Maiorana–McFarland Functions with High Second–Order Nonlinearity

Nicholas Kolokotronis[1] and Konstantinos Limniotis[2,3]

[1] Department of Computer Science and Technology
University of Peloponnese
End of Karaiskaki Street, 22100 Tripolis, Greece
Email: nkolok@uop.gr

[2] Hellenic Data Protection Authority
Kifisias 1–3, 11523 Athens, Greece
Email: klimniotis@dpa.gr

[3] Department of Informatics and Telecommunications
National and Kapodistrian University of Athens
University Campus, 15784 Athens, Greece
Email: klimn@di.uoa.gr

Abstract. The second–order nonlinearity, and the best quadratic approximations, of Boolean functions are studied in this paper. We prove that cubic functions within the Maiorana–McFarland class achieve very high second order nonlinearity, which is close to an upper bound that was recently proved by Carlet *et al.*, and much higher than the second order nonlinearity obtained by other known constructions. The structure of the cubic Boolean functions considered allows the efficient computation of (a subset of) their best quadratic approximations.

Keywords. Boolean functions; cryptography; Maiorana–McFarland class; second–order nonlinearity.

1 Introduction

Boolean functions have a prominent role in cryptography. Their most important applications are in the analysis and design of building blocks used in symmetric cryptosystems: s–boxes [26] and filter or combining functions. A fundamental property that all Boolean functions are required to have is high nonlinearity, which is the minimum distance from all the affine functions; it determines the extent to which linear cryptanalytic attacks can be prevented. For an even number n of variables, the maximum nonlinearity is $2^{n-1} - 2^{n/2-1}$ and is only attained by bent functions; on the contrary, the maximum nonlinearity attained for an odd number of variables still remains an open problem. Recently the appearance of new attacks based on higher–order correlation [11] and low–order approximation [19, 21] (threatening the security of cryptosystems being immune to linear cryptanalytic attacks) necessitates the absence of good low–degree (not necessarily affine) approximations of the cryptographic Boolean functions employed. As a result, the rth–order nonlinearity, $r \geq 1$,

which is the minimum distance from all the functions of degree at most r, received great attention in the past few years [8, 9].

From a cryptographic viewpoint, not much is known about the structure of functions with maximum rth–order nonlinearity, as these values are unknown in general. Even the second–order nonlinearity is unknown for all Boolean functions, with the exception of a small number of variables, or some special cases [9]. Moreover, proving bounds on the rth–order nonlinearity is also a hard task, even for $r = 2$. Many results in this area are stated in terms of properties like algebraic immunity, e.g. the lower bounds in [6, 7, 24]; other lower (resp. upper) bounds can be found in [9, 10, 14–16, 18, 22, 30–32] (resp. [8, 10]). Apart from determining the rth–order nonlinearity, the computation of the best rth–order approximations of a given Boolean function is also of high importance. This problem was first studied in [25], where an algorithm to find good (not necessarily the best) rth–order approximations is given. Recently, a way to efficiently compute all the best second–order approximations of cubic Boolean functions, with arbitrary number of variables, is given in [20]; the functions f considered therein were separable, that is $f = f_1 + \cdots + f_m$ where $f_1, \ldots, f_m$ are defined on disjoint sets of variables.

In this paper, we extend the results of [20] by considering cubic Boolean functions in the well–known Maiorana–McFarland class. We prove that, under certain assumptions, the second–order nonlinearity of cubic functions in this class is equal to their nonlinearity (hence, their best affine approximations are among their best quadratic approximations); the analysis utilizes properties of quadratic vectorial perfect–nonlinear Boolean functions. Compared to the second–order nonlinearity obtained by other constructions, or general lower bounds on its value (not necessarily associated with an explicit construction), our result is superior and very close to an upper bound recently proved by Carlet *et al.* [8]. In addition, we determine (a subset of) their best quadratic approximations and establish a link with the directional derivatives of the functions considered.

The rest of the paper is organized as follows. Section 2 provides the background and settles the notation. The second–order nonlinearity of cubic functions in the Maiorana–McFarland class is studied in Section 3, where we prove that it is equal to their nonlinearity. Finally, concluding remarks are given in Section 4.

2 Preliminaries

Let $\mathbb{B}_n$ be the set of all Boolean functions on n variables, i.e. mappings $f : \mathbb{F}_2^n \to \mathbb{F}_2$ with $\mathbb{F}_2 = \{0, 1\}$, and let $x = (x_1, \ldots, x_n)$. It is well–known that any $f \in \mathbb{B}_n$ admits a unique polynomial representation in the quotient ring $\mathbb{F}_2[x_1, \ldots, x_n]/\langle x_1^2 + x_1, \ldots, x_n^2 + x_n\rangle$

$$f(x) = \sum_{u \in \mathbb{F}_2^n} c_u x^u, \quad c_u \in \mathbb{F}_2 \tag{1}$$

called the *algebraic normal form* (ANF) of f [23], where $u = (u_1, \ldots, u_n)$ and $x^u = \prod_i x_i^{u_i}$. The algebraic degree of f, denoted by $\deg f$, is the maximum degree of the monomials in (1) whose coefficients are nonzero; the linear functions $ax^t = \sum_i a_i x_i$ are denoted by φ_a. For $\deg f \leq r$, the truth table of f is a codeword of the Reed–Muller code $\mathsf{RM}(r, n)$ [23],

and we write $f \in \mathsf{RM}(r,n)$ for simplicity. The *restriction* of f on the subspace $E \subseteq \mathbb{F}_2^n$ is denoted by $f|_E$, and $D_a f$ is its *derivative* along the direction of $a \in \mathbb{F}_2^n$, which is defined as $D_a f(x) = f(x+a) + f(x)$.

The set $\mathsf{supp}\, f$ is the subset of $\mathbb{F}_2^n$ where f takes nonzero value and $\mathsf{wt}(f) = \big|\mathsf{supp}\, f\big|$ is the Hamming weight of f; the distance between $f, g \in \mathbb{B}_n$ is given by $\mathsf{d}(f,g) = \mathsf{wt}(f+g)$. The *Walsh transform* $\chi_f(a)$ at $a \in \mathbb{F}_2^n$ of f is the integer–valued function

$$\chi_f(a) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x)+ax^t} = 2^n - 2\,\mathsf{d}(f, \varphi_a)\,. \tag{2}$$

