemma V.3 that $\mathcal{F}^2(f + \varphi_\alpha) = \mathcal{F}^2(f)$. Thus, applying Theorem V.1, we obtain

$$\sum_{\alpha \in W^\perp} \mathcal{F}^2(f + \varphi_\alpha) = 2^{m-t} \mathcal{F}^2(f) = 2^{m-t} \sum_{a \in V} \mathcal{F}^2(h_a).$$

It follows that

$$\sum_{a \in V} \mathcal{F}^2(h_a) = \mathcal{F}^2(f).$$

This property holds for any $f + \varphi_\beta$, $\beta \in \boldsymbol{F}_2^m$, since $D_e(f + \varphi_\beta)$ is balanced as soon as $D_e f$ is balanced. The upper bound on $\mathcal{L}^2(f)$ is obviously deduced. Take any $\beta \in \boldsymbol{F}_2^m$; then we obtain, as above

$$\mathcal{F}^2(f + \varphi_{\beta+\alpha}) = \mathcal{F}^2(f + \varphi_\beta)$$

for all $\alpha \in W^\perp$. This implies that the Fourier spectrum of f cannot have more than 2^t magnitudes. □

Several corollaries can be deduced. We study, for instance, the case where W has codimension 1.

Corollary V.7: Assume that $f \in \mathcal{B}_m$ is such that $\overline{E_f} \subset H_\alpha$, some subspace of codimension 1. Denote by (h_1, h_2) the decomposition of f with respect to H_α. Then we have

i) for any $\beta \in \boldsymbol{F}_2^m$

$$\mathcal{F}^2(f + \varphi_\beta) = \max(\mathcal{F}^2(h_1 + \ell), \mathcal{F}^2(h_2 + \ell))$$

where $(h_1 + \ell, h_2 + \ell + \varepsilon)$ is the decomposition of $f + \varphi_\beta$, ℓ is a linear function in $\mathcal{B}_{m-1}$ and ε is constant;

ii) for every linear function ℓ in $\mathcal{B}_{m-1}$, at least one term in the pair $(\mathcal{F}(h_1 + \ell), \mathcal{F}(h_2 + \ell))$ is zero;

iii)

$$\mathcal{L}(f) = \max(\mathcal{L}(h_1), \mathcal{L}(h_2))$$
$$\mathcal{N}(f) = 2^{m-2} + \min(\mathcal{N}(h_1), \mathcal{N}(h_2))$$

and

$$\mathcal{V}(f) = \mathcal{V}(h_1) + \mathcal{V}(h_2).$$

Proof: Since $\mathcal{F}(f) = \mathcal{F}(h_1) + \mathcal{F}(h_2)$, we have, according to Theorem V.5

$$\mathcal{F}^2(f) = \mathcal{F}^2(h_1) + \mathcal{F}^2(h_2) = (\mathcal{F}(h_1) + \mathcal{F}(h_2))^2.$$

Thus, $\mathcal{F}(h_1)\mathcal{F}(h_2) = 0$ providing either $\mathcal{F}(h_1) = 0$ or $\mathcal{F}(h_2) = 0$—where both can be zero. This property holds when we consider $f + \varphi_\beta$ for any $\beta \in \boldsymbol{F}_2^m$ and the decomposition of $f + \varphi_\beta$ which is actually of the form $(h_1 + \ell, h_2 + \ell + \varepsilon)$, where ℓ is linear and ε is constant. We obviously have that $\mathcal{F}^2(h_2 + \ell + \varepsilon) = \mathcal{F}^2(h_2 + \ell)$. Then i) and ii) are clearly proved and the values of $\mathcal{L}(f)$ and $\mathcal{N}(f)$ given in iii) are easily deduced.

Now we compute $\mathcal{V}(f)$, by using (7) and the previous properties. We denote by L_{m-1} the set of all linear functions in $\mathcal{B}_{m-1}$. We then deduce

