# Recursive lower bounds on the nonlinearity profile of Boolean functions and their applications

Claude Carlet *

**Abstract**

The nonlinearity profile of Boolean functions is an important cryptographic criterion, whose role against attacks on stream and block ciphers has been illustrated by many papers. We introduce a method for lower bounding its values and we deduce bounds on the second order nonlinearity for several classes of cryptographic Boolean functions, including the Welch and the inverse functions (which are used in the S-boxes of the AES). In the case of inverse function, we are able to bound the whole profile and to show the good behavior of this function with respect to this criterion.

**Keywords**: stream cipher, block cipher, Boolean function, nonlinearity profile

## 1 Introduction

Boolean functions are central objects for the design and the security of symmetric cryptosystems (stream ciphers and block ciphers), see [2, 3]. In cryptography, the most usual representation of these functions is the *algebraic normal form* (ANF):

$$f(x_1, \ldots, x_n) = \sum_{I \subseteq \{1,\ldots,n\}} a_I \prod_{i \in I} x_i,$$

where the $a_I$'s are in $F_2$. The terms $\prod_{i \in I} x_i$ are called *monomials*. The *algebraic degree* $d^\circ f$ of a Boolean function $f$ equals the maximum degree of those monomials whose coefficients are nonzero in its (unique) algebraic normal form. *Affine functions* are those Boolean functions of algebraic degrees at most 1.

A characteristic of Boolean functions, called their nonlinearity profile, plays an important role with respect to the security of the cryptosystems in which they are involved. Let $f : F_2^n \to F_2$ be an $n$-variable Boolean function. For every non-negative integer $r \leq n$, we denote by $nl_r(f)$ the minimum Hamming distance between $f$ and all functions of algebraic degrees at most $r$ (in the case of $r = 1$, we shall simply write $nl(f)$). In other words, $nl_r(f)$ equals the distance from $f$ to the Reed-Muller code $RM(r, n)$ of length $2^n$ and of order $r$. This parameter is called the *r-th order nonlinearity* of $f$ (simply the nonlinearity in the case $r = 1$). The *nonlinearity profile* of the function is the sequence of those values $nl_r(f)$ for $r$ ranging from 1 to $n - 1$.

The cryptographic relevance of this parameter has been illustrated by (e.g.) Courtois, Golic, Iwata-Kurosawa, Knudsen-Robshaw, Maurer and Millan [10, 14, 15, 17, 21, 22]. Very little is known on $nl_r(f)$ for $r > 1$. The best known upper bound [7] on $nl_r(f)$ has asymptotic version:

$$nl_r(f) = 2^{n-1} - \frac{\sqrt{15}}{2} \cdot (1 + \sqrt{2})^{r-2} \cdot 2^{n/2} + O(n^{r-2}).$$

*University of Paris 8 (MAATICAH). Also with INRIA, Projet CODES (Address: BP 105 - 78153, Le Chesnay Cedex, France). Email: claude.carlet@inria.fr

It can be proved [9, 4] that, for every positive real number such that $c^2 \log_2(e) > 1$ (e.g. for $c = 1$), there exist, for sufficiently large values of $n$, functions with $r$-th order nonlinearity greater than

$$2^{n-1} - c\sqrt{\sum_{i=0}^{r} \binom{n}{i}} \quad 2^{\frac{n-1}{2}} \approx 2^{n-1} - \frac{c\, n^{r/2}\, 2^{n/2}}{\pi^{1/4}\, r^{(2r+1)/4}\, 2^{3/4}}.$$

This proves that the best possible $r$-th order nonlinearity of $n$-variable Boolean functions is asymptotically equivalent to $2^{n-1}$, and that its difference with $2^{n-1}$ is polynomially proportional to $2^{n/2}$, whatever is the fixed value of $r$. But the proof of this fact is obtained by counting the number of functions having upper bounded $r$-th order nonlinearity (or more precisely by upper bounding this number) and it does not help obtaining explicit functions with non-weak $r$-th order nonlinearity.