The function $f \in \mathbb{B}_n$ is said to be *balanced* if $\mathsf{wt}(f) = 2^{n-1}$; equivalently, when $\chi_f(0) = 0$ due to (2). The *rth–order nonlinearity* of f is defined as its minimum distance from all the Boolean functions of degree at most r

$$\mathsf{nl}_r(f) = \mathsf{d}(f, \mathsf{RM}(r,n)) \tag{3}$$

where $\mathsf{nl}(f) = \mathsf{nl}_1(f)$. Functions in $\mathsf{RM}(r,n)$ whose distance from f equals $\mathsf{nl}_r(f)$ are called *best rth–order approximations* of f and comprise the set $\mathsf{A}_r(f)$ —likewise, $\mathsf{A}(f) = \mathsf{A}_1(f)$. By its definition, $\mathsf{nl}_r(f)$ only depends on the monomials in the ANF of f whose degree is greater than r; $\mathsf{nl}_r(f)$ is also invariant under any affine transformation of f. The bounds

$$\max_{f \in \mathbb{B}_n} \mathsf{nl}_2(f) \le 2^{n-1} - 2^{n/2-1}\sqrt{15} + \mathcal{O}(1) \tag{4a}$$

$$\max_{f \in \mathbb{B}_n} \mathsf{nl}_2(f) \ge 2^{n-1} - 2^{n/2-1} n\sqrt{\ln 2} + \mathcal{O}(1) \tag{4b}$$

are proved in [8, 10] respectively, and are the best known bounds for the maximum value of $\mathsf{nl}_2(f)$. The exact value of $\mathsf{max}_{f \in \mathbb{B}_n} \mathsf{nl}_r(f)$ remains unknown for $r > 1$, apart from some special cases. When $r = 1$ and n is even, only *bent functions* f achieve the highest possible nonlinearity $\mathsf{nl}(f) = 2^{n-1} - 2^{n/2-1}$; such functions satisfy $|\chi_f(a)| = 2^{n/2}\ \forall a \in \mathbb{F}_2^n$ and the unique function $f^\perp$ defined by $(-1)^{f^\perp(a)} = 2^{-n/2}\chi_f(a)$ is also bent [13] —referred to as the *dual* of f. The derivatives $D_a f$, $a \neq 0$, of bent functions are all balanced.

2.1 Vectorial Boolean Functions

A multi–output mapping $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$, with $m > 1$, is called *vectorial* Boolean function and is represented as $F = (f_1, \ldots, f_m)$, where $f_1, \ldots, f_m \in \mathbb{B}_n$ are its coordinate functions. Most of the notions given so far are extend in a natural way. The algebraic degree of F is the maximum of its coordinate functions' degrees. Linear functions $(b_1 x^t, \ldots, b_m x^t) = xB$ are denoted by Φ_B, and B is the $n \times m$ matrix whose columns are $b_1, \ldots, b_m$; $\mathsf{im}\ \Phi_B$ and $\mathsf{ker}\ \Phi_B$ are the image and kernel of Φ_B respectively ($\mathsf{coim}\ \Phi_B = \mathbb{F}_2^n / \mathsf{ker}\ \Phi_B$ is also used next). The pre–image of $c \in \mathbb{F}_2^m$ is comprised by those $x \in \mathbb{F}_2^n$ for which $F(x) = c$ and is denoted by $F^{-1}(c)$. Its cardinality is given by [5, Proposition 1]

$$\big|F^{-1}(c)\big| = 2^{-m} \sum_{a \in \mathbb{F}_2^m} (-1)^{ca^t} \chi_{\varphi_a \circ F}(0)\,. \tag{5}$$

The rth–order nonlinearity of F equals the minimum among $\mathsf{nl}_r(\varphi_a \circ F)$, $a \neq 0$, where $\circ$ stands for the composition of functions [5]. If $\varphi_a \circ F$, $a \neq 0$, are all bent functions, then $\mathsf{nl}(F) = 2^{n-1} - 2^{n/2-1}$ and F is called *perfect nonlinear* (PN); they are known to exist if n is even and $m \leq n/2$. Since $|\chi_{\varphi_a \circ F}(0)|$ equals 2^n if $a = 0$ and $2^{n/2}$ otherwise, when F is a PN function, (5) becomes

$$\begin{aligned} |F^{-1}(c)| &= 2^{-m}\left(2^n + 2^{n/2}\sum_{a\neq 0}(-1)^{F^\star(a)+ca^t}\right) \\ &= 2^{n/2-m}(2^{n/2} + \chi_{F^\star}(c) - 1) \end{aligned} \tag{6}$$

where the Boolean function $F^\star : \mathbb{F}_2^m \to \mathbb{F}_2$ is defined as $F^\star(a) = (\varphi_a \circ F)^\perp(0)$ $\forall a \neq 0$ and $F^\star(0) = 0$ [5]. For all $b \neq 0$, the derivatives $D_bF(x) = F(x+b) + F(x)$ of a PN function are balanced, i.e. the pre–images of all the elements of $\mathbb{F}_2^m$ have cardinality 2^{n-m}. Due to (6), vectorial functions with $F^\star = 0$ (or linear $F^\star$ in general) have the highest disparity

$$|F^{-1}(0)| - |F^{-1}(c)| = 2^{n/2}, \quad \forall c \neq 0$$

among the pre–images of the elements of $\mathbb{F}_2^m$ (*see* also [26]). In the sequel, we also use the notation Fa^t instead of $\varphi_a \circ F$ whenever convenient.

2.2 Quadratic Boolean Functions

Let $f \in \mathbb{B}_n$ be a quadratic function, whose quadratic terms are xQx^t, where Q is a strictly upper–triangular $n \times n$ matrix. If $\mathcal{Q} = Q + Q^t$, then we will find it convenient to define $\mathsf{rank}\, f = \mathsf{rank}\, \mathcal{Q}$, as many properties of f depend on $\mathsf{rank}\, \mathcal{Q}$. In particular, it is known that $\mathsf{rank}\, f = 2h$ for some $1 \leq h \leq \lfloor n/2 \rfloor$ [23]; the nonlinearity of $f \in \mathbb{B}_n$ is given by

$$\mathsf{nl}(f) = 2^{n-1} - 2^{n-1-h}.$$

When n is even, f is bent if and only if $\mathsf{rank}\, f = n$ [12, 27], which in turn is equivalent to $\ker \Phi_{\mathcal{Q}} = \{0\}$ and f is called non–degenerate [4]. The number of best affine approximations of f is equal to 2^{2h}; they are determined as shown next.