$$\begin{aligned}
2^m \mathcal{V}(f) &= \sum_{\beta \in \boldsymbol{F}_2^m} \mathcal{F}^4(f + \varphi_\beta) \\
&= \sum_{\ell \in L_{m-1},\, \varepsilon \in \boldsymbol{F}_2} (\mathcal{F}^2(h_1 + \ell) + \mathcal{F}^2(h_2 + \ell + \varepsilon))^2 \\
&= 2 \sum_{\ell \in L_{m-1}} (\mathcal{F}^2(h_1 + \ell) + \mathcal{F}^2(h_2 + \ell))^2 \\
&= 2 \sum_{\ell \in L_{m-1}} (\mathcal{F}^4(h_1 + \ell) + \mathcal{F}^4(h_2 + \ell)) \\
&= 2 \left(2^{m-1} \mathcal{V}(h_1) + 2^{m-1} \mathcal{V}(h_2)\right) \\
&= 2^m \left(\mathcal{V}(h_1) + \mathcal{V}(h_2)\right)
\end{aligned}$$

completing the proof. □

Example V.7: Let $f \in \mathcal{B}_m$ such that $D_a f$ is a linear nonconstant function, for some a. So it is clear that $D_e D_a f$ is constant for any e.

Set $D_a f = \varphi_\alpha$ and recall that H_α denotes the kernel of φ_α. It is clear that $D_e D_a f = 1$ if and only if $e \notin H_\alpha$. Thus, according to Proposition II.5, $D_e f$ is balanced for any $e \notin H_\alpha$. This implies that $\overline{E_f} \subset H_\alpha$. It is very easy to construct such a function f. For instance,

$$f(x_1, \ldots, x_5) = x_1 x_2 x_3 + x_1 x_3 x_4 + x_2 x_5$$

is such that $D_a f$, $a = (0, 0, 0, 0, 1)$, is equal to x_2.

Remark V.9: For any bent function f we have $\overline{E_f} = \{0\}$. Hence, f satisfies the hypothesis of Theorem V.5 for any W. For such a function, Theorem V.5 i), becomes

$$\sum_{a\in V} \mathcal{F}^2(h_a) = 2^m \quad \text{implying} \quad 2^m \le \sum_{a\in V} \mathcal{L}^2(h_a).$$

Note that the property on the left holds for any $f+\varphi_\beta$ and the corresponding decomposition. If V has dimension 1 or 2 we can apply Lemmas B.1 and B.2. So we obtain again some results given by Theorem V.3 and V.4.

Appendix I

We briefly recall some properties of quadratic functions. More can be found in [22, Ch. 15] and [15]. In this appendix, f denotes a Boolean function of degree 2 of m variables. The *associated symplectic form* of f is the mapping from $(\boldsymbol{F}_2^m)^2$ to $\boldsymbol{F}_2$

$$\Psi(u, v) = f(0) + f(u) + f(v) + f(u+v)$$

where $(u, v) \in (\boldsymbol{F}_2^m)^2$. The *kernel* of Ψ is defined as follows:

$$\mathcal{E}_f = \{u \in \boldsymbol{F}_2^m | \forall v \in \boldsymbol{F}_2^m \colon \Psi(u, v) = 0\}.$$

The set $\mathcal{E}_f$ is a $\boldsymbol{F}_2$-subspace of $\boldsymbol{F}_2^m$ of dimension $m-2h$, where $2h$ is the *rank* of Ψ. This rank satisfies

- $1 \le h \le m/2$ for even m, and
- $1 \le h \le (m-1)/2$ for odd m.

Obviously, $\mathcal{E}_{f+\ell} = \mathcal{E}_f$ for any linear function ℓ. The Fourier spectrum of f (and thus the weight distribution of the corresponding coset $\Omega_f + R(1, m)$) only depends on h (cf. [22, p. 441]). For such a coset, the weights are $\{2^{m-1}, 2^{m-1} \pm 2^{m-h-1}\}$ and the corresponding numbers of codewords $\{2^{m+1}-2^{2h+1}, 2^{2h}\}$.

So the quadratic functions are three-valued unless $h = m/2$ for even m. In this case, the function is bent and its Fourier spectrum is $\{\pm 2^{m/2}\}$.

Proposition A.1: An element a is in $\mathcal{E}_f$ if and only if the function $D_a f$ is constant. The subspace $\mathcal{E}_f$ is the linear space of f.

Moreover, f is balanced if and only if there is $a \in \mathcal{E}_f$ such that $D_a f = 1$.

Proof: Note that $D_a f$ is constant if and only if

$$f(v) + f(a+v) = \varepsilon$$

for all v, where ε denotes a constant—either 0 or 1. But $a \in \mathcal{E}_f$ means

$$f(0) + f(a) = f(v) + f(a+v) \qquad \forall v$$

or, equivalently, $D_a f(v) = D_a f(0)$, for all v. This proves the first sentence of the proposition.

