On those Boolean functions that are coset leaders of first order Reed-Muller codes

Claude Carlet* and Serge Feukoua†

Abstract

In this paper, we study the class of those Boolean functions that are coset leaders of first order Reed-Muller codes. We study their properties and try to better understand their structure (which seems complex), by studying operations on Boolean functions that can provide coset leaders (we show that these operations all provide coset leaders when the operands are coset leaders, and that some can even produce coset leaders without the operands being coset leaders). We characterize those coset leaders that belong to the well known classes of direct sums of monomial Boolean functions and Maiorana-McFarland functions. Since all the functions of Hamming weight at most 2^{n-2} are automatically coset leaders, we are interested in constructing infinite classes of coset leaders having possibly Hamming weight larger than 2^{n-2}.

Keywords: Boolean functions, coset leader, Hamming Weight.

1 Introduction

The Reed-Muller code $RM(r, n)$ (of length 2^n and order r) is the linear code equal to the vector space of all n-variable Boolean functions $f : \mathbb{F}_2^n \mapsto \mathbb{F}_2$ (represented by the last column of their truth tables) of algebraic degree at most r; in particular, when $r = 1$, it equals the vector space of all affine Boolean functions (see [12] for more details on Reed-Muller codes).

A coset leader of a linear code C of length N (that is, of an $\mathbb{F}_2$-subspace of $\mathbb{F}_2^N$) is any vector e of length N that has minimum Hamming weight among all the elements of the coset of C to which it belongs, that is, among the elements of $e + C$.

*LAGA, Department of Mathematics, University of Paris 8 (and Paris 13 and CNRS), Saint–Denis cedex 02, France, and University of Bergen, Norway. E-mail: claude.carlet@gmail.com. The research of this author is partly supported by the Trond Mohn Foundation and Norwegian Research Council.

†Serge Feukoua, University of Yaoundé 1, Faculty of Sciences, Department of Mathematics. E-mail: feukouaf@yahoo.fr.

Coset leaders play a role in coding theory; they can be used in maximum likelihood decoding analysis (see [1]): if a word g is sent, and the decoder receives the word h, the word $e = h - g$ is called the error vector and belongs to the coset $h + C$ of C. The decoding of the word h depends on the weight t of e. For instance, it is known that if $t > \frac{d-1}{2}$, where d is the minimal distance of C, then there may be errors of weight t which are not uniquely decodable (see [10]). Helleseth and Klove defined in [10] the notion of false neighbor of a codeword $h \in C$ which is a nonzero codeword $g \in C$ such that $w_H(h - g) \leq w_H(h)$. An error e has no false neighbor if and only if it is the unique coset leader in its coset.

Coset leaders also present an interest from the cryptographic viewpoint. A coset leader of $RM(1, n)$ is a Boolean function f whose Hamming weight $w_H(f)$ is the minimum of the set $\{w_H(f + l),\ l \in RM(1, n)\}$. In other words, a coset leader is a Boolean function whose Hamming weight is equal to its nonlinearity, an important notion in cryptography, see [4].

It is straightforward that any Boolean function in n variables and of Hamming weight at most 2^{n-2} is a coset leader of $RM(1, n)$: if $w_H(f) \leq 2^{n-2}$, then for every nonzero affine function a, we have (according to the triangular inequality) that $w_H(f + a) \geq w_H(a) - w_H(f) \geq 2^{n-2}$. The case of coset leaders of Hamming weight larger than 2^{n-2} has been little studied. It may be illusory to hope to ever find a simple characterization and/or a nice structure of the set of coset leaders, since coset leaders represent an important part of all Boolean functions: any coset of $RM(1, n)$ contains at least one coset leader, by definition, and the number of cosets of $RM(1, n)$ being equal to 2^{2^n-n-1}, the number of coset leaders of $RM(1, n)$ is at least 2^{2^n-n-1} (it is in fact, significantly larger, since for instance there are 2^n coset leaders in each coset of $RM(1, n)$ containing bent functions). Note that by the covering radius bound, the nonlinearity of any function f, and therefore the Hamming weight of any coset leader, can not be larger than $\lfloor 2^{n-1} - 2^{\frac{n}{2}-1} \rfloor$ (for n even, this bound is tight - and achieved by the so-called bent functions [15] - and for $n \geq 9$ odd, the maximum nonlinearity is unknown). The property of being bent is clearly preserved by the addition of affine functions.

Determining the coset leaders in a given coset $f + RM(1, n)$ is closely related to determining the Walsh transform of f (see the definition in Section 2). Indeed, $f(x) + u \cdot x + \epsilon$ (where "$\cdot$" is the inner product chosen for defining the Walsh transform) is a coset leader if and only if $W_f(u)$ equals $(-1)^\epsilon \max\{|W_f(a)|; a \in \mathbb{F}_2^n\}$. All the numerous papers which calculate the Walsh spectra of some Boolean functions deal then with coset leaders. We shall not describe the contents of all. In particular, we shall not describe those dealing with quadratic functions: we know that every quadratic function is, up to the composition on the right by some affine permutation, equal to $f(x) = x_1x_2 + \cdots + x_{2k-1}x_{2k} + a \cdot x + \epsilon$, for some $k \leq \frac{n}{2}$ and $a \in \mathbb{F}_2^n, \epsilon \in \mathbb{F}_2$, and that $\max\{|W_f(a)|; a \in \mathbb{F}_2^n\}$ equals then 2^{n-k} and the coset leaders in $f + RM(1, n)$ are the functions of the form $(x_1 + a_2)(x_2 + a_1) + \cdots + (x_{2k-1} + a_{2k})(x_{2k} + a_{2k-1})$, for some $a_1, \ldots, a_{2k}$. Some papers deal with the Walsh spectrum of non-quadratic Boolean functions. For instance, in [16] was shown that the degree 3 rotation symmetric function $x_1x_2x_3 + x_2x_3x_4 + \cdots + x_nx_1x_2$ has the same nonlinearity as its

weight, after that the paper [8] made the same observation with the degree 2 rotation symmetric functions $x_1x_l + x_2x_{l+1} \cdots + x_nx_{n+l-1}$ for n even, and conjectured the same was true for degree 3 functions. In [2] is also studied the algorithmic viewpoint on coset leaders: an algorithmic process is given for finding the whole set of coset leaders of a binary code C by using the Gröbner representation of C which allows the description of a complete algorithm for the computation of its set of coset leaders.

In the present paper, we first study the properties of the set of coset leaders, and we show that the operations of direct sum, direct product and direct majority (an operation that we define) are internal in the set of coset leaders and that they provide, under rather weak hypotheses, coset leaders from Boolean functions without the operands being coset leaders themselves. We construct coset leaders f of Hamming weight $w_H(f)$ satisfying the inequalities $2^{n-2} < w_H(f) \leq \lfloor 2^{n-1} - 2^{\frac{n}{2}-1} \rfloor$ and we characterize those coset leaders in the well known classes of direct sums of monomial functions and of Maiorana-McFarland functions. Note that when a class is not a union of cosets of $RM(1,n)$ (as the class of direct sums of monomial functions, for instance), then for every function f in this class, we need to study the coset leaders in the coset $f + RM(1,n)$ and when a class is a union of cosets of $RM(1,n)$ (as the class of Maiorana-McFarland functions), it suffices to study the coset leaders in this class.
The paper is organized as follows. Section 2 recalls the background useful for our constructions and characterizations. Section 3 is devoted to the properties of coset leaders and to the study of the operations that allow to construct some of them. In Section 4, we characterize the coset leaders in specific classes.

2 Preliminaries

A function from $\mathbb{F}_2^n$ to $\mathbb{F}_2$ is called a Boolean function on $\mathbb{F}_2^n$, or an n-variable Boolean function, or a Boolean function in dimension n. The set $2^{\mathbb{F}_2^n}$ of all Boolean functions on $\mathbb{F}_2^n$ is denoted by B_n. Boolean functions can be expressed in different ways, each ensuring uniqueness. The most used one is by their multi-variate polynomial expression called the algebraic normal form (in brief, ANF) defined as follows:

Definition 1 *We call Algebraic Normal Form of a Boolean function f its n-variable representation over $\mathbb{F}_2$ belonging to $\mathbb{F}_2[x_1, \ldots, x_n]/(x_1^2 + x_1, \ldots, x_n^2 + x_n)$:*

$$f(x) = \sum_{I \subseteq \{1,2,\ldots,n\}} a_I \Big(\prod_{j \in I} x_j \Big) = \sum_{I \subseteq \{1,2,\ldots,n\}} a_I \, x^I$$

where $a_I \in \mathbb{F}_2$.

We also have the possibility of representing Boolean function by their truth-tables, that we will not consider in this work, and by their trace representation, that we shall describe

later. We will treat a function $f \in B_n$ as a vector of length 2^n and as a polynomial in n variables $x_l, \ldots, x_n$. The *algebraic degree* of a Boolean function f, denoted by $deg(f)$, is the degree of its ANF (see [11]).
For every binary vector $x \in \mathbb{F}_2^n$, the Hamming weight $w_H(x)$ of x being the number of its non zero coordinates (*i.e.* the size of the set $\{i \in N / x_i \neq 0\}$, called the support of x, where N denotes the set $\{1, \ldots, n\}$), the *Hamming weight* $w_H(f)$ of a Boolean function f on $\mathbb{F}_2^n$ is also the size of the support of the function denoted by $supp(f)$, *i.e.* of the set $\{x \in \mathbb{F}_2^n / f(x) = 1\}$ (the set $\{x \in \mathbb{F}_2^n / f(x) = 0\}$ being called the cosupport of the fonction f, and denoted by $cosupp(f)$). The *Hamming distance* between two Boolean functions f and g equals the Hamming weight of their sum, that is, $|\{x \in \mathbb{F}_2^n; f(x) \neq g(x)\}|$.
We define in what follows the notions of affine equivalence, of affine invariance and some notation useful in Section 3 and in Section 4.

Definition 2 *Two Boolean functions f and g are said affinely equivalent if there exists L, an affine automorphism of $\mathbb{F}_2^n$, such that $f = g \circ L$ where $\circ$ is the operation of composition. If L is a simple permutation of the input bits, then f and g are called permutation-equivalent.*
Two Boolean functions f and g are said EA-equivalent if g is affinely equivalent to the sum of f and an affine function.

Recall that an affine automorphism of $\mathbb{F}_2^n$ is a function $L : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} \mapsto M \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix}$
where M is a nonsingular $n \times n$ matrix.
A parameter associated to a function is called an affine invariant if it is preserved by affine equivalence. For instance, the Hamming weight and the algebraic degree are affine invariants.

Notation 1 *Let f and g be two Boolean functions.*
The notation $f \sim g$ will be used for "f and g are affinely equivalent".
We shall denote by $Var(f)$ the set $\{i \mid x_i$ appears in the ANF of $f\}$, which is not preserved by affine equivalence.