Proposition 1 ([20]). *If $\varphi_\lambda + \delta$ is a best affine approximation of the quadratic function $f \in \mathbb{B}_n$, then all its best affine approximations are determined by*

$$\mathsf{A}(f) = \{D_a(f + \varphi_\lambda) : a \in \mathbb{F}_2^n\} + \varphi_\lambda + \delta. \tag{7}$$

Since $|\mathsf{A}(f)| = 2^{2h}$, by computing the best affine approximations according to (7), results in obtaining each element of $\mathsf{A}(f)$ multiple times; this can be avoided if we restrict $a \in \mathbb{F}_2^n$ to take values from $\mathsf{coim}\, \Phi_{\mathcal{Q}}$ [20]. From (2), (3), we see that the function $\varphi_l + \epsilon$ is a best affine approximation of f if and only if $\chi_f(l) = (-1)^\epsilon 2^{n-h}$; however, Proposition 1 asserts that this holds if and only if $l = \lambda + a\mathcal{Q}$ and $\epsilon = \delta + D_a(f + \varphi_\lambda)(0)$ for some $a \in \mathsf{coim}\, \Phi_{\mathcal{Q}}$. Consequently, the nonzero values of the Walsh transform of f are given by

$$\chi_f(\lambda + a\mathcal{Q}) = (-1)^{\delta + D_a(f+\varphi_\lambda)(0)}\, 2^{n-h}, \qquad \forall a \in \mathsf{coim}\, \Phi_{\mathcal{Q}}. \tag{8}$$

If f is bent, then either $\delta = 0$ or $\delta = 1$ is a best affine approximation of f (that is, we can choose $\lambda = 0$). Recalling the definition of the dual function, the previous analysis implies that $f^\perp(a\mathcal{Q}) = f(a) + f(0) + \delta$ or equivalently

$$f^\perp(x) = f(x\mathcal{Q}^{-1}) + f(0) + \delta, \qquad \forall x \in \mathbb{F}_2^n \tag{9}$$

due to the invertibility of $\mathcal{Q}$. It is easily established from (9) that $\deg f^\perp = 2$, $\delta = f^\perp(0)$, and $\mathcal{Q}^{-1}$ is the symplectic matrix associated with $f^\perp$. On the other hand, if f is not bent then the dual function is not necessarily unique; however, the restriction $(f + \varphi_\lambda)|_{\mathsf{coim}\, \Phi_\mathcal{Q}}$ is a bent function on $2h$ variables and admits a unique dual. Working as above, we have

$$f^\perp(\lambda + x\mathcal{Q}) = (f + \varphi_\lambda)(x) + f(0) + \delta, \qquad \forall x \in \mathsf{coim}\, \Phi_\mathcal{Q}\,. \tag{10}$$

Likewise, we see that it holds $\delta = f^\perp(\lambda)$ and $\mathsf{rank}\, f^\perp = \mathsf{rank}\, f$. Since $y = x\mathcal{Q} \in \mathsf{im}\, \Phi_\mathcal{Q}$, we have proved the following result.

Proposition 2. *With the above notation, the dual function of f is any quadratic function $f^\perp \in \mathbb{B}_n$ such that $f^\perp|_{\lambda + \mathsf{im}\, \Phi_\mathcal{Q}} + f^\perp(\lambda) = (f + \varphi_\lambda)|_{\mathsf{coim}\, \Phi_\mathcal{Q}} + f(0)$.*

3 Maiorana–McFarland Class of Cubic Functions

In this section, we find the second–order nonlinearity and best quadratic approximations of cubic functions that belong to the general Maiorana–McFarland class. For a quadratic $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ and cubic $g : \mathbb{F}_2^n \to \mathbb{F}_2$, the functions in this class have the form

$$f(x, y) = F(x)y^t + g(x), \qquad (x, y) \in \mathbb{F}_2^n \times \mathbb{F}_2^m\,. \tag{11}$$

This construction is widely used to get Boolean functions simultaneously satisfying many cryptographic properties [2, 3, 29]. Next, we confine ourselves to quadratic PN functions F and $g \in \mathsf{RM}(2, n)$ —equivalently, we may assume that $g = 0$, as we are only interested in studying the second–order nonlinearity of f. The weight of functions in the Maiorana–McFarland class (for any g) is determined as shown below.

Proposition 3. *Let the Maiorana–McFarland function f be given by (11); then, we have*

$$\mathsf{wt}(f) = 2^{n+m-1} - 2^{m-1} \sum_{x \in F^{-1}(0)} (-1)^{g(x)}\,.$$

Proof. We may compute the weight of f by summing–up the weights of all its restrictions $f|_{x=u}$ $\forall u \in \mathbb{F}_2^n$, which are the affine functions $F(u)y^t + g(u)$. If $F(u) = 0$, their weight is either zero or 2^m (i.e. equal to $2^m g(u)$); otherwise, they have weight 2^{m-1}. Hence

$$\begin{aligned}\sum_{u \in \mathbb{F}_2^n} \mathsf{wt}(f|_{x=u}) &= 2^m \sum_{u \in F^{-1}(0)} g(u) + 2^{m-1}\Big(2^n - |F^{-1}(0)|\Big) \\ &= 2^{n+m-1} - 2^{m-1}\Big(|F^{-1}(0)| - 2\,\mathsf{wt}(g|_{F^{-1}(0)})\Big)\end{aligned}$$

where $\mathsf{wt}(g|_{F^{-1}(0)}) = \sum_{u \in F^{-1}(0)} g(u)$, immediately leading to the desired result. ∎

Remark 1. To determine the second–order nonlinearity of Maiorana–McFarland functions, we need to write quadratic functions on $n+m$ variables ($x \in \mathbb{F}_2^n$ and $y \in \mathbb{F}_2^m$) as follows

$$g(x,y) = (x\ y)\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}\begin{pmatrix} x^t \\ y^t \end{pmatrix} + (a\ b)\begin{pmatrix} x^t \\ y^t \end{pmatrix} + \epsilon = p(x) + L(x)y^t + q(y) \tag{12}$$

where p, q are the quadratic functions $p(x) = xAx^t + ax^t + \epsilon$ and $q(y) = yBy^t + by^t$ on n and m variables respectively, whereas $L : \mathbb{F}_2^n \to \mathbb{F}_2^m$ is the linear mapping Φ_C. Note that not all A, B, C can simultaneously be zero, as otherwise g would not be quadratic.

Let φ_λ, $\lambda \in \mathbb{F}_2^m$, be a best linear approximation of q (as given above); then, we can set $q(y) = yBy^t + (b+\lambda)y^t$ and $L(x) = \Phi_C(x) + \lambda$ instead. This simplifies our analysis, as q now has minimum weight $2^{m-1} - 2^{m-1-\mathsf{rank}\, q/2}$ and does not affect g. □

Theorem 1. *Let n be even, $m \le n/2$, and $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ be an arbitrary quadratic PN function. The second–order nonlinearity of the cubic function $f(x,y) = F(x)y^t$ satisfies*

$$\mathsf{nl}_2(f) \ge 2^{n+m-1} - 2^{n/2-1}\left(2^{n/2} + 2^m - 1\right). \tag{13}$$

Proof. Let $g \in \mathbb{B}_{n+m}$ be a quadratic function given by (12) and let q have minimum weight (*see* Remark 1), where $\mathsf{rank}\, q = 2h$. If we set $G = F + L$, then G is also a quadratic PN function. From (2) we immediately get that $\mathsf{d}(f,g) = 2^{n+m-1} - \frac{1}{2}\chi_{f+g}(0)$, which yields

$$\mathsf{d}(f,g) = 2^{n+m-1} - \frac{1}{2}\sum_{x\in\mathbb{F}_2^n}\sum_{y\in\mathbb{F}_2^m}(-1)^{p(x)+G(x)y^t+q(y)} \tag{14}$$

$$= 2^{n+m-1} - \frac{1}{2}\sum_{x\in\mathbb{F}_2^n}(-1)^{p(x)}\chi_q\big(G(x)\big). \tag{15}$$