If $D_a f = 1$ for some a then f is balanced (see Proposition II.5). Conversely, suppose that f is balanced. Denote by τ the dimension of $\mathcal{E}_f$. Recall that $\tau = m-2h$ and that the number of 0's in the Fourier spectrum of f is equal to $2^{m+1} - 2^{2h+1}$. Note that f cannot be bent, so that the dimension of $\mathcal{E}_f$ is at least 1. We assume that for any $a \in \mathcal{E}_f$, $a \neq 0$, $D_a f = 0$ and we are going to prove that this is impossible. Define the subspace

$$B = \{\varphi_b \,|\, \mathcal{E}_f \subset \mathrm{Ker}\varphi_b\}$$

of the space of linear functions $\{\varphi_b \,|\, b \in \boldsymbol{F}_2^m\}$ (where $\varphi_0 = 0$). The number of hyperplanes of $\boldsymbol{F}_2^m$ containing $\mathcal{E}_f$ is equal to $2^{m-\tau}-1$. Thus, the cardinality of B is $2^{m-\tau}$ (by adding φ_0). We then have $2^m - 2^{m-\tau}$ functions φ_u, $u \neq 0$, such that $\varphi_u(a) = 1$ for some $a \in \mathcal{E}_f$. But for such φ_u, we have

$$D_a(f+\varphi_u) = D_a f + 1 = 1$$

implying that $f+\varphi_u$ is balanced. Therefore, $f+\varphi_u+1$ is balanced too providing at all $2^{m+1} - 2^{2h+1}$ zero values in the Fourier spectrum of f. We have proved that any balanced function $f+\varphi_u$ is such that $u \neq 0$ (since $u \notin B$). This contradicts that f itself is balanced. □

Appendix II

Lemma B.1: Let m be an integer, $m \ge 0$, and let X and Y be two integers. Then the condition $X^2 + Y^2 = 2^{m+1}$ implies

- if m is even then $X^2 = Y^2 = 2^m$;
- if m is odd then $X^2 = 2^{m+1}$ and $Y = 0$ or vice versa.

Proof: This lemma can be proved by induction, as it is shown in [34], but this result was first stated and proven by Jacobi in 1828. His proof relies on the fact that the number of solutions $(X_1, \ldots, X_k) \in \mathbb{Z}^k$ of the equation

$$X_1^2 + \cdots + X_k^2 = N \tag{35}$$

is exactly the coefficient c_N of x^N in the expansion of θ^k, where

$$\theta = \sum_{i=-\infty}^{+\infty} x^{i^2} = 1 + 2x + 2x^4 + 2x^9 + \cdots.$$

In 1828, Jacobi proved that (see [35] and [36])

$$\theta^2 = 1 + 4\sum_{N=1}^{+\infty} \left(\sum_{d|N,\, d\equiv 1 \bmod 4} 1 - \sum_{d|N,\, d\equiv 3 \bmod 4} 1 \right) x^N$$

which means that the number of solutions of (35) satisfies

- if there exists a divisor d of N, $d \equiv 3 \bmod 4$, which occurs in N to an odd power, then $c_N = 0$;
- else

$$c_N = 4\left(\sum_{d|N,\, d\equiv 1 \bmod 4} 1 - \sum_{d|N,\, d\equiv 3 \bmod 4} 1 \right).$$

Then we have $c_{2^{m+1}} = 4$; this means that for our case ($N = 2^{m+1}$ and $X_i \ge 0$ for all i), the only solutions are the ones presented in the lemma. □

Lemma B.2: Let m be an integer, $m \ge 1$, and let X, Y, Z and T be four integers. Then the condition

$$X^2 + Y^2 + Z^2 + T^2 = 2^{m+2}$$

implies

- if m is even, then either $X^2 = Y^2 = Z^2 = T^2 = 2^m$ or $X^2 = 2^{m+2}$ and $Y^2 = Z^2 = T^2 = 0$;
- if m is odd, then $X^2 = Y^2 = 2^{m+1}$ and $Z = T = 0$.