A class of Boolean functions in even dimension plays an important role thanks to its exceptional properties and its relation with design theory, cryptography, coding theory and sequences for telecommunications:

Definition 3 *[15, 7, 14] A Boolean function over $\mathbb{F}_2^n$ (n even) is bent if its Hamming distance to the set of all n-variable affine Boolean functions (the nonlinearity of f) equals $2^{n-1} - 2^{n/2-1}$ (which is optimal).*

Proposition 1 *[15, 7, 14] An n-variable Boolean function is bent if and only if its Hamming distance to any affine function equals $2^{n-1} \pm 2^{\frac{n}{2}-1}$. In particular, if f is a bent Boolean function over $\mathbb{F}_2^n$, then $w_H(f) = 2^{n-1} \pm 2^{\frac{n}{2}-1}$. If f is quadratic, the converse is true.*
Bent functions are those functions such that, for every non zero vector $a \in \mathbb{F}_2^n$, the derivative $D_a F = F(x) + F(x + a)$ is balanced.

We shall use the notion of Fourier and Walsh transform defined as follows:

Definition 4 *The Fourier transform of a function f over $\mathbb{F}_2^n$ and valued in $\mathbb{Z}$ is denoted by $\widehat{f}$ and defined as:*

$$\widehat{f}(u) = \sum_{x \in \mathbb{F}_2^n} f(x)(-1)^{u \cdot x} \text{ for all } u \in \mathbb{F}_2^n,$$

where " $\cdot$ " is some chosen inner product, that is, where $x \cdot y$ is a bilinear form such that $x \cdot y = 0$ for every $y \in \mathbb{F}_2^n$ if and only if $x = 0$ (i.e. the only element orthogonal to $\mathbb{F}_2^n$ is 0). For a Boolean function f, we obtain by considering it as valued in $\{0, 1\} \subset \mathbb{Z}$:

$$\widehat{f}(u) = \sum_{x \in supp(f)} (-1)^{u \cdot x} \text{ for all } u \in \mathbb{F}_2^n.$$

The Walsh transform of f, denoted by W_f, is the Fourier transform of the sign function $f_\chi(x) = (-1)^{f(x)}$:

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

These transforms satisfy the so-called *inverse Fourier formulas* $\sum_{u \in \mathbb{F}_2^n} \widehat{f}(u)(-1)^{u \cdot x} = 2^n f(x)$ and $\sum_{u \in \mathbb{F}_2^n} W_f(u)(-1)^{u \cdot x} = 2^n (-1)^{f(x)}$ and the *Parseval relation* $\sum_{u \in \mathbb{F}_2^n} W_f^2(u) = 2^{2n}$.

Bent functions are those functions whose Walsh transform takes the values $\pm 2^{n/2}$ only. A class of n-variable Boolean functions which generalizes bent functions is the class of plateaued functions, whose Walsh transform takes only the values 0 and $\pm\mu$, where μ is necessarily a power of 2, say $\mu = 2^r$, with $r \geq \frac{n}{2}$ if n is even and $r \geq \frac{n+1}{2}$ if n is odd (this positive number μ is called the *amplitude* of the plateaued function, see [5]).

Let f be an n-variable Boolean function. It is clear that $(-1)^{f(x)} = 1 - 2f(x)$ which implies

$$\widehat{f}(u) = 2^{n-1}\delta_0 - \frac{1}{2} W_f(u), \tag{1}$$

where δ_0 is the Dirac function at the zero vector. In particular, $u = 0$ yields:

$$w_H(f) = 2^{n-1} - \frac{1}{2} W_f(0). \tag{2}$$

If f is the direct sum of two functions, that is, $f(x,y) = f_1(x) + f_2(y)$, where $x \in \mathbb{F}_2^n$ and $y \in \mathbb{F}_2^m$, then we have:

$$W_f(a,b) = W_{f_1}(a)W_{f_2}(b). \tag{3}$$

In particular, if $f_1(x) = \prod_{i=1}^{n} x_i$ is the monomial function of degree n (that is, equals the Dirac function at the all-1 vector) and if f_2 is the monomial function of degree m, then $\widehat{f_1}(a) = (-1)^{w_H(a)}$, $\widehat{f_2}(b) = (-1)^{w_H(b)}$ and therefore the direct sum f satisfies:

$$W_f(a,b) = (2^n\delta_0(a) - 2(-1)^{w_H(a)})(2^m\delta_0(b) - 2(-1)^{w_H(b)}). \tag{4}$$

If f is the direct product of two functions, that is, $f(x,y) = f_1(x)f_2(y)$, then we have:

$$\begin{aligned}
&\widehat{f}(a,b) = \widehat{f_1}(a)\widehat{f_2}(b), \text{ i.e.,}\\
&2^{n+m+1}\delta_0(a,b) - 2W_f(a,b) = (2^n\delta_0(a) - W_{f_1}(a))(2^m\delta_0(b) - W_{f_2}(b)), \text{ i.e.,}\\
&2W_f(a,b) = 2^{n+m}\delta_0(a,b) + 2^n\delta_0(a)W_{f_2}(b) + 2^m\delta_0(b)W_{f_1}(a) - W_{f_1}(a)W_{f_2}(b). \quad (5)
\end{aligned}$$

The nonlinearity $nl(f)$ of a Boolean function f over $\mathbb{F}_2^n$ is the minimum Hamming distance $d_H(f,h) = |\{x \in \mathbb{F}_2^n; f(x) \neq h(x)\}|$ between f and affine functions h (in other words, the distance from f to $RM(1,n)$). We have:

$$nl(f) = 2^{n-1} - \frac{1}{2}\max_{a\in\mathbb{F}_2^n} | W_f(a) | . \tag{6}$$

Thanks to the Parseval Relation, the maximum of $W_f^2(a)$ is larger than or equal to its arithmetic mean $\frac{2^{2n}}{2^n} = 2^n$, and we have then the so-called covering radius bound:

$$nl(f) \leq 2^{n-1} - 2^{n/2-1}.$$

3 Properties of coset leaders

In this paper, since we study only the coset leaders of the first order Reed-Muller codes, we shall omit specifying "of $RM(1,n)$" when speaking of these coset leaders.
Let us recall first the coset leader definition.

Definition 5 *An n-variable Boolean function f with $n \geq 2$ is called a* coset leader *of the first order Reed-Muller code $RM(1,n)$ if for all $l \in RM(1,n)$, $w_H(f+l) \geq w_H(f)$.*
By abuse of language, given any Boolean function f, we shall call "coset leaders of f" the coset leaders in the coset $f + RM(1,n)$.

This latter expression will ease our presentation when we shall study classes of functions that are not unions of cosets of $RM(1,n)$. We have the following easy result:

Lemma 1 *An n variable Boolean function f is a coset leader of the first order Reed-Muller code $RM(1,n)$ if and only if $nl(f) = w_H(f)$, that is, $W_f(0) = \max_{a\in\mathbb{F}_2^n} \mid W_f(a) \mid$, or equivalently, $W_f(0) \geq |W_f(a)|$ for all $a \in \mathbb{F}_2^n$, or still equivallently $\widehat{f}(0) \leq 2^{n-1} - \max_{a\neq 0} |\widehat{f}(a)|$.*

Proof. The first part of the statement is by definition of the nonlinearity. The rest is a direct consequence of Relations (2) and (6). □

Given a Boolean function f and a vector a, denoting the function $a \cdot x$ by $\ell_a(x)$, we have $W_{f+\ell_a}(0) = W_f(a)$ and $W_{f+\ell_a+1}(0) = -W_f(a)$; then we have:

Lemma 2 *For every n-variable Boolean function f, every vector a and every bit ϵ, the function $f + \ell_a + \epsilon$ is a coset leader if and only if $|W_f(a)|$ is maximal over $\mathbb{F}_2^n$ and either $W_f(a) > 0$ and $\epsilon = 0$, or $W_f(a) < 0$ and $\epsilon = 1$.*

Remark 1 *In the case of a bent function f, we have $\max_{a\in\mathbb{F}_2^n} \mid W_f(a) \mid= 2^{\frac{n}{2}}$, and by Lemma 1, the bent coset leaders of $RM(1,n)$ are all the bent functions of Hamming weight $2^{n-1} - 2^{\frac{n}{2}-1}$. Note that we have $2^{n-1} - 2^{\frac{n}{2}-1} > 2^{n-2}$ for any $n \geq 4$.*

The property of being a coset leader is an affine invariant:

Lemma 3 *Let n be a positive integer, and let f and g be two n-variable Boolean functions with $n \geq 2$ such that $f \sim g$. Then, f is a coset leader of $RM(1,n)$ if and only if g is also a coset leader of $RM(1,n)$.*

Proof. Let f be a coset leader in n variables and let L be an affine automorphism of $\mathbb{F}_2^n$ such that $g = f \circ L$. We have that g is not a coset leader if and only if there exists an affine function l such that $w_H(f \circ L + l) < w_H(f \circ L)$, that is, $w_H((f + l \circ L^{-1}) \circ L) < w_H(f \circ L)$ or equivalently $w_H(f + l \circ L^{-1}) < w_H(f)$, that is, f is not a coset leader. □

As we said in the introduction, by the covering radius bound, the nonlinearity of a function can not be larger than $\lfloor 2^{n-1} - 2^{\frac{n}{2}-1} \rfloor$, that is, the Hamming weight of a coset leader can not be larger than $\lfloor 2^{n-1} - 2^{\frac{n}{2}-1} \rfloor$. We have the following characterization (which excludes functions of weight at most 2^{n-2} and bent functions since the case of these two categories of functions has been completely clarified):

Proposition 2 *Let f be a non-bent n-variable Boolean function of Hamming weight larger than 2^{n-2}. For being a coset leader of $RM(1,n)$, f needs to satisfy $w_H(f) = 2^{n-2} + e$, where $1 \leq e < 2^{n-2} - 2^{\frac{n}{2}-1}$, and then, f is a coset leader if and only if, for every non constant affine function l, we have $w_H(f(l+1)) \geq e$.*

Proof. By hypothesis we have $w_H(f) = 2^{n-2} + e$ with $e \geq 1$. The inequality $e < 2^{n-2} - 2^{\frac{n}{2}-1}$ is a necessary condition, as we observed above. Assuming it is satisfied, we have $w_H(f+1) = 2^n - w_H(f) = 2^n - (2^{n-2} + e) = 3 \cdot 2^{n-2} - e > 2^{n-2} + e = w_H(f)$, and the

inequality $w_H(f+l) < w_H(f)$ is then possible only if l is not constant. Hence, we have that f is a coset leader of $RM(1,n)$ if and only if, for every non constant affine function l, we have $w_H(f+l) \geq w_H(f)$, that is, $w_H(f\,l) \leq 2^{n-2}$ (since $w_H(f+l) = w_H(f)+2^{n-1}-2w_H(f\,l)$), or equivalently, $w_H(f\,(l+1)) \geq e$. □

Every n-variable Boolean function can be viewed as an $(n+1)$-variable Boolean function that does not depend on its last input variable. We show in the next proposition that this does not change its status of being (or not being) a coset leader.