Computing the $r$-th order nonlinearity of a given function with algebraic degree strictly greater than $r$ is a hard task for $r > 1$. Even the second order nonlinearity is unknown for all functions except for a few peculiar ones and for functions in small numbers of variables. A nice algorithm due to G. Kabatiansky and C. Tavernier and improved and implemented by Fourquet et al. [13, 16, 12] works well for $r = 2$ and $n \leq 11$ ( in some cases, $n \leq 13$), only. It can be applied for higher orders, but it is then efficient only for very small numbers of variables. No better algorithm is known.
Proving lower bounds on the $r$-th order nonlinearity of functions (and therefore proving their good behavior with respect to this criterion) is also a quite difficult task, even for the second order. Until recently, there had been only one attempt, by Iwata-Kurosawa [15], to construct functions with lower bounded $r$-th order nonlinearity. But the obtained value, $2^{n-r-3}(r+5)$, of the lower bound was small. A lower bound on the $r$-th order nonlinearity of functions with given algebraic immunity[1] has been given in [6] and improved in [5]. It gives better results than those of [15] for functions $f$ with good algebraic immunity $AI(f)$ (i.e. with $AI(f)$ not much smaller than $\lceil n/2 \rceil$), but the corresponding values of the lower bound, which is roughly equal to $\max\left(\sum_{i=0}^{AI(f)-r-1} \binom{n}{i}, 2\sum_{i=0}^{AI(f)-r-1} \binom{n-r}{i}\right)$, are small too.

In the present paper, we introduce a new method for lower bounding the nonlinearity profile of a given function and we deduce, for some classes of functions, explicit lower bounds on the second order nonlinearity (extendable in some cases to bounds on higher order nonlinearities, but the expressions become then more complex). Most interestingly, we obtain lower bounds for the whole nonlinearity profile of the inverse functions.
The paper is organized as follows. After some recalls and some simple observations done at Section 2, we give the general lower bounds at Section 3. We apply them at Section 4 to the Maiorana-McFarland functions, to the functions of univariate degree $2^t - 1$ on the field $F_{2^n}$, and to some classes of functions whose first order nonlinearities are known good (the Welch functions, some related functions, and the inverse functions), to deduce bounds on their second order nonlinearities. In Section 5, we obtain, for every $r$, a lower bound on the $r$-th order nonlinearity of the inverse function, and we deduce that it is asymptotically equivalent to $2^{n-1}$.

# 2 Some simple facts

In this section, we recall some known facts on the nonlinearity profile and we make some easy observations.

• Adding to a function $f$ a function of algebraic degree at most $r$ clearly does not change the $r$-th order nonlinearity of $f$.

[1] The algebraic immunity is a parameter quantifying the resistance to basic algebraic attacks.

• Since $RM(r,n)$ is invariant under affine automorphism, composing a Boolean function by an affine automorphism does not change its $r$-th order nonlinearity (i.e. $nl_r$ is affine invariant).

• The minimum distance of $RM(r,n)$ being equal to $2^{n-r}$ for every $r \leq n$, we have $nl_r(f) \geq 2^{n-r-1}$ for every function $f$ of algebraic degree exactly $r+1 \leq n$. Moreover, any minimum weight function $f$ of algebraic degree $r+1$ (that is, the indicator of any $(n-r-1)$-dimensional flat, see [20]), has $r$-th order nonlinearity equal to $2^{n-r-1}$ since a closest function of algebraic degree at most $r$ to $f$ is clearly the null function.

• As observed by Iwata and Kurosawa [15] (for instance), if $f_0$ is the $(n-1)$-variable restriction of $f$ to the linear hyperplane $H$ of equation $x_n = 0$ and $f_1$ the restriction of $f$ to the affine hyperplane $H'$ of equation $x_n = 1$, then we have $nl_r(f) \geq nl_r(f_0) + nl_r(f_1)$ since, for every function $g$ of algebraic degree at most $r$, the restrictions of $g$ to $H$ and $H'$ having both algebraic degree at most $r$, we have $d_H(f,g) \geq nl_r(f_0) + nl_r(f_1)$ where $d_H$ denotes the Hamming distance (obviously, this inequality is more generally valid if $f_0$ is the restriction of $f$ to any linear hyperplane $H$ and $f_1$ its restriction to the complement of $H$).