Based on the analysis of Section 2.2, we see that $\chi_q(v)$ is nonzero if and only if $v \in \mathsf{im}\, \Phi_{\mathcal{B}}$, with $\mathcal{B} = B + B^t$. Hence, as $G^{-1}(\mathsf{im}\, \Phi_{\mathcal{B}})$ is the union of the pre–images of the elements in $\mathsf{im}\, \Phi_{\mathcal{B}}$, we can write (15) as follows

$$\mathsf{d}(f,g) = 2^{n+m-1} - \frac{1}{2}\sum_{x\in G^{-1}(\mathsf{im}\ \Phi_{\mathcal{B}})}(-1)^{p(x)}\chi_q\big(G(x)\big). \tag{16}$$

Case 1: If $B = 0$, then $\mathsf{im}\, \Phi_{\mathcal{B}} = \{0\}$ and q is identically zero according to Remark 1, as $\lambda = b$ (the best linear approximation of the linear function by^t is the function itself); this implies that $\chi_q(0) = 2^m$ and (16) becomes

$$\begin{aligned}\mathsf{d}(f,g) &= 2^{n+m-1} - 2^{m-1}\sum_{x\in G^{-1}(0)}(-1)^{p(x)} \\ &\ge 2^{n+m-1} - 2^{m-1}\left|G^{-1}(0)\right|\end{aligned} \tag{17}$$

which holds with equality if $p|_{G^{-1}(0)}$ is identically zero. The fact that G is a quadratic PN function and (6) give the inequality $\left|G^{-1}(0)\right| \le 2^{n/2-m}(2^{n/2} + 2^m - 1)$, which leads to

$$\mathsf{d}(f,g) \ge 2^{n+m-1} - 2^{n/2-1}\left(2^{n/2} + 2^m - 1\right) \tag{18}$$

if combined with (17); this holds with equality if and only if $G^\star = 0$.

Case 2: Let us next assume $B \neq 0$ (hence $h \neq 0$). Then, we have $\chi_q(v) = (-1)^{q^\perp(v)} 2^{m-h}$ $\forall v \in \mathsf{im}\ \Phi_{\mathcal{B}}$ and substitution in (16) yields (*see* also [9])

$$\mathsf{d}(f,g) = 2^{n+m-1} - 2^{m-h-1} \sum_{x \in G^{-1}(\mathsf{im}\ \Phi_{\mathcal{B}})} (-1)^{(q^\perp \circ G)(x) + p(x)} \tag{19}$$

$$\geq 2^{n+m-1} - 2^{m-h-1} \sqrt{\sum_{x \in G^{-1}(\mathsf{im}\ \Phi_{\mathcal{B}})} \sum_{u \in x + G^{-1}(\mathsf{im}\ \Phi_{\mathcal{B}})} (-1)^{D_u(q^\perp \circ G)(x) + D_u p(x)}}. \tag{20}$$

In contrast with the case $B = 0$ and (17), we see from (19) that the trivial lower bound

$$\mathsf{d}(f,g) \geq 2^{n+m-1} - 2^{m-h-1} \left| G^{-1}(\mathsf{im}\ \Phi_{\mathcal{B}}) \right|$$

cannot be attained, as this would require that $(q^\perp \circ G + p)|_{G^{-1}(\mathsf{im}\ \Phi_{\mathcal{B}})}$ is identically zero, and therefore $\mathsf{deg}\, p = \mathsf{deg}(q^\perp \circ G) = 4$ —contradicting our hypothesis that p is quadratic. However, the minimization of the weight of $(q^\perp \circ G + p)|_{G^{-1}(\mathsf{im}\ \Phi_{\mathcal{B}})}$ becomes viable, if we note that both $G(x)$ and $G(x+u)$ in (20) take values in $\mathsf{im}\ \Phi_{\mathcal{B}}$, and thus

$$\begin{aligned} D_u(q^\perp \circ G)(x) &= q^\perp(G(x)) + q^\perp(G(x+u)) \\ &= q^\perp(y) + q^\perp(y+s) = D_s q^\perp(y) = D_s q^\perp(G(x)) \\ &= (D_s q^\perp \circ G)(x) \end{aligned}$$

where $y, s \in \mathsf{im}\ \Phi_{\mathcal{B}}$ are such that $G(x) = y$ and $G(x+u) = y+s \Leftrightarrow D_u G(x) = s$. If $\mathsf{c}(f,g)$ denotes the expression inside the square root of (20), the above allow us to rewrite it as

$$\begin{aligned} \mathsf{c}(f,g) &= \sum_{\substack{y \in \mathsf{im}\ \Phi_{\mathcal{B}} \\ s \in \mathsf{im}\ \Phi_{\mathcal{B}}}} \sum_{x \in G^{-1}(y)} \sum_{u \in x + G^{-1}(y+s)} (-1)^{(D_s q^\perp \circ G)(x) + D_u p(x)} \\ &= \sum_{s \in \mathsf{im}\ \Phi_{\mathcal{B}}} \sum_{x \in G^{-1}(\mathsf{im}\ \Phi_{\mathcal{B}})} \sum_{\substack{u \in \mathbb{F}_2^n : \\ D_u G(x) = s}} (-1)^{(D_s q^\perp \circ G)(x) + D_u p(x)} \end{aligned} \tag{21}$$

and our task becomes to minimize the weight of the quadratic function $D_s q^\perp \circ G + D_u p$, for most values of u, s. From (10), the fact that $q(0) = 0$, and that q has minimum weight (hence, 0 is a best linear approximation of q), we obtain

$$q^\perp(y\mathcal{B}) = q(y) \Rightarrow D_s q^\perp(y\mathcal{B}) = D_v q(y), \quad \forall y, v \in \mathsf{coim}\ \Phi_{\mathcal{B}}$$

with $s = v\mathcal{B}$; moreover $D_v q(y) = v\mathcal{B}y^t + q(v) = \varphi_v(y\mathcal{B}) + q(v)$. The above, and the fact that $G(x) \in \mathsf{im}\ \Phi_{\mathcal{B}}$ (compare with $y\mathcal{B}$), imply that (21) is written as

$$\mathsf{c}(f,g) = \sum_{v \in \mathsf{coim}\ \Phi_{\mathcal{B}}} \sum_{x \in G^{-1}(\mathsf{im}\ \Phi_{\mathcal{B}})} \sum_{\substack{u \in \mathbb{F}_2^n : \\ D_u G(x) = v\mathcal{B}}} (-1)^{(\varphi_v \circ G)(x) + D_u p(x) + q(v)}$$