Proposition 3 *Let f be an n-variable Boolean function. Then f is a coset leader of $RM(1,n)$ if and only if, seen as an $(n+1)$-variable function, it is a coset leader of $RM(1,n+1)$.*

Proof. For $k \geq n$, let us denote by $w_H^{(k)}(f)$ the Hamming weight of f seen as a k-variable Boolean function. Let f be a coset leader, then seeing f as in $n+1$ variables, we have $w_H^{(n+1)}(f) = 2w_H^{(n)}(f)$. For all affine function $l(x)$ in $n+1$ variables, if $n+1 \in Var(l)$ then $f+l$ is balanced in $\mathbb{F}_2^{n+1}$ and we have then $w_H^{(n+1)}(f+l) = 2^n > w_H^{(n+1)}(f)$. If $n+1 \notin Var(a)$, then $f+l$ does not depend on its last input coordinate and since f is a coset leader of $RM(1,n)$, we have $w_H^{(n+1)}(f+l) = 2w_H^{(n)}(f+l) \geq 2w_H^{(n)}(f) = w_H^{(n+1)}(f)$. Hence, seen as an $(n+1)$-variable function, f is a coset leader of $RM(1,n+1)$. The converse is straightforward. □

Consider the 5-variable functions $f_1 = x_1x_2x_3 + x_4x_5$ and $f_2 = x_4x_5 + x_5$. According to Relation (4), the Walsh transform of f_1 is valued in $\{\pm 4, \pm 12\}$ and we have $W_{f_1}(0) = 12$, meaning that f_1 is a coset leader, while f_2 is also a coset leader of $RM(1,5)$ since $w_H(f_2) = 2^3$ (recall that any Boolean function of Hamming weight at most 2^{n-2} is a coset leader). But the sum $f_1 + f_2 = x_1x_2x_3 + x_5$ is a balanced function meaning that it is not a coset leader of $RM(1,5)$. The following lemma gives a sufficient condition under which the sum of two coset leaders is a coset leader.

Lemma 4 *Let f_1 and f_2 be two coset leaders of $RM(1,n)$. If, for every affine function l, we have:*

$$\min\{w_H(lf_1), w_H(lf_2)\} \leq 2w_H(lf_1f_2),$$

then, $f_1 + f_2$ is a coset leader of $RM(1,n)$.

Proof. For every affine function l, the inequality $w_H(f_1 + f_2 + l) = w_H(f_1 + l) + w_H(f_2) - 2w_H((f_1 + l)f_2) \geq w_H(f_1) + w_H(f_2) - 2w_H(f_1f_2) - 2w_H(lf_2) + 4w_H(lf_1f_2)$ holds since f_1 is a coset leader and $w_H((f_1 + l)f_2) = w_H(f_1f_2) + w_H(lf_2) - 2w_H(lf_1f_2)$. Since $w_H(f_1) + w_H(f_2) - 2w_H(f_1f_2) = w_H(f_1 + f_2)$, then we have $w_H(f_1 + f_2 + l) - w_H(f_1 + f_2) \geq 4w_H(lf_1f_2) - 2w_H(lf_2)$. The proof ends by observing that f_2 is a coset leader and then we have also $w_H(f_1 + f_2 + l) - w_H(f_1 + f_2) \geq 4w_H(lf_1f_2) - 2w_H(lf_1)$ for all affine function

l. □

It is difficult to study the general structure of the set of coset leaders, for instance by determining the most general operations that are internal to this class. But some particular operations behave well with respect to coset leaders. For instance, we saw that the sum of two coset leaders is in general not a coset leader, but the situation is different when the sum is direct:

Proposition 4 *The direct sum of two coset leaders is a coset leader.*

Proof. Given two coset leaders f_1 and f_2 in n and m variables, respectively, we have according to Lemma 1 that $W_{f_1}(0) = \max_{a\in\mathbb{F}_2^n} \mid W_{f_1}(a) \mid$ and $W_{f_2}(0) = \max_{b\in\mathbb{F}_2^m} \mid W_{f_2}(b) \mid$. According to Relation (3), we decuce then $W_f(0,0) = \max_{a\in\mathbb{F}_2^n, b\in\mathbb{F}_2^m} \mid W_f(a,b) \mid$. This completes the proof. □

Note that such a direct sum can have Hamming weight larger than 2^{n+m-2}.

Conversely to Proposition 4, if the direct sum of two general Boolean functions is a coset leader, then these functions are either both coset leaders, or they are both the complements of coset leaders, since the maximum of the product of the two independent non-negative sequences $|W_{f_1}(a)|$ and $|W_{f_2}(b)|$ is achieved when each of these sequences reaches its maximum. Let us see that the situation is different with the direct product. We begin with a characterization.

Proposition 5 *The direct product of an n-variable function f_1 and an m-variable function f_2 is a coset leader if and only if:*

$$\begin{aligned}&\forall a \in \mathbb{F}_2^n \setminus \{0\}, \forall b \in \mathbb{F}_2^m \setminus \{0\},\\ &\quad 2^{n+m} + 2^n W_{f_2}(0) + 2^m W_{f_1}(0) - W_{f_1}(0)W_{f_2}(0) \geq\\ &\quad \max\Big((2^n - W_{f_1}(0))|W_{f_2}(b)|\ ,\ (2^m - W_{f_2}(0))|W_{f_1}(a)|\ ,\ |W_{f_1}(a)W_{f_2}(b)|\Big). \qquad (7)\end{aligned}$$

Proof. Let f be the direct product of f_1 and f_2. Relation (5) and the facts that $2^n - W_{f_1}(0) \geq 0$ and $2^m - W_{f_2}(0) \geq 0$ directly imply that f is a coset leader if and only if we have Inequality (7) (in which the three numbers whose maximum is taken correspond respectively to "$a = 0, b \neq 0$"; "$a \neq 0, b = 0$" and "$a \neq 0, b \neq 0$"). □

We first deduce that:

Corollary 1 *The direct product of two coset leaders is a coset leader.*

Proof. Since $W_{f_1}(0) \geq 0$ and $W_{f_2}(0) \geq 0$, we have:

$$\begin{aligned}&2^{n+m} + 2^n W_{f_2}(0) + 2^m W_{f_1}(0) - W_{f_1}(0)W_{f_2}(0) =\\ &2^m(2^n + W_{f_1}(0)) + (2^n - W_{f_1}(0))W_{f_2}(0) \geq\\ &(2^n - W_{f_1}(0))|W_{f_2}(0)|,\end{aligned}$$

and

$$2^{n+m} + 2^n W_{f_2}(0) + 2^m W_{f_1}(0) - W_{f_1}(0)W_{f_2}(0) =$$
$$2^n(2^m + W_{f_2}(0)) + (2^m - W_{f_2}(0))W_{f_1}(0) \geq$$
$$(2^m - W_{f_2}(0))|W_{f_1}(0)|.$$

Hence, since f_1 and f_2 are coset leaders, we have:

$$2^{n+m} + 2^n W_{f_2}(0) + 2^m W_{f_1}(0) - W_{f_1}(0)W_{f_2}(0) \geq (2^n - W_{f_1}(0))|W_{f_2}(b)|$$

and

$$2^{n+m} + 2^n W_{f_2}(0) + 2^m W_{f_1}(0) - W_{f_1}(0)W_{f_2}(0) \geq (2^m - W_{f_2}(0))|W_{f_1}(a).$$

We have also $2^{n+m} + 2^n W_{f_2}(0) + 2^m W_{f_1}(0) - 2W_{f_1}(0)W_{f_2}(0) = 2^{n+m} + (2^n - W_{f_1}(0))W_{f_2}(0) + (2^m - W_{f_2}(0))W_{f_1}(0) \geq 0$ which implies $2^{n+m} + 2^n W_{f_2}(0) + 2^m W_{f_1}(0) - W_{f_1}(0)W_{f_2}(0) \geq |W_{f_1}(a)W_{f_2}(b)|$ for every nonzero a and b, and f is a coset leader, according to Proposition 5. □

But an n-variable coset leader having Hamming weight less than 2^{n-1} and an m-variable coset leader having Hamming weight less than 2^{m-1}, the Hamming weight of their direct product, equal to the product of these Hamming weights, is smaller than 2^{n+m-2}, and the result of Corollary 1 is then trivial.

Let us see now that there are functions f_1 and f_2 for which Proposition 5 applies and such that f has Hamming weight larger than 2^{n+m-2}. In fact, we shall see in the next corollary that a direct product is always a coset leader, unless one of the functions has a very low nonlinearity. Indeed, let f_1 have nonlinearity $2^{n-1} - \frac{N}{2}$ and f_2 have nonlinearity $2^{m-1} - \frac{M}{2}$, for some integers N and M. Then we have $-N \leq W_{f_1}(a) \leq N$, for every a and $-M \leq W_{f_2}(b) \leq M$, for every b. Hence, we have $2^{n+m} + 2^n W_{f_2}(0) + 2^m W_{f_1}(0) - W_{f_1}(0)W_{f_2}(0) \geq 2^{n+m} - 2^n M - 2^m N - NM$ and $\max\Big((2^n - W_{f_1}(0))|W_{f_2}(b)|\ ,\ (2^m - W_{f_2}(0))|W_{f_1}(a)|\ ,\ |W_{f_1}(a)W_{f_2}(b)|\Big) \leq \max\Big((2^n + N)M\ ,\ (2^m + M)N\ ,\ NM\Big)$. We deduce, taking $N = 2^n\lambda$ and $M = 2^m\mu$:

Corollary 2 *Let f_1 be any n-variable function of nonlinearity $2^{n-1}(1-\frac{\lambda}{2})$, where $0 < \lambda \leq 1$ and let f_2 be any m-variable function of nonlinearity $2^{m-1}(1-\frac{\mu}{2})$, where $0 < \mu \leq 1$. Assume that:*

$$\max(\lambda + 2\mu + 2\lambda\mu\ ,\ 2\lambda + \mu + 2\lambda\mu) \leq 1. \tag{8}$$

Then the direct product of f_1 and f_2 is a coset leader.

Condition (8) is rather weak, since if both λ and μ are not larger than $\frac{-3+\sqrt{17}}{4} \approx .28$ then it is satisfied. Indeed, $\lambda \leq \frac{-3+\sqrt{17}}{4}$ implies $3\lambda + 2\lambda^2 \leq 1$, and then $\lambda \leq \frac{-3+\sqrt{17}}{4}$ and $\mu \leq \frac{-3+\sqrt{17}}{4}$

imply $\max(2\mu + \lambda + 2\lambda\mu \ , \ \mu + 2\lambda + 2\lambda\mu) \leq 1$. The only pairs (f_1, f_2) that do not satisfy Condition (8) are then such that f_1 has very low nonlinearity (not much larger than $(.72) \cdot 2^{n-1}$) or f_2 has very low nonlinearity (not much larger than $(.72) \cdot 2^{m-1}$).