- Moreover, if $f_0 = f_1$, then there is equality since if $g$ is the best approximation of algebraic degree at most $r$ of $f_0 = f_1$, then $g$ now viewed as an $n$-variable function lies at distance $2nl_r(f_0)$ from $f$.

- Since $nl_r$ is affine invariant, this implies that, if there exists a nonzero vector $a \in F_2^n$ such that $f(x+a) = f(x)$, then the best approximation of $f$ by a function of algebraic degree $r$ is achieved by a function $g$ such that $g(x+a) = g(x)$ and $nl_r(f)$ equals twice the $r$-th order nonlinearity of the restriction of $f$ to any linear hyperplane $H$ excluding $a$.

- Note that the equality $nl_r(f) = 2nl_r(f_0)$ is also true if $f_0$ and $f_1$ differ by a function of algebraic degree at most $r-1$ since the function $x_n(f_0+f_1)$ has then algebraic degree at most $r$.

- The $r$-th order nonlinearity of the restriction of a function $f$ to a hyperplane is lower bounded by means of the $r$-th order nonlinearity of $f$ (this simple result will be a very useful tool in the sequel):

**Proposition 1** *Let $f$ be an $n$-variable Boolean function, $r$ a positive integer smaller than $n$ and $H$ an affine hyperplane of $F_2^n$. Then the $r$-th order nonlinearity of the restriction $f_0$ of $f$ to $H$ (viewed as an $(n-1)$-variable function) satisfies:*

$$nl_r(f_0) \geq nl_r(f) - 2^{n-2}.$$

*Proof*: We assume without loss of generality that $H = F_2^{n-1} \times \{0\}$. Let $g$ be any $(n-1)$-variable function of algebraic degree at most $r$. Let us extend it to an $n$-variable function of algebraic degree at most $r$, that we shall still denote by $g$. Then we have:

$$d_H(f_0, g) = 2^{n-2} - \frac{1}{2} \sum_{x \in H} (-1)^{f(x)+g(x)} =$$

$$2^{n-2} - \frac{1}{4} \left( \sum_{x \in F_2^n} (-1)^{f(x)+g(x)} + \sum_{x \in F_2^n} (-1)^{f(x)+g(x)+x_n} \right) =$$

$$2^{n-2} - \frac{1}{4} \left(2^n - 2d_H(f,g) + 2^n - 2d_H(f, g+x_n)\right) \geq -2^{n-2} + nl_r(f).$$

□

## 3 Lower bounds on the nonlinearity profile of a function by means of the nonlinearity profiles of its derivatives

*Notation*: We denote by $D_a f$ the so-called derivative of $f$ in the direction of $a \in F_2^n$:

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

Applying such discrete derivation several times to a function $f$ leads to the so-called higher order derivatives $D_{a_1} \cdots D_{a_k} f(x) = \sum_{u \in F_2^k} f(x + \sum_{i=1}^k u_i a_i)$.
Note that if $a_1, \cdots, a_k$ are not linearly independent then $D_{a_1} \cdots D_{a_k} f$ is null and, if they are linearly independent, then the set $\{x + \sum_{i=1}^k u_i a_i;\ u \in F_2^k\}$ is a $k$-dimensional flat. Note also that every derivation reduces the algebraic degree of $f$ at least by 1.

We first give a (rather weak, but tight) lower bound on the $r$-th order nonlinearity of any function $f$, knowing a lower bound on the $(r-1)$-th order nonlinearity of at least one of its derivatives (in nonzero directions).