Proof: In the same way, this result can be obtained by induction, but was stated by Jacobi in 1828, since we have

$$\theta^4 = 1 + 8 \sum_{s\ge 1,\, 4\nmid s} \frac{sx^s}{1-x^s}$$

where s runs through all positive integers which are not multiples of 4. Then we have $c_{2^{m+2}} = 24$; Jacobi then proved that for our case, the solutions presented in the lemma are the only ones. □

ACKNOWLEDGMENT

The authors wish to thank one anonymous referee for many useful comments which greatly improved the manuscript.

REFERENCES

[1] T. Kasami, "Weight distributions of Bose–Chaudhuri–Hocquenghem codes," in *Proc. Conf. Combinatorial Mathematics and its Applications*. Chapel Hill, NC: Univ. North Carolina Press, 1968, pp. 335–357.

[2] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, pp. 656–715, 1949.

[3] M. Matsui, "Linear cryptanalysis method for DES cipher," in *Advances in Cryptology—EUROCRYPT'93 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1993, vol. 765, pp. 386–397.

[4] A. Canteaut and M. Trabbia, "Improved fast correlation attacks using parity-check equations of weight 4 and 5," in *Advances in Cryptology—EUROCRYPT 2000 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 2000, vol. 1807, pp. 573–588.

[5] A. F. Webster and S. E. Tavares, "On the design of S-boxes," in *Advances in Cryptology—CRYPTO'85 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1985, vol. 219, pp. 523–534.

[6] B. Preneel, W. V. Leekwijck, L. V. Linden, R. Govaerts, and J. Vandewalle, "Propagation characteristics of Boolean functions," in *Advances in Cryptology—EUROCRYPT'90 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1991, vol. 437, pp. 155–165.

[7] B. Preneel, R. Govaerts, and J. Vandewalle, "Boolean functions satisfying higher order propagation criteria," in *Advances in Cryptology—EUROCRYPT'91 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1992, vol. 547, pp. 141–152.

[8] Y. Zheng, X.-M. Zhang, and H. Imai, "Restriction, terms and nonlinearity of Boolean functions," *Theor. Comput. Sci.*, vol. 226, no. 1–2, pp. 207–223, 1999.

[9] X.-M. Zhang and Y. Zheng, "GAC—The criterion for global avalanche characterics of cryptographic functions," *J. Univ. Comput. Sci.*, vol. 1, no. 5, pp. 320–337, 1995.

[10] A. Canteaut, C. Carlet, P. Charpin, and C. Fontaine, "Propagation characteristics and correlation-immunity of highly nonlinear Boolean functions," in *Advances in Cryptology—EUROCRYPT 2000 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 2000, vol. 1807, pp. 507–522.

[11] O. S. Rothaus, "On bent functions," *J. Combin. Theory Ser. A*, vol. 20, pp. 300–305, 1976.

[12] C. Fontaine, "On some cosets of the first-order Reed–Muller code with high minimum weight," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1237–1243, May 1999.

[13] T. Helleseth, T. Kløve, and J. Mykkeltveit, "On the covering radius of binary codes," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 627–628, Sept. 1978.

[14] J. H. Evertse, "Linear structures in block ciphers," in *Advances in Cryptology—EUROCRYPT' 87 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1988, vol. 304, pp. 249–266.

[15] C. Carlet, "Codes de Reed–Muller, codes de Kerdock et de Preparata," Ph.D. dissertation, Univ. Paris 6, Paris, France, 1990.

[16] C. Carlet, "Partially-bent functions," *Des., Codes Cryptogr.*, no. 3, pp. 135–145, 1993.

[17] E. F. Assmus and J. Key, "Polynomials codes and finite geometry," in *Handbook of Coding Theory—Part 2: Connections*, V. S. Pless, W. C. Huffman, and R. A. Brualdi, Eds. Amsterdam, The Netherlands: Elesevier, 1998, ch. 16, pp. 1269–1343.

[18] P. Charpin, "Codes cycliques étendus invariants sous le groupe affine,", Thése d'Etat, Univ. Paris 7, Paris, France, 1987. LITP 87-6.

[19] P. Charpin, "Open problems on cyclic codes," in *Handbook of Coding Theory—Part 1*, V. S. Pless, W. C. Huffman, and R. A. Brualdi, Eds. Amsterdam, The Netherlands: Elsevier, 1998, ch. 11, pp. 963–1063.

[20] S. D. Berman, "On the theory of group codes," *Kibernetica*, vol. 1, no. 1, pp. 31–39, 1967.

[21] V. Pless, "Power moment identities on weight distributions in error-correcting codes," *Inform. Contr.*, vol. 3, pp. 147–152, 1963.