Let us now visit a third secondary construction of Boolean functions: given three functions f_1, f_2 and f_3 in n, m and r variables, respectively, we call the *direct majority* of f_1, f_2 and f_3 the Boolean function which takes value 1 if and only if a majority of these three functions takes value 1, that is: $f(x, y, z) = f_1(x)f_2(y) + f_1(x)f_3(z) + f_2(y)f_3(z)$. It is known from [3] and recalled in [4, Proposition 85], that given three functions h_1, h_2, h_3 over $\mathbb{F}_2^N$, we have $W_{h_1} + W_{h_2} + W_{h_3} = W_s + 2W_f$, where $s = h_1 + h_2 + h_3$ and $f = h_1h_2 + h_1h_3 + h_2h_3$. We can apply this to the functions $h_1(x, y, z) = f_1(x), h_2(x, y, z) = f_2(y)$ and $h_3(x, y, z) = f_3(z)$. We have then $W_{h_1}(a, b, c) = 2^{m+r}W_{f_1}(a)\delta_0(b)\delta_0(c)$, $W_{h_2}(a, b, c) = 2^{n+r}W_{f_2}(b)\delta_0(a)\delta_0(c)$ and $W_{h_3}(a, b, c) = 2^{n+m}W_{f_3}(c)\delta_0(a)\delta_0(b)$, and since we know by iterating (3) that the direct sum s satisfies $W_s(a, b, c) = W_{f_1}(a)W_{f_2}(b)W_{f_3}(c)$, we have then:

$$\forall a \in \mathbb{F}_2^n, \forall b \in \mathbb{F}_2^m, \forall c \in \mathbb{F}_2^r, \quad 2W_f(a, b, c) = \tag{9}$$

$$2^{m+r}W_{f_1}(a)\delta_0(b)\delta_0(c) + 2^{n+r}W_{f_2}(b)\delta_0(a)\delta_0(c) + 2^{n+m}W_{f_3}(c)\delta_0(a)\delta_0(b) - W_{f_1}(a)W_{f_2}(b)W_{f_3}(c).$$

We deduce:

Proposition 6 *The direct majority f of three Boolean functions f_1, f_2 and f_3 in n, m and r variables, respectively, is a coset leader if and only if, for every nonzero a, b, c, we have:*

$$2^{m+r}W_{f_1}(0) + 2^{n+r}W_{f_2}(0) + 2^{n+m}W_{f_3}(0) - W_{f_1}(0)W_{f_2}(0)W_{f_3}(0) \geq$$

$$\max\Big(|W_{f_1}(a)|(2^{m+r} - W_{f_2}(0)W_{f_3}(0)), |W_{f_2}(b)|(2^{n+r} - W_{f_1}(0)W_{f_3}(0)),$$
$$|W_{f_3}(c)|(2^{n+m} - W_{f_1}(0)W_{f_2}(0)), |W_{f_1}(0)W_{f_2}(b)W_{f_3}(c)|,$$
$$W_{f_1}(a)W_{f_2}(0)W_{f_3}(c)|, |W_{f_1}(a)W_{f_2}(b)W_{f_3}(0)|, |W_{f_1}(a)W_{f_2}(b)W_{f_3}(c)|\Big).$$

Corollary 3 *If f_1, f_2 and f_3 are three coset leaders in n, m and r variables, respectively, then their direct majority is a coset leader.*

Proof. According to Proposition 6, the direct majority of f_1, f_2 and f_3 is a coset leader if and only if, for every triple $(a, b, c) \in \mathbb{F}_2^n \times \mathbb{F}_2^m \times \mathbb{F}_2^r$ having at most one zero term, we have:

$$2^{m+r}W_{f_1}(0) + 2^{n+r}W_{f_2}(0) + 2^{n+m}W_{f_3}(0) - W_{f_1}(0)W_{f_2}(0)W_{f_3}(0) \geq |W_{f_1}(a)W_{f_2}(b)W_{f_3}(c)|.$$

Hence, if $2^{m+r}W_{f_1}(0) + 2^{n+r}W_{f_2}(0) + 2^{n+m}W_{f_3}(0) \geq 2W_{f_1}(0)W_{f_2}(0)W_{f_3}(0)$, then f is a coset leader. But this condition is always satisfied since we have $W_{f_1}(0) \leq 2^n < \sqrt{\frac{3}{2}}\,2^n$, $W_{f_2}(0) \leq 2^m < \sqrt{\frac{3}{2}}\,2^m$ and $W_{f_3}(0) \leq 2^r < \sqrt{\frac{3}{2}}\,2^r$. □

Remark 2 *More generally, let g be any r-variable Boolean function and $f_1, \ldots, f_r$ be Boolean functions in $n_1, \ldots, n_r$ variables, respectively. Then we can consider the Boolean function $f(x^{(1)}, \ldots, x^{(r)}) = g(f_1(x^{(1)}), \ldots, f_r(x^{(r)}))$, where $x^{(1)} \in \mathbb{F}_2^{n_1}, \ldots, x^{(r)} \in \mathbb{F}_2^{n_r}$ are "disjoint" (that is, independent) variables. In the case of the direct sum, we have $r = 2$ and $g(y_1, y_2) = y_1 + y_2$ and in the case of the direct product, we have $r = 2$ and $g(y_1, y_2) = y_1 y_2$. We have:*

$$\begin{aligned} W_f(a^{(1)}, \ldots, a^{(r)}) &= \sum_{(x^{(1)}),\ldots,x^{(r)}) \in \mathbb{F}_2^{n_1+\cdots+n_r}} (-1)^{g(f_1(x^{(1)}),\ldots,f_r(x^{(r)}))+\sum_{i=1}^r a^{(i)}\cdot x^{(i)}} \\ &= 2^{-r} \sum_{\substack{(x^{(1)}),\ldots,x^{(r)}) \in \mathbb{F}_2^{n_1+\cdots+n_r} \\ y,z \in \mathbb{F}_2^r}} (-1)^{g(y)+\sum_{i=1}^r a^{(i)}\cdot x^{(i)}+\sum_{i=1}^r z_i(y_i+f_i(x^{(i)}))} \\ &= 2^{-r} \sum_{z \in \mathbb{F}_2^r} W_g(z) \prod_{i=1}^r W_{z_i f_i}(a^{(i)}). \end{aligned} \tag{10}$$

In the case of the direct sum, we have $W_g(y_1, y_2) = \sum_{y \in \mathbb{F}_2^2} (-1)^{y_1+y_2+z_1 y_1+z_2 y_2} = 4\delta_{(1,1)}(z)$, where $\delta_{(1,1)}(z)$ equals 1 if $(z_1, z_2) = (1, 1)$ and equals 0 otherwise. Relation (10) *gives* (3). *In the case of the direct product, we have $W_g(y_1, y_2) = \sum_{y \in \mathbb{F}_2^2} (-1)^{y_1 y_2+z_1 y_1+z_2 y_2} = 2(-1)^{z_1 z_2}$ and this gives* (5).

Note that the only other symmetric function in 2 variables is $g(y_1, y_2) = y_1 y_2 + y_1 + y_2$, which is not interesting since applying such g to (f_1, f_2) gives the complement of what we get by applying the direct product to $(f_1 + 1, f_2 + 1)$; we have seen that such latter functions are most often coset leaders; their complements are then not.

In the case of the direct majority, we have $g(y_1, y_2, y_3) = y_1 y_2 + y_1 y_3 + y_2 y_3$. The Walsh transform of g is known from [9], but what we have recalled above gives a fast way for calculating it: denoting by l_1, l_2 and l_3 the three coordinate functions over $\mathbb{F}_2^3$ and by s_1 the Boolean function $l_1 + l_2 + l_3$, we have $W_{l_1}(z) = \sum_{y \in \mathbb{F}_2^3} (-1)^{(z_1+1)y_1+z_2 y_2+z_3 y_3} = 8\,\delta_{(1,0,0)}(z)$, $W_{l_2}(z) = 8\,\delta_{(0,1,0)}(z)$, $W_{l_3}(z) = 8\,\delta_{(0,0,1)}(z)$, $W_{s_1}(z) = 8\,\delta_{(1,1,1)}(z)$ and $W_g(z) = 4\,(\delta_{(1,0,0)}(z) + \delta_{(0,1,0)}(z) + \delta_{(0,0,1)}(z)) - \delta_{(1,1,1)}(z))$. Note that the Walsh transform of g vanishes on a whole hyperplane[1]. *Relation* (10) *gives then* (9).

We can also take $g(y_1, y_2, y_3) = y_1 y_2 y_3$. Then, similarly to the case of the direct product of two functions, we have, thanks to the relation $\widehat{f}(a, b, c) = \widehat{f_1}(a)\widehat{f_2}(b)\widehat{f_3}(c)$, that such a direct product of three functions is a coset leader under the sufficient condition that none of the three functions is of very weak nonlinearity.

Other examples of g functions could be tried but either they would no longer be symmetric, or they would be in at least 4 variables and the number of terms in Relation (10) *would increase.*

[1] This is equivalent to saying that g has a linear structure, see [4, Proposition 29]; here, we have $g(x + 1, y + 1, z + 1) = g(x, y, z) + 1$.

4 Coset leaders in classical classes of Boolean functions

4.1 Direct sums of monomials

In this subsection and the two next ones, "$\cdot$" is the usual inner product: $a \cdot x = \sum_{i=1}^{n} a_i x_i$.

Definition 6 *Let f be a non constant Boolean function in n variables. We call f a direct sum of monomials (in brief, DSM) if the following holds for its ANF $f(x) = \sum_{I \subseteq \{1,2,\ldots,n\}} a_I x^I$:*

$$\forall\, I, J \text{ such that } a_I = a_J = 1,\ I \cap J \in \{\emptyset, I \cup J\}$$

Note that the class of DSM is not a union of cosets of the first order Reed-Muller code. We shall have then not only to determine those DSM functions that are coset leaders but also the sums of DSM and affine functions that are coset leaders.

We also need the following definition:

Definition 7 ***(Direct Sum Vector [13]).*** *Let f be a DSM function, its direct sum vector is the following sequence of length $k + 1 = deg(f) + 1$:*

$$m_f = [m_0, m_1, m_2, \ldots, m_k],$$

where m_i is the number of monomials of degree i in the ANF of f ($i \geq 0$):

$$m_i = |\{a_I = 1,\ \textit{such that } |I| = i\}|.$$

The function f associated to the direct sum vector $m_f = [m_0, m_1, m_2, \ldots, m_k]$, has $M = \sum_{i=0}^{k} m_i$ distinct monomials in its ANF and n variables with $n \geq \sum_{i=0}^{k} i m_i$.
Clearly, two DSM functions having the same direct sum vector and the same number of variables are permutation-equivalent and two DSM functions having the same part $[m_2, \ldots, m_k]$ and the same number of variables are EA-equivalent.

Remark 3 *The number of DSM functions of direct sum vector $[m_0, m_1, m_2, \ldots, m_k]$ equals*

$$2 \prod_{i=1}^{k} \left(\binom{n - (i-1)m_{i-1}}{i m_i} \times \frac{\binom{i m_i}{i}\binom{i m_i - i}{i} \cdots \binom{2i}{i}\binom{i}{i}}{m_i!} \right).$$

Indeed, the number of functions equal to the direct sum of m_1 monomials of degree 1 clearly equals $\binom{n}{m_1} = \binom{n}{m_1} \frac{\binom{m_1}{1}\binom{m_1-1}{1} \cdots \binom{2}{1}\binom{1}{1}}{m_1!}$; such function being chosen, the number of functions equal to the direct sum of m_2 monomials of degree 2 having disjoint variables with the m_1 already chosen monomials equals the number of choices of $2m_2$ variables among the $n - m_1 = n - (2-1)m_{2-1}$ remaining variables, times the number of direct sums of monomials of degree 2 in these variables, which yields $\binom{n-(2-1)m_{2-1}}{2m_2} \times \frac{\binom{2m_2}{2}\binom{2m_2-2}{2} \cdots \binom{2\times 2}{2}\binom{2}{2}}{m_2!}$ possibilities . One continues the process until $i = k$ and obtains the result by multiplying results obtained for each $i = 1 \cdots k$.