$$= \sum_{v\in \mathsf{coim}\ \Phi_{\mathcal{B}}} \sum_{u\in\mathbb{F}_2^n} \sum_{\substack{x\in G^{-1}(\mathsf{im}\ \Phi_{\mathcal{B}})\\ \cap (D_uG)^{-1}(v\mathcal{B})}} (-1)^{(\varphi_v\circ G)(x)+D_up(x)+q(v)} \tag{22}$$

where the last equality is obtained by rearranging the two inner sums. By Proposition 1, and the analysis in Sections 2.1–2.2, the best affine approximations of $\varphi_v \circ G$, $v \neq 0$, are

$$\begin{aligned}\mathsf{A}(\varphi_v \circ G) &= \{D_u(\varphi_v \circ G) : u \in \mathbb{F}_2^n\} + G^\star(v)\\ &= \{\varphi_v \circ (D_uG) : u \in \mathbb{F}_2^n\} + G^\star(v)\end{aligned}$$

since $\varphi_v \circ G$ is bent, for $v \neq 0$. The restriction in the flat $(D_uG)^{-1}(v\mathcal{B})$ of the best affine approximation $(\varphi_v \circ D_uG)(x) + G^\star(v)$ is constant as it is equal to $v\mathcal{B}v^t + G^\star(v) = G^\star(v)$. Therefore, (22) is maximized if and only if the affine function $D_up + q(v)$ satisfies

$$\left.\begin{aligned}(D_up)|_{(D_uG)^{-1}(v\mathcal{B})} = 0\ \forall v \in \mathsf{coim}\ \Phi_{\mathcal{B}}\\ (G^\star + q)(v) = 0\ \forall v \in \mathsf{coim}\ \Phi_{\mathcal{B}}\end{aligned}\right\} \quad\Leftrightarrow\quad \begin{cases}(D_up)|_{(D_uG)^{-1}(\mathsf{im}\ \Phi_{\mathcal{B}})} = 0\\ (G^\star + q)|_{\mathsf{coim}\ \Phi_{\mathcal{B}}} = 0\end{cases} \tag{23}$$

for most values of u, v. It turns out that the first constraint of (23) can easily be satisfied, e.g. if we let $p = \varphi_\omega \circ G$ for some fixed vector $\omega \in \ker\ \Phi_{\mathcal{B}}$, since then for all $u \in \mathbb{F}_2^n$ we get

$$D_up(x) = v\mathcal{B}\omega^t = 0, \quad \forall x \in (D_uG)^{-1}(v\mathcal{B}), v \in \mathsf{coim}\ \Phi_{\mathcal{B}}$$

or equivalently $p|_{G^{-1}(\mathsf{im}\ \Phi_{\mathcal{B}})} = 0$. By this result, the fact that $q^\perp(v\mathcal{B}) = q(v)\ \forall v \in \mathsf{coim}\ \Phi_{\mathcal{B}}$, and (6), (19) we have

$$\begin{aligned}\mathsf{d}(f,g) &\geq 2^{n+m-1} - 2^{m-h-1} \sum_{x\in G^{-1}(\mathsf{im}\ \Phi_{\mathcal{B}})} (-1)^{(q^\perp\circ G)(x)}\\ &= 2^{n+m-1} - 2^{m-h-1} \sum_{v\in\mathsf{coim}\ \Phi_{\mathcal{B}}} (-1)^{q(v)} |G^{-1}(v\mathcal{B})|\\ &= 2^{n+m-1} - 2^{n/2-h-1} \sum_{v\in\mathsf{coim}\ \Phi_{\mathcal{B}}} (-1)^{q(v)} \left(2^{n/2} + \chi_{G^\star}(v\mathcal{B}) - 1\right).\end{aligned} \tag{24}$$

By comparing (19) and (22), we conclude that the second constraint of (23) enforces the *independent* minimization of the function's $(q^\perp \circ G)|_{G^{-1}(\mathsf{im}\ \Phi_{\mathcal{B}})}$ weight. Note that $q|_{\mathsf{coim}\ \Phi_{\mathcal{B}}}$ is a quadratic function on $2h$ variables, and by hypothesis its weight equals $2^{2h-1} - 2^{h-1}$. On the other hand, $G^\star|_{\mathsf{coim}\ \Phi_{\mathcal{B}}}$ is not quadratic in general. However, even if we manage to choose the mapping L such that $G^\star = q$ in order to satisfy (23) and minimize $\mathsf{d}(f,g)$ —in any other case, the value of $\mathsf{d}(f,g)$ would be higher according to the above analysis— and thus substitute $\chi_{G^\star}(v\mathcal{B}) = (-1)^{q(v)}2^{m-h}\ \forall v \in \mathsf{coim}\ \Phi_{\mathcal{B}}$ in (24), we eventually get

$$\begin{aligned}\mathsf{d}(f,g) &\geq 2^{n+m-1} - 2^{n/2-h-1} \sum_{v\in\mathsf{coim}\ \Phi_{\mathcal{B}}} (-1)^{q(v)} \left(2^{n/2} + (-1)^{q(v)}2^{m-h} - 1\right)\\ &= 2^{n+m-1} - 2^{n/2-1}\left(2^{n/2} + 2^m - 1\right)\end{aligned} \tag{25}$$

as a direct consequence of $\sum_{v\in\mathsf{coim}\ \Phi_{\mathcal{B}}} (-1)^{q(v)} = 2^h$, and the fact $|\mathsf{coim}\ \Phi_{\mathcal{B}}| = 2^{2h}$. Hence, from (18), (25) we conclude that the second–order nonlinearity of f satisfies (13). ■

Example 1. The Boolean function $f(x,y) = x_1x_3y_1 + x_1x_4y_2 + x_2x_3y_2 + x_2x_4y_1 + x_2x_4y_2$ corresponds to the last representative of cubic forms in $\mathsf{RM}(3,6)$ given in [17, Theorem 6.1]. We may write f as follows

$$\begin{aligned} f(x,y) &= (x_1x_3 + x_2x_4)y_1 + (x_2x_3 + (x_1+x_2)x_4)y_2 \\ &= f_1(x)y_1 + f_2(x)y_2 \end{aligned}$$

where $F = (f_1, f_2)$ is easily seen to be a quadratic PN function. As proved in [17], f has the highest second–order nonlinearity (i.e. 18) amongst all cubic functions on 6 variables. Indeed, we get $\mathsf{nl}_2(f) \geq 2^5 - 2(2^2 + 2^2 - 1) = 32 - 2 \cdot 7 = 18$ by Theorem 1, and thus the lower bound holds with equality. □

Corollary 1. *Let n be even, $m \leq n/2$, and $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ be an arbitrary quadratic PN function. The second–order nonlinearity of the cubic function $f(x,y) = F(x)y^t$ satisfies*

$$\mathsf{nl}_2(f) = 2^{n+m-1} - 2^{n/2-1}\big(2^{n/2} + 2^m - 1\big). \tag{26}$$

Moreover, if F is such that $F^\star$ is linear, then $\mathsf{nl}_2(f) = \mathsf{nl}(f)$.