[22] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam, The Netherlands: North-Holland, 1977.

[23] E. R. Berlekamp and L. R. Welch, "Weight distributions of the cosets of the $(32, 6)$ Reed–Muller code," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 203–207, Jan. 1972.

[24] A. Canteaut, "On the weight distributions of optimal cosets of the first-order Reed–Muller codes," *IEEE Trans. Inform. Theory*, vol. 47, pp. 407–413, Jan. 2001.

[25] A. Canteaut, P. Charpin, and H. Dobbertin, "Binary m-sequences with three-valued crosscorrelation: A proof of Welch conjecture," *IEEE Trans. Inform. Theory*, vol. 46, pp. 4–8, Jan. 2000.

[26] ——, "Weight divisibility of cyclic codes, highly nonlinear functions on GF(2^m) and crosscorrelation of maximum-length sequences," *SIAM J. Discr. Math.*, vol. 13, no. 1, pp. 105–138, 2000.

[27] H. Dobbertin, "Another proof of Kasami's Theorem," *Des., Codes Cryptogr.*, vol. 17, no. 1/3, pp. 177–180, 1999.

[28] C. Carlet, P. Charpin, and V. Zinoviev, "Codes, bent functions and permutations suitable for DES-like cryptosystems," *Des., Codes Cryptogr.*, vol. 15, no. 2, pp. 125–156, 1998.

[29] E. Filiol and C. Fontaine, "Highly nonlinear balanced Boolean functions with a good correlation-immunity," in *Advances in Cryptology—EUROCRYPT'98 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1998, vol. 1403, pp. 475–488.

[30] H. Dobbertin, "Construction of bent functions and balanced Boolean functions with high nonlinearity," in *Fast Software Encryption (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1994, vol. 1008, pp. 61–74.

[31] N. J. Patterson and D. H. Wiedemann, "The covering radius of the $[2^{15}, 16]$ Reed–Muller code is at least 16276," *IEEE Trans. Inform. Theory*, vol. IT-36, no. 2, p. 443, 1983.

[32] X.-D. Hou, "On the covering radius of $R(1, m)$ in $R(3, m)$," *IEEE Trans. Inform. Theory*, vol. 42, no. 3, pp. 1035–1037, 1996.

[33] J. F. Dillon, "Elementary Hadamard Difference sets," Ph.D. dissertation, University of Maryland, 1974.

[34] X.-M. Zhang and Y. Zheng, "Characterizing the structures of cryptographic functions satisfying the propagation criterion for almost all vectors," *Designs, Codes and Cryptography*, vol. 7, no. 1, pp. 111–134, 1996.

[35] C. G. J. Jacobi, "Correspondance mathématique entre Legendre et Jacobi," *J. Reine Angew. Math.*, no. 80, pp. 205–279, 1875.

[36] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 3rd ed. London, U.K.: Clarendon, 1954.

[37] C. Carlet, "Two new classes of bent functions," in *Advances in Cryptology—EUROCRYPT'93 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1994, vol. 765, pp. 77–101.

[38] ——, "On the propagation criterion of degree ℓ and order k," in *Advances in Cryptology—EUROCRYPT'98 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1998, vol. 1403, pp. 462–474.

[39] ——, "On cryptographic propagation criteria for Boolean functions," *Inform. Comput.*, no. 151, pp. 32–56, 1999.

[40] T. Kasami, "The weight enumerators for several classes of subcodes of the second order binary Reed–Muller codes," *Inform. Contr.*, vol. 18, pp. 369–394, 1971.

[41] S. Maitra and P. Sarkar, "Highly nonlinear resilient functions optimizing Siegenthaler's inequality," in *Advances in Cryptology—CRYPTO'99 (Lecture Notes in Computer Science)*. Berlin, Germany: Springer-Verlag, 1999, vol. 1666, pp. 198–215.

[42] R. J. McEliece, "Weight congruence for p-ary cyclic codes," *Discr. Math.*, vol. 3, pp. 177–1925, 1972.

[43] T. Siegenthaler, "Correlation-immunity of nonlinear combining functions for cryptographic applications," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 776–780, Sept. 1984.

[44] G. Xiao and J. L. Massey, "A spectral characterization of correlation-immune combining functions," *IEEE Trans. Inform. Theory*, vol. 34, pp. 569–571, May 1988.