In the following, given a Boolean function f in n variables, the Walsh spectrum $\{W_f(a),\ a \in \mathbb{F}_2^n\}$ of f is denoted by $WS(f)$ and the set $\{|W_f(a)|,\ a \in \mathbb{F}_2^n\}$ (called the extended Walsh spectrum of f) is denoted by $EWS(f)$. Note that Lemma 1 is equivalent to f is a coset leader if and only if $W_f(0) = \max EWS(f)$.

Remark 4 *The Walsh spectrum of DSM functions has been covered by the paper [6] when determining their nonlinearity, but the details we need here were skipped. Consider a DSM function f of direct sum vector $m_f = [m_0, m_1, m_2, \ldots, m_k]$. Note that if $m_1 \neq 0$, then according to Relation (4), we have $W_f(0) = 0$, and f having weight 2^{n-1}, it can not then be a coset leader. So all DSM coset leaders satisfy $m_1 = 0$. Note also that if $n > \sum_{i=1}^{k} im_i$, then for all vector $a \in \mathbb{F}_2^n$ such that there exists $i_0 \in supp(a) \setminus Var(f)$, we have*

$$W_f(a) = \sum_{x_{i_0} \in \mathbb{F}_2} (-1)^{x_{i_0}} \sum_{\substack{x \in \mathbb{F}_2^n \\ x_{i_0}=0}} (-1)^{f(x)+a\cdot x} = 0,$$

and the coset leaders of f are necessarily without any such term x_{i_0}. From Relation (1) and the fact that a function of Hamming weight 1 has it Fourier transform valued in ± 1, the Walsh transform of each degree i monomial in a DSM function f is valued in $\{\pm 2, 2^i - 2\}$. The integer m_i being the number of degree i monomials in f, then given $a \in \mathbb{F}_2^n$, let $t_i \leq m_i$ be the number of degree i monomials in f whose Walsh transform at a equals $2^i - 2$, which means that $m_i - t_i$ is the number of degree i monomials in f whose Walsh Transform at a equals ± 2. Hence, if $m_0 = m_1 = 0$ then in the expression of $W_f(a)$, we have $2^i - 2$ raised at the power t_i for every i, and 2 raised at the power $p = \sum_{i=2}^{k}(m_i - t_i)$ (and we have $\sum_{i=2}^{k} t_i = \sum_{i=2}^{k} m_i - p$). Therefore, the Walsh transform of an n-variable DSM function f of direct sum vector $m_f = [m_0 = 0, m_1 = 0, m_2, \ldots, m_k]$ with $n > \sum_{i=1}^{k} im_i$ is valued in the following sets:

$$\{0\};$$

$$\left\{ \pm 2^{n-\sum_{i=2}^{k} im_i} \prod_{i=2}^{k} (2^i - 2)^{m_i} \right\};$$

$$\left\{ \pm 2^{n-\sum_{i=2}^{k} im_i+1} \prod_{i=2}^{k} (2^i - 2)^{t_i} where\ 0 \leq t_i \leq m_i;\ \sum_{i=1}^{k} t_i = \sum_{i=2}^{k} m_i - 1 \right\};$$

$$\left\{ \pm 2^{n-\sum_{i=2}^{k} im_i+2} \prod_{i=2}^{k} (2^i - 2)^{t_i} where 0 \leq t_i \leq m_i;\ \sum_{i=2}^{k} t_i = \sum_{i=2}^{k} m_i - 2 \right\};$$

$$\ldots$$

$$\left\{ \pm 2^{n-\sum_{i=2}^{k} im_i+\sum_{i=2}^{k} m_i-1} (2^j - 2),\ 2 \leq j \leq k \right\};$$

$$\left\{ \pm 2^{n-\sum_{i=2}^{k} im_i+\sum_{i=2}^{k} m_i} \right\}.$$

The union of these sets equals:

$$WS(f) = \left\{\pm 2^{n-\sum_{i=1}^{k} im_i+p} \prod_{i=2}^{k}(2^i-2)^{t_i} \text{ where } 0 \le t_i \le m_i,\ p \ge 0 \text{ and } p + \sum_{i=1}^{k} t_i = \sum_{i=1}^{k} m_i \right\} \cup \{0\}$$

and if $n = \sum_{i=1}^{k} im_i$ *this set becomes:*

$$WS(f) = \left\{ \pm 2^{p} \prod_{i=2}^{k}(2^i - 2)^{t_i} \text{ where } 0 \le t_i \le m_i,\ p \ge 0 \text{ and } p + \sum_{i=1}^{k} t_i = \sum_{i=1}^{k} m_i \right\}.$$

The value of $W_f(0)$ *among these is* $2^{n-\sum_{i=2}^{k} im_i} \prod_{i=2}^{k}(2^i-2)^{m_i}$ *which is maximal over* $EWS(f)$*, since for all* $i \ge 2$*,* $2^i - 2 \ge 2$*.*
Note that $WS(f)$ *for the case* $m_0 = 1$ *or* $m_1 \ne 0$ *is obtained easily by observing that* $f = h + b \cdot x + \epsilon$ *where* h *is a DSM with* $m_0 = m_1 = 0$*,* $b \cdot x$ *is the linear part of* f *and* $\epsilon = 0$ *or 1, which implies* $W_f(a) = (-1)^{\epsilon} W_h(a + b)$*.*

Recall that, by abuse of language, given a class C of Boolean functions, we call "coset leaders of the functions in C" the coset leaders of $RM(1,n)$ that belong to $\bigcup_{f \in C}(f + RM(1,n))$.

Theorem 1 *1) Any DSM function* f *in* n *variables with direct sum vector:*

$$m_f = [m_0, m_1, m_2, \ldots, m_k],$$

with $k \ge 2$ *and* $n \ge \sum_{i=0}^{k} im_i$*, is a coset leader of* $RM(1,n)$ *if and only if* $m_0 = m_1 = 0$*.*

2) The coset leaders of DSM functions are the functions $h + \ell_a + \epsilon$ *where:*

- h *is a DSM function with* $m_0 = m_1 = 0$*,*
- $\ell_a(x) = a \cdot x$ *is such that* $a(i) = 0$ *for every* $i \notin K_1 \cup \cdots \cup K_{m_2}$ *where* $K_1, \ldots, K_{m_2}$ *are the pairwise disjoint pairs of* $\{1, 2, \ldots, n\}$ *such that the degree* 2 *part of* h *is given by* $\sum_{i=1}^{m_2} \left(\prod_{j \in K_i} x_j \right)$ *and where we denote by* $a(i)$ *the vector* $(a_j)_{j \in K_i}$ *of length* $Card(K_i)$*,*
- $\epsilon = 0$ *if*

$$\prod_{\substack{i=1,\ldots,m_2 \\ a(i) \ne 0}} (-1)^{w_H(a(i))+1} = 1,$$

and $\epsilon = 1$ *otherwise.*

Proof. 1) Every monomial function of degree $i \ge 2$ in i variables is a coset leader, since its Hamming weight equals 1 and the coset of $RM(1,i)$ which contains it does not contain the 0 function. According to Proposition 4, every DSM such that $m_0 = m_1 = 0$ and

$n = \sum_{i=0}^{k} im_i$ is then a coset leader. If $m_0 = m_1 = 0$ and $n > \sum_{i=0}^{k} im_i$, then denoting by f' the same function as f viewed as in $\sum_{i=0}^{k} im_i$ variables, $WS(f) = \{W_f(a), a \in \mathbb{F}_2^n\}$ is the union of $\{0\}$ and of the set obtained by multiplying each value in $WS(f')$ by $2^{n-\sum_{i=0}^{k} im_i}$. Hence, $W_f(0)$ remains maximal in $EWS(f)$, that is, f is a coset leader.
Let us show now that these functions are the only coset leaders in the class of DSM functions: if $m_0 = 1$ and $m_1 = 0$, then $f = g + 1$ where g is a DSM function with $m_0 = m_1 = 0$ and is then a coset leader, which implies that f is not a coset leader; and if $m_1 > 0$, then according to Relation (3), we have $W_f(0) = 0$ meaning that f is balanced and then, it can not be a coset leader. This completes the proof of 1).
2) Given a DSM function f, any function $g \in f + RM(1, n)$ can be written as $g = h + \ell_a + \epsilon$ where $\epsilon = 0$ or 1 and h is the DSM function with $m_0 = m_1 = 0$ equal to f deprived of its degree 0 and 1 monomials. Let us denote by $K_1, \ldots, K_{m_2}, K_{m_2+1}, \ldots, K_{m_2+m_3}, \ldots, K_M$, where $M = \sum_{i=2}^{k} m_i$, the disjoint subsets of $\{1, \ldots, n\}$ such that the degree s part of h is given by $\sum_{i=m_2+\cdots+m_{s-1}+1}^{m_2+\cdots+m_s} \left(\prod_{j \in K_i} x_j \right)$ and let $K_{M+1} = \{1, \ldots, n\} \backslash (K_1 \cup \cdots \cup K_M)$ corresponding to the variables of h that do not appear in its ANF. The sets $K_1, \ldots, K_{M+1}$ form a partition of the set $\{1, 2, \ldots, n\}$ and we have $h = \sum_{i=1}^{M} \left(\prod_{j \in K_i} x_j \right)$ and $K_{M+1} \cap Var(h) = \emptyset$.
If $\epsilon = 0$, then $g = h + \ell_a$ has the same Walsh spectrum as h, and is then, according to Lemma 2, a coset leader if and only if $W_g(0) = W_h(a)$ is maximal over $EWS(h)$. This needs first that $supp(a) \subseteq Var(h)$, that is, $supp(a) \cap K_{M+1} = \emptyset$, since if there exists $i_0 \in supp(a) \setminus Var(h)$ then $W_h(a) = 0$ meaning that $W_h(a)$ is not optimal. Assuming that $supp(a) \subseteq Var(h)$, we have according to Relation (4) iterated:

$$W_h(a) = 2^{n-\sum_{i=2}^{k} im_i} \prod_{i=1}^{M} \left(2^{Card(K_i)} \delta_0(a(i)) - 2(-1)^{w_H(a(i))} \right) \tag{11}$$

and we have $W_h(a) = \max EWS(h)$ if and only if each factor $2^{Card(K_i)} \delta_0(a(i)) - 2(-1)^{w_H(a(i))}$ has maximal absolute value (which is equivalent to $a(i) = 0$ for all $i = m_2 + 1, \ldots, M$), and $\prod_{\substack{i=1,\ldots,m_2 \\ a(i) \neq 0}} (-1)^{w_H(a(i))+1} = 1$.
If $\epsilon = 1$, then we have the same situation but with $\prod_{\substack{i=1,\ldots,m_2 \\ a(i) \neq 0}} (-1)^{w_H(a(i))+1} = -1$. □

The next result is a characterization of coset leaders of DSM functions having Hamming weight larger than 2^{n-2}.

Lemma 5 *Let f be a DSM coset leader in n variables of direct sum vector $m_f = [m_0 = 0, m_1 = 0, m_2, \ldots, m_k]$ with $k \geq 2$ and $n \geq \sum_{i=1}^{k} im_i$. Then, $w_H(f) > 2^{n-2}$ if and only if*

$$\prod_{i=2}^{k} (2^i - 2)^{m_i} < 2^{\sum_{i=2}^{k} im_i - 1}$$

More generally, every coset leader $g = h + a \cdot x + \epsilon$ of a DSM function, defined in Theorem 1, where h is a DSM with $m_0 = m_1 = 0$, has Hamming weight larger than 2^{n-2} if and only if $\prod_{i=2}^{k} (2^i - 2)^{m_i} < 2^{\sum_{i=2}^{k} (im_i) - 1}$.