Proof. It is well–known [3, 29], that the nonlinearity of Boolean functions that belong to the Maiorana–McFarland class satisfies

$$2^{n+m-1} - 2^{m-1} \max_{v \in \mathbb{F}_2^m} |F^{-1}(v)| \leq \mathsf{nl}(f) \leq 2^{n+m-1} - 2^{m-1} \left\lceil \sqrt{\max_{v \in \mathbb{F}_2^m} |F^{-1}(v)|} \right\rceil . \tag{27}$$

The particular class of functions f considered here, satisfies the lower bound of (27) with equality. Indeed $\mathsf{wt}(f) = 2^{n+m-1} - 2^{m-1}|F^{-1}(0)|$ by Proposition 3, and adding the linear function $\varphi_{(0,v)}(x,y)$ results in just using the pre–image $(F+v)^{-1}(0) = F^{-1}(v)$, instead of $F^{-1}(0)$, to compute the weight of $f + \varphi_{(0,v)}$. Therefore, if $F^\star$ (*see* Section 2.1) is a linear function, then from (6) and (27) we conclude that

$$\begin{aligned} \mathsf{nl}(f) &= 2^{n+m-1} - 2^{n/2-1}\big(2^{n/2} + \max_{v \in \mathbb{F}_2^m} \chi_{F^\star}(v) - 1\big) \\ &= 2^{n+m-1} - 2^{n/2-1}\big(2^{n/2} + 2^m - 1\big). \end{aligned} \tag{28}$$

By comparing (13) and (28), we find that Theorem 1 states $\mathsf{nl}_2(f) \geq \mathsf{nl}(f)$, which in turn yields $\mathsf{nl}_2(f) = \mathsf{nl}(f)$, due to the upper bound $\mathsf{nl}_2(f) \leq \mathsf{nl}(f)$. On the other hand, if F is such that $F^\star$ is arbitrary, then we can always add to f a quadratic function $L(x)y^t$, where $L : \mathbb{F}_2^n \to \mathbb{F}_2^m$ is an affine mapping, for which $(F+L)^\star = 0$ (*see* Remark 1 and the proof of Theorem 1). Therefore, the lower bound of Theorem 1 can always be attained, and the above lead to (26). ■

Theorem 2. *Let n be even, $m \leq n/2$, and $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ be an arbitrary quadratic PN function. The best quadratic approximations of the cubic function $f(x,y) = F(x)y^t$ are*

$$\mathsf{A}_2(f) \supseteq \{D_{(a,b)}(f+g) : (a,b) \in \mathbb{F}_2^n \times \mathbb{F}_2^m\} + g \tag{29}$$

where the quadratic function $g(x,y) = L(x)y^t$ is such that $(F+L)^\star = 0$.

Proof. As shown in proof of Theorem 1, there always exists a quadratic Boolean function $g(x,y) = L(x)y^t$, where $L : \mathbb{F}_2^n \to \mathbb{F}_2^m$ is an affine mapping, for which $(F+L)^\star = 0$. This in turn implies that all the nonzero linear combinations $\varphi_v \circ (F+L)$, $v \in \mathbb{F}_2^m$, have weight $2^{n-1} - 2^{n/2-1}$, and therefore (*see* proof of Corollary 1)

$$\begin{aligned}\mathsf{wt}(f+g) &= 2^{n+m-1} - 2^{m-1}|(F+L)^{-1}(0)| \\ &= 2^{n+m-1} - 2^{n/2-1}(2^{n/2} + 2^m - 1) = \mathsf{nl}_2(f+g).\end{aligned}$$

Then, (29) is a direct consequence of the fact that $(f+g)(x,y)$ and $(f+g)(x+a,y+b)$ have the same weight for all $(a,b) \in \mathbb{F}_2^n \times \mathbb{F}_2^m$. Hence, all quadratic functions of the form $D_{(a,b)}(f+g)+g$ are amongst the best quadratic approximations of f. ■

Remark 2. A well–known construction for quadratic PN functions $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$ is given in [26] that satisfy $F^\star = 0$. More precisely, if Δ is the companion–form matrix associated with a primitive linear feedback shift–register (LFSR) of length $n/2$, then

$$F(x) = F(z,w) = \left(zw^t, z\Delta w^t, \ldots, z\Delta^{m-1}w^t\right) = z\,\Psi(w) \tag{30}$$

where $z, w \in \mathbb{F}_2^{n/2}$ and Ψ an $\frac{n}{2} \times m$ matrix. In this case, any nonzero linear combination $\sum_{i=1}^m v_i \Delta^{i-1}$ has full rank and $\varphi_v \circ F$ is bent, with weight $2^{n-1} - 2^{n/2-1}$. □

Example 2. Let $f(x,y) = f_1(x)y_1 + f_2(x)y_2 + f_3(x)y_3$ be the cubic Boolean function on 9 variables provided in [1]

$$\begin{aligned}f(x) = {} & (x_3x_4 + x_1x_5 + (x_2+x_3)x_6)y_1 + (x_1x_4 + x_2x_5 + x_3x_6)y_2 \\ & + (x_2x_4 + (x_1+x_3)x_5 + x_1x_6)y_3\end{aligned}$$

which belongs to the Maiorana–McFarland class. Note that $F = (f_1, f_2, f_3)$ corresponds to a quadratic PN function obtained via the construction described in Remark 2, with

$$\Delta = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

of order 7 (since $\Delta^7 = I$). The coordinate functions of F are equal to $f_1(z,w) = z\Delta^{-1}w^t$, $f_2(z,w) = zw^t$ and $f_3(z,w) = z\Delta w^t$ respectively, where $z = (x_1,x_2,x_3)$, $w = (x_4,x_5,x_6)$. As proved in [1], f has the highest second–order nonlinearity (i.e. 196) amongst all cubic functions on 9 variables. Indeed, from Corollary 1 we also get that

$$\mathsf{nl}_2(f) = 2^8 - 2^2(2^3 + 2^3 - 1) = 256 - 4 \cdot 15 = 196.$$

Moreover, all nonzero linear combinations $\sum_{i=1}^3 v_i f_i$ yield bent functions of weight 28, i.e. we have $F^\star = 0$, and as a result, all the quadratic functions $D_{(a,b)}f$ $\forall (a,b) \in \mathbb{F}_2^6 \times \mathbb{F}_2^3$ are amongst the best quadratic approximations of f according to Theorem 2. □

Next, we prove a new lower bound on the maximum second–order nonlinearity of Boolean functions $f \in \mathbb{B}_n$, by applying the results of Theorem 1. The bound is tight, as shown in

Corollary 1, and close to the upper bound (4a) (*see* Table 1). Moreover, it improves (4b) for moderate values of n (in contrast with (4b), our bound is obtained by studying cubic functions only). The advantage in our case is that, an explicit construction is also given that attains this bound. We note that (31) agrees with the Ax–McEliece theorem [23], as $\mathsf{nl}_2(f) \equiv 0 \pmod{2^{\lfloor n/3 \rfloor}}$.