Proof. Recall from Remark 4 and Theorem 1 that $W_f(0) = 2^{n-\sum_{i=2}^{k} im_i} \prod_{i=2}^{k} (2^i - 2)^{m_i}$. Then, the inequality $\prod_{i=2}^{k} (2^i - 2)^{m_i} < 2^{\sum_{i=2}^{k} im_i - 1}$ is equivalent to $W_f(0) < 2^{n-1}$, that is, $w_H(f) > 2^{n-2}$.
The proof is completed by observing that every coset leader $g = h + a \cdot x + \epsilon$ of a DSM function (where h has no monomial of degree 1 or 2) is such that $WS(g) = (-1)^{\epsilon} WS(h)$, which implies $W_g(0) = W_h(0)$ since, both g and h being coset leaders, we have $W_g(0) = \max_u |W_g(u)|$ and $W_h(0) = \max_u |W_h(u)|$. □

Lemma 5 yields

Corollary 4 *Let f be a degree k DSM function in n variables of direct sum vector $m_f = [m_0 = 0, m_1 = 0, m_2, \ldots, m_k]$ with $k \geq 2$ and $n \geq \sum_{i=1}^{k} im_i$. Then we have:*

1) *f is a coset leader of $RM(1, n)$ of Hamming weight greater than 2^{n-2} if and only if $\sum_{i=2}^{k} m_i \log_2(\frac{2^{i-1}}{2^{i-1}-1}) > 1$, where $\log_2(x)$ is the binary logarithm of x.*

2) *if for all $i = 2, \ldots, k$, we have $m_i > \frac{1}{(k-1)\log_2(\frac{2^{i-1}}{2^{i-1}-1})}$, then this condition is satisfied.*

Proof. From Lemma 5, f is a coset leader of $RM(1, n)$ of weight greater than 2^{n-2} if and only if the inequality $\prod_{i=2}^{k} (2^i - 2)^{m_i} < 2^{\sum_{i=2}^{k} im_i - 1}$ holds. This inequality is equivalent to

$$\log_2\Big(\prod_{i=2}^{k} (2^i - 2)^{m_i}\Big) < \log_2\Big(\frac{1}{2}\prod_{i=2}^{k} 2^{im_i}\Big) \Leftrightarrow$$

$$\sum_{i=2}^{k} m_i \log_2(2^i - 2) < -1 + \sum_{i=2}^{k} im_i \Leftrightarrow$$

$$\sum_{i=2}^{k} m_i \log_2\Big(\frac{2^{i-1}}{2^{i-1} - 1}\Big) > 1,$$

which proves 1), and 2) is straightforward. □

Remark 5 *Corollary 4 allows to generate easily DSM coset leaders of $RM(1, n)$ with Hamming weight greater than 2^{n-2}. For instance for $n = 8$ all the coset leaders of $RM(1, 8)$ of weight greater than $2^6 = 64$ are known and we have:*

- *for $k = 2$ the possible direct sum vectors are* $[0, 0, 2], [0, 0, 3], [0, 0, 4]$.
- *for $k = 3$, the possible direct sum vectors are* $[0, 0, 1, 1], [0, 0, 2, 1], [0, 0, 1, 2]$.
- *for $k = 4$, the possible direct sum vectors are* $[0, 0, 1, 0, 1], [0, 2, 0, 1]$.
- *for $k = 5$, the unique direct sum vector is* $[0, 0, 1, 0, 0, 1]$

- *for $k = 6$, the unique direct sum vector is* $[0, 0, 1, 0, 0, 0, 1]$
- *for $k = 7$ or 8, there is no DSM coset leaders of $RM(1, 8)$ of weight greater than $2^6 = 64$.*

We shall provide in Corollary 5 an example of an infinite class of coset leaders whose Hamming weight is larger than 2^{n-2}. We need first the following result.

Lemma 6 *Let μ be a positive integer with $\mu \geq 3$. Setting $n_\mu = 2 + 3 + \cdots + \mu = \frac{(\mu-1)(\mu+2)}{2}$, we have*

$$\prod_{i=2}^{\mu}(2^i - 2) < 2^{n_\mu - 1}$$

Proof. We prove it by induction on μ. For $\mu = 3$, we have $n_3 = 5$ and $\prod_{i=2}^{3}(2^i - 2) = 12 < 2^{5-1} = 16$. Now assume that, for some $\mu \geq 3$, the inequality $\prod_{i=2}^{\mu}(2^i - 2) < 2^{n_\mu - 1}$ holds. By observing that $n_{\mu+1} = n_\mu + \mu + 1$, we have $\prod_{i=2}^{\mu+1}(2^i - 2) = (2^{\mu+1} - 2)\prod_{i=2}^{\mu}(2^i - 2) < 2^{\mu+1}\prod_{i=2}^{\mu}(2^i - 2)) < 2^{\mu+1}2^{n_\mu - 1} = 2^{n_{\mu+1}-1}$ which ends the proof. □

Corollary 5 *let μ and n be two integers such that $\mu \geq 3$ and $n = \frac{(\mu+2)(\mu-1)}{2}$. Every n-variable function f equivalent to the direct sum of $\mu-1$ monomials equal to $x_1x_2+x_3x_4x_5+\cdots+x_{n-\mu+1}x_{n-\mu+2}\ldots x_n$ is a coset leader of weight $w_H(f) = 2^{n-1} - \prod_{i=3}^{\mu}(2^i - 2) > 2^{n-2}$.*

Proof. According to Lemma 3 it suffices to show the result for $h(x) = x_1x_2 + x_3x_4x_5 + \cdots + x_{n-\mu+1}x_{n-\mu+2}\ldots x_n$. The direct sum vector of h as a DSM function is such that $m_0 = m_1 = 0$, $m_i = 1$ for all $i = 2, \ldots, \mu$ and $n = \sum_{i=2}^{\mu} i$. Then, according to Theorem 1, f is a coset leader and from Relation (4), $W_f(0) = \prod_{i=2}^{\mu}(2^i - 2)$. From Lemma 6, we have $\prod_{i=2}^{\mu}(2^i - 2) < 2^{n-1}$ and this ends the proof according to lemma 5. □

A function of the form $x_1x_2 + x_3x_4x_5 + \cdots + x_{n-\mu+1}x_{n-\mu+2}\ldots x_n$ is called a triangular function.

4.2 An infinite class of coset leaders that are sums of monomials and not affinely equivalent to DSM functions

Proposition 7 *Let n be an integer with $n \geq 6$. Let f be an n- variable Boolean function such that:*

$$f \sim \prod_{j=1}^{n-2} x_{i_j} + x_{i_{n-1}}x_{i_n} + x_{i_p}x_{i_q}x_{i_r}$$

where $1 \leq i_1 < i_2 < \cdots < i_n \leq n$, $p, q \in \{1, .., n-2\}$, $p \neq q$, and $r \in \{n-1, n\}$.
Then f is a coset leader of $RM(1, n)$ of weight $2^{n-2} + 2$ which is not affinely equivalent to a DSM function.

Proof. Let us calculate $WS(g)$ for $g = \prod_{j=1}^{n-2} x_{i_j} + x_{i_{n-1}} x_{i_n} + x_{i_p} x_{i_q} x_{i_r}$. We assume without loss of the generality that $p = 1, q = 2$ and $r = n-1$. For all $a = (a_1, a_2, \ldots, a_n) \in \mathbb{F}_2^n$, we denote by $a^{i_1 \cdots i_t}$ the vector of length $n-t$ obtained from a by erasing the coordinates $a_{i_1}, \ldots a_{i_t}$. We have, distinguishing the cases $x_{i_1} = 0$ and $x_{i_1} = 1$:

$$W_g(a) = \sum_{x \in \mathbb{F}_2^n} (-1)^{\prod_{j=1}^{n-2} x_{i_j} + x_{i_{n-1}} x_{i_n} + x_{i_1} x_{i_2} x_{i_{n-1}} + a \cdot x} =$$

$$\sum_{x \in \mathbb{F}_2^n; x_{i_1} = 0} (-1)^{x_{i_{n-1}} x_{i_n} + a^{i_1} \cdot x^{i_1}} + (-1)^{a_{i_1}} \sum_{x \in \mathbb{F}_2^n; x_{i_1} = 1} (-1)^{\prod_{j=2}^{n-2} x_{i_j} + x_{i_{n-1}} x_{i_n} + x_{i_2} x_{i_{n-1}} + a^{i_1} \cdot x^{i_1}} =$$

$$\sum_{x \in \mathbb{F}_2^n; x_{i_1} = 0} (-1)^{(x_{i_{n-1}} + a_{i_n})(x_{i_n} + a_{i_{n-1}}) + a_{i_{n-1}} a_{i_n} + a^{i_1 i_{n-1} i_n} \cdot x^{i_1 i_{n-1} i_n}} +$$

$$(-1)^{a_{i_1}} \Big(\sum_{x \in \mathbb{F}_2^n; x_{i_1} = 1, x_{i_2} = 0} (-1)^{x_{i_{n-1}} x_{i_n} + a^{i_1 i_2} \cdot x^{i_1 i_2}} + (-1)^{a_{i_2}} \sum_{x \in \mathbb{F}_2^n; x_{i_1} = 1, x_{i_2} = 1} (-1)^{\prod_{j=2}^{n-2} x_{i_j} + x_{i_{n-1}} x_{i_n} + x_{i_{n-1}} + a^{i_1 i_2} \cdot x^{i_1 i_2}} \Big) =$$

$$2^{n-2} (-1)^{a_{i_{n-1}} a_{i_n}} \delta_0(a^{i_1 i_{n-1} i_n}) + (-1)^{a_{i_1}} \sum_{x \in \mathbb{F}_2^n; x_{i_1} = 1, x_{i_2} = 0} (-1)^{(x_{i_{n-1}} + a_{i_n})(x_{i_n} + a_{i_{n-1}}) + a_{i_{n-1}} a_{i_n} + a^{i_1 i_2 i_{n-1} i_n} \cdot x^{i_1 i_2 i_{n-1} i_n}} +$$

$$(-1)^{a_{i_1} + a_{i_2}} A \times B$$

where

$$A = \sum_{x^{i_1 i_2 i_{n-1} i_n} \in \mathbb{F}_2^{n-4}} (-1)^{\prod_{j=3}^{n-2} x_{i_j} + a^{i_1 i_2 i_{n-1} i_n} \cdot x^{i_1 i_2 i_{n-1} i_n}}$$

and

$$B = \sum_{(x_{n-1}, x_n) \in \mathbb{F}_2^2} (-1)^{x_{i_{n-1}} (x_{i_n} + 1) + a^{i_1 i_2 \cdots i_{n-2}} \cdot (x_{n-1}, x_n)}.$$

We have then, $A = 2^{n-4} \delta_0(a^{i_1 i_2 i_{n-1} i_n}) - 2(-1)^{w_H(a^{i_1 i_2 i_{n-1} i_n})}$ and

$$B = \sum_{(x_{n-1}, x_n) \in \mathbb{F}_2^2} (-1)^{(x_{i_{n-1}} + a_{i_n})(x_{i_n} + a_{i_{n-1}} + 1) + a_{i_n} + a_{i_{n-1}} a_{i_n}} = 2(-1)^{a_{i_n} + a_{i_{n-1}} a_{i_n}}.$$

Hence,

$$\left.\begin{array}{l} W_g(a) = 2^{n-2} (-1)^{a_{i_{n-1}} a_{i_n}} \delta_0(a^{i_1 i_{n-1} i_n}) + 2^{n-3} (-1)^{a_{i_1} + a_{i_{n-1}} a_{i_n}} \delta_0(a^{i_1 i_2 i_{n-1} i_n}) + \\ (-1)^{a_{i_1} + a_{i_2} + a_{i_n} + a_{i_{n-1}} a_{i_n}} (2^{n-3} \delta_0(a^{i_1 i_2 i_{n-1} i_n}) - 4(-1)^{w_H(a^{i_1 i_2 i_{n-1} i_n})}) \end{array}\right\}, \tag{12}$$

implying that $WS(g) = \{\pm 3 \cdot 2^{n-3} \pm 4; \pm 2^{n-3} \pm 4; \pm(2^{n-1} - 4); \pm 2^{n-2} \pm 4; \pm 4\}$. We can easily check that for $n \geq 6$, $\max EWS(g) = 2^{n-1} - 4$ and by taking $a = 0$ in Relation (12), we have $W_g(0) = 2^{n-1} - 4$ meaning by Lemma 1 that g is a coset leader.