Corollary 2. *For $n \geq 3$, the maximum possible second–order nonlinearity amongst $f \in \mathbb{B}_n$ satisfies the lower bound*

$$\max_{f \in \mathbb{B}_n} \mathsf{nl}_2(f) \geq 2^{n-1}\left(1 - 2^{-\lfloor n/3 \rfloor}\right)\left(1 - 2^{-\lfloor n/3 \rfloor - \varepsilon}\right) \tag{31}$$

where $\varepsilon = 1$ if $n \equiv 2 \pmod 3$, and $\varepsilon = 0$ otherwise.

Proof. For each value of n, we seek for a cubic function $g(x, y) = F(x)y^t$ with at most n variables and for a quadratic PN function $F : \mathbb{F}_2^{2k} \to \mathbb{F}_2^m$, for some integers k, m, subject to the constraints $2k + m \leq n$ and $m \leq k$. The second–order nonlinearity of $g \in \mathbb{B}_{2k+m}$ is

$$\mathsf{nl}_2(g) = 2^{k-1}(2^k - 1)(2^m - 1)$$

according to Corollary 1. If $\|$ denotes the concatenation of (the truth tables of) Boolean functions, then the cubic function $f \in \mathbb{B}_n$ obtained by concatenating 2^{n-2k-m} copies of g, that is $f = g \,\|\, \cdots \,\|\, g$, has second–order nonlinearity

$$\begin{aligned} \mathsf{nl}_2(f) &= 2^{n-2k-m}\,\mathsf{nl}_2(g) = 2^{n-k-m-1}(2^k - 1)(2^m - 1) \\ &= 2^{n-1}(1 - 2^{-k})(1 - 2^{-m}) \end{aligned} \tag{32}$$

by recursively applying the identity $\mathsf{nl}_2(g \,\|\, g) = 2\,\mathsf{nl}_2(g)$ [9]. By using arguments as those given in [20], we conclude that (32) is maximized if and only if k, m take their maximum possible values subject to $|k - m| \leq 1$. Thus, if $n \not\equiv 2 \pmod 3$ we have that $k = m = \lfloor \frac{n}{3} \rfloor$ satisfy the constraints imposed above, whereas for $n \equiv 2 \pmod 3$ we get $k = \lfloor \frac{n}{3} \rfloor + 1$ and $m = \lfloor \frac{n}{3} \rfloor$. ∎

The exact values for the maximum second–order nonlinearity that a function $f \in \mathbb{B}_n$ can achieve (i.e. the covering radius of $\mathsf{RM}(2, n)$) are known only for $3 \leq n \leq 6$ [28]; its value is $1, 2, 6$ and 18 respectively. It is conjectured in [17] that the exact value of the maximum second–order nonlinearity is attained by a coset of $\mathsf{RM}(2, n)$ in $\mathsf{RM}(3, n)$ (i.e. by a cubic function). By the work of [1, 17] we know that for $7 \leq n \leq 9$ the maximum second–order nonlinearity of cubic functions equals $40, 88$ and 196 respectively. On the other hand, the bounds given in (4) are determined over all $f \in \mathbb{B}_n$ (*see* [8] and [10, Section 2]). However, by inspecting Table 1 and [20, Table I], we see that (31) is close to the upper bound (4a). Other general lower bounds, on the maximum second–order nonlinearity, are those given in [6, Theorem 1], [7, Theorem 2], [18, Proposition 5.1] and [24, Theorem 10], but are not included in the comparison as they are shown in [20, Table I] to be worse than a previous lower bound developed by the authors [20, Theorem 6] for separable cubic functions.

Several lower bounds on the maximum second–order nonlinearity, obtained by means of explicit Boolean function constructions, have been recently presented in the literature

Table 1: Lower bounds on the maximum second–order nonlinearity of cubic functions $f \in \mathbb{B}_n$.

n	(31)	[9]	[14]	[15]	[16]	[20]	[22]	[30]	[31]	[32]
3	1	1	–	–	–	1	–	–	–	–
4	2	2	–	2	–	2	–	–	2	2
5	6	6	–	4	5	6	6	1	4	4
6	18	12	15	10	10	14	16	10	17	8
7	36	36	30	20	32	28	36	19	34	16
8	84	72	60	52	64	68	78	64	84	62
9	196	176	120	104	166	148	166	128	168	124
10	392	352	378	256	331	296	351	330	386	248
11	840	802	756	512	768	664	737	661	772	496
12	1800	1604	1524	1187	1536	1400	1536	1535	1689	1318
13	3600	3468	3048	2374	3372	2800	3184	3071	3378	2636
14	7440	6936	7139	5296	6744	6032	6567	6742	7172	5272
15	15376	14605	14278	10592	14336	12496	13488	13485	14344	10544
16	30752	29210	28556	23027	28672	24992	27608	28669	29877	24561
17	62496	60517	57112	46054	59744	52576	56341	57341	59754	49122
18	127008	121034	122758	98304	119487	107744	114688	119482	122888	98244
19	254016	247951	245516	196608	245760	215488	232952	238968	245776	196488
20	512064	495902	491278	414071	491520	446528	472273	491513	501129	431562

[9, 14–16, 18, 20, 22, 30–32]. From these constructions, only the analysis provided in [20] is based on the ANF representation (1) of Boolean functions; in all other cases an equivalent representation, namely the *trace representation* of Boolean functions [23, Chapter 13], is used. Furthermore, all the constructions provide cubic functions (from [9, Section IV.D] we consider the modified Welch function, which is shown to have the best second–order nonlinearity among those given therein). The lower bounds in Table 1 are [14, Theorem 2], [15, Theorem 3], [16, Theorem 4], [22, Corollary 3], [30, Theorem 5], [31, Theorems 1,2], and [32, Theorem 1]. The entries with light gray color could not be obtained via the above constructions, due to restrictions in their parameters; we applied the identity $\mathsf{nl}_2(f \parallel f) = 2\,\mathsf{nl}_2(f)$ [9] to fill–in the gaps and facilitate the comparison. As depicted in Table 1, the second–order nonlinearity of the functions treated in this paper, clearly outer–performs the second–order nonlinearity of other known constructions.

4 Conclusions

Estimating cryptographic Boolean functions of guaranteed high second–order nonlinearity is known to be a difficult task. In this paper, it was proved that cubic functions $f(x, y) = F(x)y^t$ lying in the general Maiorana–McFarland class, with F being a perfect nonlinear function, achieve high second–order nonlinearity; its value was shown to be much higher than that obtained by other constructions (and close to an upper bound that was recently proved by Carlet *et al.*). Apart from their second–order nonlinearity, a subset of their best quadratic approximations was efficiently determined by means of their directional derivatives.

These results suggest that constructions based on perfect nonlinear mappings seem

to be the right way in order to obtain functions with high first–order and second–order nonlinearity. Our approach opens new directions for further research; other well–known constructions providing functions with high first–order nonlinearity need to be analyzed in terms of second–order nonlinearity, whereas extending these results to higher degree Boolean functions remains an open problem.

References

[1] E. Brier and P. Langevin, "Classification of boolean cubic forms in nine variables," in E. Biglieri and V. Tarokh (eds.) *IEEE Inform. Theory Workshop*, pp. 179–182, 2003. [results at: http://www.univ-tln.fr/~langevin/project/cubics/]

[2] P. Camion, C. Carlet, P. Charpin, and N. Sendrier, "On correlation-immune functions," in J. Feigenbaum (eds.) *Crypto 1991*, LNCS 576, pp. 86–100, Springer, Heidelberg, 1992.

[3] C. Carlet, "On the confusion and diffusion properties of Maiorana-McFarland's and extended Maiorana-McFarland's functions," *J. Complexity*, vol. 20, pp. 182–204, 2004.

[4] C. Carlet and J. Yucas, "Piecewise constructions of bent and almost optimal boolean functions," *Des. Codes Cryptogr.*, vol. 37, pp. 449–464, 2005.

[5] C. Carlet, "Vectorial boolean functions for cryptography," chapter in *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*, E.Y. Crama and P.L. Hammer (eds.) Cambridge University Press, pp. 398–469, 2010.

[6] C. Carlet, "On the higher order nonlinearities of algebraic immune functions," in C. Dwork (ed.) *Crypto 2006*, LNCS 4117, pp. 584–601, Springer, Heidelberg, 2006.

[7] C. Carlet, D. Dalai, K. Gupta, and S. Maitra, "Algebraic immunity for cryptographically significant Boolean functions: analysis and construction," *IEEE Trans. Inform. Theory*, vol. 52, pp. 3105–3121, 2006.

[8] C. Carlet and S. Mesnager, "Improving the upper bounds on the covering radii of binary Reed-Muller codes," *IEEE Trans. Inform. Theory*, vol. 53, pp. 162–173, 2007.

[9] C. Carlet, "Recursive lower bounds on the nonlinearity profile of boolean functions and their applications," *IEEE Trans. Inform. Theory*, vol. 54, pp. 1262–1272, 2008.

[10] G. Cohen and S. Litsyn, "On the covering radius of Reed-Muller codes," *Discrete Math.*, vol. 106–107, pp. 147–55, 1992.

[11] N. Courtois and W. Meier, "Algebraic attacks on stream ciphers with linear feedback," in E. Biham (eds.) *Eurocrypt 2003*, LNCS 2656, pp. 345–359, Springer, Heidelberg, 2003.

[12] J. Dillon, *Elementary Hadamard Difference Sets*. Ph.D. Thesis, University of Maryland, 1974.

[13] H. Dobbertin and G. Leander, "A survey of some recent results on bent functions," in T. Helleseth et al. (eds.) *Sequences and Their Applications*, LNCS 3486, pp. 1–29, Springer, Heidelberg, 2004.

[14] S. Gangopadhyay, S. Sarkar, and R. Telang, "On the lower bounds of the second order nonlinearity of some Boolean functions," *Inform. Sci.*, vol. 180, no. 2, pp. 266–273, 2010.

[15] M. Garg and S. Gangopadhyay, "Good second-order nonlinearity of a bent function via Niho power function," *IACR Cryptology ePrint Archive*, report 171, 2011. [available at: http://eprint.iacr.org/2011/171.pdf]

[16] R. Gode and S. Gangopadhyay, "On second order nonlinearities of cubic monomial Boolean functions," *IACR Cryptology ePrint Archive*, report 502, 2009. [available at: http://eprint.iacr.org/2009/502.pdf]

[17] X. Hou, "$GL(m, 2)$ acting on $R(r, m)/R(r-1, m)$," *Discrete Math.*, vol. 149, pp. 99–122, 1996.

[18] T. Iwata and K. Kurosawa, "Probabilistic higher order differential attack and higher order bent functions," in K.Y. Lam, E. Okamoto, and C. Xing (eds.) *Asiacrypt 1999*, LNCS 1716, pp. 62–74, Springer, Heidelberg, 1999.

[19] L. Knudsen and M. Robshaw, "Non-linear approximations in linear cryptanalysis," in U. Maurer (eds.) *Eurocrypt 1996*, LNCS 1070, pp. 224–236, Springer, Heidelberg, 1996.

[20] N. Kolokotronis, K. Limniotis, and N. Kalouptsidis, "Best affine and quadratic approximations of particular classes of boolean functions," *IEEE Trans. Inform. Theory*, vol. 55, no. 11, pp. 5211–5222, Nov. 2009.

[21] K. Kurosawa, T. Iwata, and T. Yoshiwara, "New covering radius of Reed-Muller codes for t-resilient functions," *IEEE Trans. Inf. Theory*, vol. 50, pp. 468–475, 2004.

[22] X. Li, Y. Hu, and J. Gao, "The lower bounds on the second order nonlinearity of cubic Boolean functions," *IACR Cryptology ePrint Archive*, report 009, 2010. [available at: http://eprint.iacr.org/2010/009.pdf]

[23] F. MacWilliams and N. Sloane, *The Theory of Error Correcting Codes*. North-Holland, Amsterdam, 1977.

[24] S. Mesnager, "Improving the lower bound on the higher order nonlinearity of Boolean functions with prescribed algebraic immunity," *IEEE Trans. Inform. Theory*, vol. 54, pp. 3656–3662, 2008.

[25] W.L. Millan, *Analysis and Design of Boolean Functions for Cryptographic Applications*. Ph.D. Thesis, Queensland University of Technology, 1997.

[26] K. Nyberg, "Perfect nonlinear s–boxes," in D. Davies (eds.) *Eurocrypt 1991*, LNCS 547, pp. 378–386, 1991.

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

[28] J. Schatz, "The second-order Reed-Muller code of length 64 has covering radius 18," *IEEE Trans. Inf. Theory*, vol. 27, pp. 529–530, 1981.

[29] J. Seberry, X. Zhang, and Y. Zheng, "On constructions and nonlinearity of correlation immune boolean functions," in T. Helleseth (eds.) *Eurocrypt 1993*, LNCS 765, pp. 181–199, Springer, Heidelberg, 1994.

[30] D. Singh, "Second order nonlinearities of some classes of cubic Boolean functions based on secondary constructions," *Int'l J. Comput. Sci. Inform. Technol.*, vol. 2, no. 2, pp. 786–791, 2011.

[31] G. Sun and C. Wu, "The lower bounds on the second order nonlinearity of three classes of Boolean functions with high nonlinearity," *Inform. Sci.*, vol. 179, no. 3, pp. 267–278, 2009.

[32] G. Sun and C. Wu, "The lower bound on the second–order nonlinearity of a class of Boolean functions with high nonlinearity," *Appl. Algebra Engrg. Comm. Comput.* (*AAECC*), vol. 22, pp. 37–45, 2011.