Function g can not be affinely equivalent to a DSM function h of direct sum vector $m_h = [m_0 = 0, m_1 = 0, m_2, \ldots, m_k]$. Indeed, suppose that it is, then if $n > \sum_{i=2}^{k} i m_i$, from Remark 4, $WS(h)$ contains 0 which is not the case for $WS(g)$, a contradiction, and if $n = \sum_{i=2}^{k} i m_i$, then h having necessarily degree $n-2$, it has 2 monomials and from Remark 4 again, $WS(h)$ conains 4 values which is not the case for $WS(g)$ (note that two affinely equivalent functions have the same Walsh spectrum). □

4.3 Maiorana-McFarland functions

Now we characterize coset leaders by the Walsh transform in the whole class of Maiorana-McFarland functions. Let us recall their definition:

Definition 8 *Let n and r be any positive integers such that $r \leq n$. We denote by MM_r the class of n-variable Boolean functions of the form:*

$$f(x, y) = x \cdot \phi(y) + g(y);\ x \in \mathbb{F}_2^r, y \in \mathbb{F}_2^{n-r}, \tag{13}$$

where ϕ is a function from $\mathbb{F}_2^{n-r}$ to $\mathbb{F}_2^r$ and g is an $(n-r)$-variable Boolean function. We call Maiorana-McFarland's functions (in brief, MM functions) the functions of such general form.

Note that MM_1 equals the whole space of n-variable functions (hence, speaking in this case of Maiorana-McFarland's functions is more a viewpoint on the functions than a specific definition) and $MM_r \subset MM_{r-1}$, for every $r \geq 2$.
Note that since functions ϕ and g defined in any function $f \in MM_r$ are in $n - r$ variables, then any function in MM_r is of degree at most $n - r + 1$. The original class of Maiorana-McFarland's bent functions, introduced by Maiorana, McFaland and Dillon, is the subclass of $MM_{\frac{n}{2}}$ in which ϕ is a permutation.

Clearly, each class MM_r is a union of cosets of the first order Reed-Muller code.

From [4, Proposition 53], for all function f given by Relation (13), we have:

$$W_f(u, v) = 2^r \sum_{y \in (\phi)^{-1}(u)} (-1)^{g(y)+v\cdot y};\ u \in \mathbb{F}_2^r, v \in \mathbb{F}_2^{n-r}. \tag{14}$$

This means that

$$w_H(f) = 2^{n-1} - 2^{r-1} \sum_{y \in (\phi)^{-1}(0)} (-1)^{g(y)} \tag{15}$$

and that,

$$nl(f) = 2^{n-1} - 2^{r-1} \max_{u \in \mathbb{F}_2^r,\, v \in \mathbb{F}_2^{n-r}} \left| \sum_{y \in (\phi)^{-1}(u)} (-1)^{g(y)+v\cdot y} \right|.$$

Hence, from Lemma 1, we have

Lemma 7 *Let n and r be two positive integers such that $r \leq n$. For all $f \in MM_r$, f is a coset leader of $RM(1, n)$ if and only if*

$$\max_{u \in \mathbb{F}_2^r, v \in \mathbb{F}_2^{n-r}} \left| \sum_{y \in (\phi)^{-1}(u)} (-1)^{g(y)+v\cdot y} \right| = \sum_{y \in (\phi)^{-1}(0)} (-1)^{g(y)}.$$

Remark 6 *Given $f = x \cdot \phi(y) + g(y)$ in MM_r and by setting $M_{u,v,\epsilon} = \{y \in (\phi)^{-1}(u)/\ g(y) + v \cdot y = \epsilon\}$ where $(u, v) \in \mathbb{F}_2^r \times \mathbb{F}_2^{n-r}$ and $\epsilon \in \mathbb{F}_2$, we have from Lemma 7 that every function $g = f + u \cdot x + v \cdot y + \epsilon$ is a coset leader if and only if $card(M_{u,v,\epsilon}) - card(M_{u,v,1+\epsilon})$ is positive and is maximal over $\mathbb{F}_2^r \times \mathbb{F}_2^{n-r}$. Indeed, it suffices to observe that $\sum_{y \in (\phi)^{-1}(u)} (-1)^{g(y)+v \cdot y} = card(M_{u,v,0}) - card(M_{u,v,1})$.*

A simple case where a Maiorana-McFarland function is a coset leader is the following:

Corollary 6 *Let n and r be two positive integers such that $r \leq n$. For all $f(x, y) = x \cdot \phi(y) + g(y) \in MM_r$, if g vanishes on $\phi^{-1}(0)$ and if $Card((\phi)^{-1}(0)) \geq card((\phi)^{-1}(u))$ for all $u \in \mathbb{F}_2^r$, then f is a coset leader of $RM(1, n)$.*

Remark 7 *Let n and r be two positive integers such that $r \leq n$. Assume that a function $f = x \cdot \phi(y) + g(y) \in MM_r$ is a coset leader of $RM(1, n)$. Then, according to Relation (15), f has Hamming weight larger than 2^{n-2} if and only if $\sum_{y \in (\phi)^{-1}(0)} (-1)^{g(y)} < 2^{n-r-1}$.*

Remark 8 *If ϕ is a permutation as it is the case for functions in the Maiorana-McFarland original class of bent functions, then $\phi^{-1}(u)$ has only one element for all u and any function $f(x, y) = x \cdot \phi(y) + g(y) \in MM_{\frac{n}{2}}$ is a coset leader if and only if $g(\phi^{-1}(0)) = 0$. But we knew already that, more generally, the bent functions which are coset leaders are those of Hamming weight $2^{n-1} - 2^{\frac{n}{2}-1}$.*

It is clearly out of reach to precisely determine all the coset leaders of Maiorana-McFarland functions: this is obvious for class MM_1, since we have seen that MM_1 equals to whole class of n-variable Boolean functions. Let us look at the large values of r.
Before that, assume that the function ϕ considered the function $f = x \cdot \phi(y) + g(y)$ in MM_r is constant then we have:

Lemma 8 *The coset leaders of every function $f = x \cdot \phi(y) + g(y)$ in MM_r, when ϕ is constant, are the coset leaders of g (viewed as functions in n variables).*

Proof Assume ϕ is constant and takes value $u_0 \in \mathbb{F}_2^r$. For all $(u, v) \in \mathbb{F}_2^r \times \mathbb{F}_2^{n-r}$, if $u \neq u_0$ then $\phi^{-1}(u) = \emptyset$ and we have $\sum_{y \in (\phi)^{-1}(u)} (-1)^{g(y)+v \cdot y} = 0$, which means that the coset leaders of f are obtained when $u = u_0$ only. Then we have $\sum_{y \in (\phi)^{-1}(u_0)} (-1)^{g(y)+v \cdot y} = \sum_{y \in \mathbb{F}_2^{n-r}} (-1)^{g(y)+v \cdot y}$ and this ends the proof since g is a Boolean function in $n - r$ variables. □

Theorem 2 *Let n and r be two positive integers such that $r \leq n$.*

1. *If $r = n$, the zero function is the unique coset leader in MM_r.*
2. *If $r = n - 1$, the coset leaders in MM_r are the zero function and, for every ϕ taking two distinct values $u_0 = \phi(0)$ and $u_1 = \phi(1)$, the four functions:*

$$f(x, y) = x \cdot (\phi(y) + u_i) + v(y + i);\ x \in \mathbb{F}_2^{n-1},\ y \in \mathbb{F}_2;\ i, v \in \mathbb{F}_2.$$

3. *If $r = n - 2$, then the coset leaders $f(x, y)$; $x \in \mathbb{F}_2^{n-2}; y \in \mathbb{F}_2^2$, of MM_r are as follows, according to the size of the image set of ϕ:*

   a) *[ϕ constant]*
   $f(x, y) = 0$, $f(x, y) = y_1y_2$, $f(x, y) = y_1y_2 + y_1$, $f(x, y) = y_1y_2 + y_2$, *and* $f(x, y) = y_1y_2 + y_1 + y_2 + 1$.

   b) *[ϕ taking two distinct values u_1 and u_2]*

   – *If $card(\phi^{-1}(u_1)) \neq card(\phi^{-1}(u_2))$, then up to renaming u_1, u_2, we have $\phi^{-1}(u_1) = \{z_1, z_2, z_3\}$, $\phi^{-1}(u_2) = \{z_4\}$ and the corresponding coset leaders are the functions:*

   $$f(x, y) = x \cdot (\phi(y) + u_1) + g(y) + v \cdot y + g(z_1) + v \cdot z_1,$$

   where g and v are such that $g(y) + v \cdot y$ is constant over $\phi^{-1}(u_1)$; the functions:

   $$f(x, y) = x \cdot (\phi(y) + u_1) + g(y) + v \cdot y + \sum_{i=1}^{3}(g(z_i) + v \cdot z_i) + 1; v \in \mathbb{F}_2^2,$$

   where g is such that $g(y) + v \cdot y$ is non-constant over $\phi^{-1}(u_1)$ for every v, and the functions:

   $$f(x, y) = x \cdot (\phi(y) + u_2) + g(y) + v \cdot y + g(z_4) + v \cdot z_4; v \in \mathbb{F}_2^2.$$

   – *If $card(\phi^{-1}(u_1)) = card(\phi^{-1}(u_2)) = 2$, then by setting $\phi^{-1}(u_1) = \{z_1, z_2\}$ and $\phi^{-1}(u_2) = \{z_3, z_4\}$, the coset leaders are the functions:*

   $$f(x, y) = x \cdot (\phi(y) + u_1) + g(y) + v \cdot y + g(z_1) + v \cdot z_1,$$

   where g is such that $g(y) + v \cdot y$ is constant over $\phi^{-1}(u_1)$, and the functions:

   $$f(x, y) = x \cdot (\phi(y) + u_2) + g(y) + v \cdot y + g(z_3) + v \cdot z_3,$$

   where g is such that $g(y) + v \cdot y$ is constant over $\phi^{-1}(u_2)$.

   c) *[ϕ taking three distinct values u_1, u_2 and u_3]. Up to renaming u_1, u_2 and u_3, we have $\phi^{-1}(u_1) = \{z_1, z_2\}$, $u_2 = \phi(z_3)$ and $u_3 = \phi(z_4)$, then the corresponding coset leaders are the functions:*

   $$f(x, y) = x \cdot (\phi(y) + u_1) + g(y) + v \cdot y + g(z_1) + v \cdot z_1;\ x \in \mathbb{F}_2^{n-2},\ y \in \mathbb{F}_2^2,$$

   where g and v are such that $g(y) + v \cdot y$ is constant over $\phi^{-1}(u_1)$, and the functions:

   $$f(x, y) = x \cdot (\phi(y) + u_i) + g(y) + v \cdot y + g(z_{i+1}) + v \cdot z_{i+1};\ i = 2, 3,\ v \in \mathbb{F}_2^2,$$

   if all the restrictions of $g(y) + v \cdot y$ to $\phi^{-1}(u_1)$ are balanced.

d) [ϕ taking four distinct values $u_1 = \phi(z_1)$, $u_2 = \phi(z_2)$, $u_3 = \phi(z_3)$ and $u_4 = \phi(z_4)$] where $\mathbb{F}_2^2 = \{z_1, z_2, z_3, z_4\}$, then the coset leaders of f are the sixteen functions:

$$f(x, y) = x \cdot (\phi(y) + u_i) + g(y) + v \cdot y + g(z_i) + v \cdot z_i; i \in \{1, 2, 3, 4\}, , v \in \mathbb{F}_2^2.$$

Proof. 1) If $r = n$, then $MM_r = RM(1, n)$ and the zero function is the unique coset leader.
2) If $r = n - 1$, then for a given function $f(x, y) = x \cdot \phi(y) + g(y)$, where ϕ and g are 1-variable functions, we have:
a) If ϕ is constant, then according to Lemma 8, the coset leaders of f are the coset leaders of g (viewed as in n variables). Since g is in 1 variable, it is affine and the unique coset leader of g is then function 0. The unique coset leader corresponding to this case is then the zero function (we could have seen this directly by observing that f itself is affine).
b) If ϕ takes two distinct values $u_0 = \phi(0)$ and $u_1 = \phi(1)$, the pre-image $\phi^{-1}(u)$ is empty when $u \notin \{u_0, u_1\}$ and equals $\{i\}$ where $i \in \mathbb{F}_2$ when $u = u_i$. Since g is affine, we can without loss of generality assume that it is the zero function and consider the coset leaders of f. The sum $\sum_{y \in \phi^{-1}(u)} (-1)^{g(y)+v \cdot y}$ equals then 0 when $u \notin \{u_0, u_1\}$ and $(-1)^{vi}$ when $u = u_i$, and the result follows from Relation (14) (or Lemma 7) and Lemma 2 with $\epsilon = vi$.
3) If $r = n - 2$, the functions ϕ and g in $f(x, y) = x \cdot \phi(y) + g(y) \in MM_r$ are 2-variable functions, and we have:
a) If ϕ is constant, then according to Lemma 8, the coset leaders of f are the coset leaders of g (viewed as in n variables). Since g is in 2 variables y_1 and y_2, it is either affine, and the unique coset leader of g is then function 0, or it has degree 2, in which case the coset leaders of g when g varies among the functions of degree 2 in two variables are the functions of Hamming weight 1, that is, the functions y_1y_2, $y_1(y_2 + 1) = y_1y_2 + y_1$, $(y_1 + 1)y_2 = y_1y_2 + y_2$ and $(y_1 + 1)(y_2 + 1) = y_1y_2 + y_1 + y_2 + 1$.
b) If ϕ takes two distinct values u_1 and u_2, noting that $card(\phi^{-1}(u)) \leq 3$ for all u, we have:
- If $\phi^{-1}(u_1) = \{z_1, z_2, z_3\}$ and $\phi^{-1}(u_2) = \{z_4\}$, then if $u \notin \{u_1, u_2\}$, the sum $\sum_{y \in \phi^{-1}(u)} (-1)^{g(y)+v \cdot y}$ equals 0; if $u = u_1$ and $g(y) + v \cdot y$ is constant over $\phi^{-1}(u_1)$, it equals $3(-1)^{g(z_1)+v \cdot z_1}$; if $u = u_1$ and $g(y) + v \cdot y$ is not constant over $\phi^{-1}(u_1)$, it equals $(-1)^{1+\sum_{i=1}^{3}(g(z_i)+v \cdot z_i)}$; and if $u = u_2$, it equals $(-1)^{g(z_4)+v \cdot z_4}$ and the result follows from Lemma 2.
-If $card(\phi^{-1}(u_1)) = card(\phi^{-1}(u_2)) = 2$, then by setting $\phi^{-1}(u_1) = \{z_1, z_2\}$ and $\phi^{-1}(u_2) = \{z_3, z_4\}$, the sum $\sum_{y \in \phi^{-1}(u)} (-1)^{g(y)+v \cdot y}$ equals 0 when $u \notin \{u_1, u_2\}$ or when $u = u_1$ and $g(y)+v \cdot y$ is balanced over $\phi^{-1}(u_1)$ or when $u = u_2$ and $g(y) + v \cdot y$ is balanced over $\phi^{-1}(u_2)$. This sum equals $2(-1)^{g(z_1)+v \cdot z_1}$ (resp. $2(-1)^{g(z_3)+v \cdot z_3}$) when $u = u_1$ and $g(y) + v \cdot y$ is constant over $\phi^{-1}(u_1)$ (resp. when $u = u_2$ and $g(y) + v \cdot y$ is constant over $\phi^{-1}(u_2)$), and the result follows again from Lemma 2.
c) The pre-image $\phi^{-1}(u)$ is empty when $u \notin \{u_1, u_2, u_3\}$. Noting that $card(\phi^{-1}(u)) \leq 2$, we can take $\phi^{-1}(u_1) = \{z_1, z_2\}$, $\phi(z_3) = u_2$ and $\phi(z_4) = u_3$, then the sum $\sum_{y \in \phi^{-1}(u)} (-1)^{g(y)+v \cdot y}$ equals 0 when $u \notin \{u_1, u_2, u_3\}$ or when $u = u_1$ and $g(y) + v \cdot y$ is balanced over $\phi^{-1}(u_1)$, it equals $2(-1)^{g(z_1)+v \cdot z_1}$ when $u = u_1$ and $g(y) + v \cdot y$ is constant over $\phi^{-1}(u_1)$, and the result follows from Lemma 2.

d) The pre-image $\phi^{-1}(u)$ is empty when $u \notin \{u_1, u_2, u_3, u_4\}$ and equals $z_i \in \mathbb{F}_2^2$ when $u = u_i$. The sum $\sum_{y \in \phi^{-1}(u)} (-1)^{g(y)+v \cdot y}$ equals then 0 when $u \notin \{u_1, u_2, u_3, u_4\}$ and $(-1)^{g(z_i)+v \cdot z_i}$ when $u = u_i$, and the result follows from Lemma 2. □

Remark 9 *Denoting the Maiorana-McFarland original class of bent functions by $MM'_{\frac{n}{2}}$, every coset leader in $MM'_{\frac{n}{2}}$ with $n \geq 6$ and $deg(g) \geq 3$ is not affinely equivalent to a DSM function. Indeed, its Walsh spectrum, which is $\{\pm 2^{\frac{n}{2}}\}$, is different from the Walsh spectrum of any DSM function of degree at least 3 given in Remark 4, since this latter Walsh spectrum contains a value divisible by $2^{deg(g)} - 2$, which is not the case for $\pm 2^{\frac{n}{2}}$.*

Conclusion

In this paper, we gave properties of the coset leaders of the first-order Reed-Muller code and started a study of their structure through the study of operations that are internal to their class. We characterized those coset leaders that belong to the well known classes of direct sums of monomial functions and Maiorana-McFarland functions. We also gave infinite classes of coset leaders having Hamming weight greater than 2^{n-2}. Many questions remain open after our work; the global structure of coset leaders is still to be understood, and the coset leaders of many classes whose Walsh transform is known (such as the class of Niho functions, majority functions and Carlet-Feng functions) remain to be determined.

References

[1] A. Barg, Complexity issue in coding theory. *In HandBook of coding theory*, Vol I, II, pages 649-724. North Holland, Amsterdam, 1998.

[2] M. Borges-Quintana, M.A. Borges-Trenard, L. Marquez-Corbella and E. Matinez-Moro. Computing Coset Leaders and Leader Codewords of Binary Codes. *Journal of Algebra and Its Applications*. 2014.

[3] C. Carlet. On bent and highly nonlinear balanced/resilient functions and their algebraic immunities. *Proceedings of AAECC-16 Conference, Lecture Notes in Computer Science* 3857, pp. 1-28, 2006.

[4] C. Carlet. Boolean Functions for Cryptography and Error Correcting Codes. *Cambridge University Press*, 562 pages, 2021.

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

[6] C. Carlet and P. Méaux. A complete study of two classes of Boolean functions: direct sums of monomials and threshold functions. To appear in *IEEE Transactions on Information Theory*.

[7] C. Carlet and S. Mesnager. Four decades of research on bent functions. Special Jubilee Issue of *Designs, Codes and Cryptography*, Vol. 78, pp. 5-50, 2016.

[8] T.W. Cusick ,P. Stanica. Fast evaluation, weights and nonlinearity of rotation-symmetric functions. *Discrete Mathematics*, Vol 258 pp. 289-301, 2002.

[9] D.K. Dalai, S. Maitra and S. Sarkar. Basic Theory in Construction of Boolean Functions with Maximum possible Annihilator Immunity. *Designs, Codes and Cryptography* 40(1), pp. 41-58, 2006.

[10] T. Helleseth, and T. Kløve. The Newton Radius of Codes *IEEE Transactions on Information Theory*, Vol. 43, No. 6, 1997.

[11] P. Langevin and P . Solé. Kernels and defaults. Proceedings of the conference *Finite Fields and Applications* Fq4, *Contemporary Mathematics* 225, pp. 77-85, 1999.

[12] F. J. MacWilliams and N. J. Sloane. *The theory of error-correcting codes*, North Holland. 1977.

[13] P. Méaux, C. Carlet, A. Journault. F-X. Standaert and C. Carlet. Towards stream ciphers for efficient FHE with low-noise ciphertexts. *In Marc Fischlin and Jean-St'ebastien Coron, editors, EUROCRYPT 2016*, Part I, volume 9665 of LNCS, pages 311Ű343. Springer, Heidelberg, May 2016.

[14] S. Mesnager. Bent Functions: Fundamentals and Results *Springer Verlag, 2016*, Version available at `http://www.math.univ-paris13.fr/~mesnager/Publications/Contents-Book-bent-Mesnager---copie---copie.pdf`

[15] O. S. Rothaus, On “bent" functions, *J. Comb. Theory*, **20** (1976), 300–305.

[16] X. Zhang, H.G. Rongquan Feng and Y. Li. Proof of a conjecture about rotation symmetric functions. *Discrete Mathematics* 311(2011), pp. 1281-1289.