**Proposition 2** *Let $f$ be an $n$-variable function and $r$ a positive integer smaller than $n$. For every nonzero $a \in F_2^n$, we have:*

$$nl_r(f) \geq \frac{1}{2} \max_{a \in F_2^n} nl_{r-1}(D_a f).$$

*Proof*: Let $a_0$ be an element such that $nl_{r-1}(D_{a_0} f) = \max_{a \in F_2^n} nl_{r-1}(D_a f)$. For every $n$-variable function $h$ of algebraic degree at most $r$, we have, denoting by $w_H$ the Hamming weight: $d_H(f,h) = w_H(f+g)$ and $w_H(D_{a_0}(f+h)) = d_H(D_{a_0} f, D_{a_0} h) \geq nl_{r-1}(D_{a_0} f)$, since the function $D_{a_0} h$ has algebraic degree at most $r-1$. So let us show that $w_H(f+h) \geq \frac{1}{2} w_H(D_{a_0}(f+h))$. Let $H$ be a linear hyperplane such that $a_0 \notin H$. The Hamming weight of the function $D_{a_0}(f+h)$ equals twice the Hamming weight of its restriction to $H$. For every $x \in H$ such that $D_{a_0}(f+h)(x) = 1$, either $x$ or $x + a_0$ belongs to the support of $f+h$. Hence, the Hamming weight of $f+h$ is at least half the Hamming weight of $D_{a_0}(f+h)$. This completes the proof. □

This bound is tight. Indeed, take for $f$ any Boolean function of algebraic degree $r+1$ and of Hamming weight $2^{n-r-1}$ (i.e. the indicator of any $(n-r-1)$-dimensional flat). The $r$-th order nonlinearity of $f$ equals its weight (see Section 2). The nonzero derivatives of $f$ are the indicators of $(n-r)$-dimensional flats and their $(r-1)$-th order nonlinearity equals their weight $2^{n-r}$.

Obviously, Proposition 2 can be repeatedly applied: for every $i$, we have

$$nl_r(f) \geq \frac{1}{2^i} \max_{a_1, \ldots, a_i \in F_2^n} nl_{r-i}(D_{a_1} \cdots D_{a_i} f).$$

This bound is also tight (take the same function as above). Proposition 2 and its iteration are particular cases of a more general result that we give in Appendix (see Proposition 8). But we see clearly the limitation of this approach since we do not get a bound which is equivalent to $2^{n-1}$.

We give now (in Corollary 1) a potentially stronger lower bound, valid when a lower bound on the $(r-1)$-th order nonlinearity is known for all the derivatives (in nonzero directions) of the function.

**Proposition 3** *Let $f$ be any $n$-variable function and $r$ a positive integer smaller than $n$. We have:*

$$nl_r(f) \geq 2^{n-1} - \frac{1}{2}\sqrt{2^{2n} - 2\sum_{a\in F_2^n} nl_{r-1}(D_a f)}.$$

*Proof*: Let $h$ be any $n$-variable function of algebraic degree at most $r$. We have:

$$\begin{aligned}\left(\sum_{x\in F_2^n}(-1)^{f(x)+h(x)}\right)^2 &= \sum_{x,y\in F_2^n}(-1)^{f(x)+f(y)+h(x)+h(y)}\\ &= \sum_{a\in F_2^n}\sum_{x\in F_2^n}(-1)^{f(x)+f(x+a)+h(x)+h(x+a)}\\ &= \sum_{a\in F_2^n}\sum_{x\in F_2^n}(-1)^{D_af(x)+D_ah(x)}.\end{aligned}$$

For every $a \in F_2^n$, the derivative $D_a h$ has algebraic degree at most $r-1$. Hence, we have $\sum_{x\in F_2^n}(-1)^{D_af(x)+D_ah(x)} = 2^n - 2d_H(D_af, D_ah) \leq 2^n - 2nl_{r-1}(D_af)$. This implies:

$$d_H(f,h) = 2^{n-1} - \frac{1}{2}\sum_{x\in F_2^n}(-1)^{f(x)+h(x)} \geq 2^{n-1} - \frac{1}{2}\sqrt{2^{2n} - 2\sum_{a\in F_2^n} nl_{r-1}(D_af)}.$$

□

This bound also is tight. Take for $f$ the indicator of any $(n-r-1)$-dimensional flat again. It has $2^{n-r-1}$ null derivatives (when $a$ belongs to the direction of the flat). The $2^n - 2^{n-r-1}$ nonzero derivatives of $f$ are the indicators of $(n-r)$-dimensional flats and have therefore $(r-1)$-th order nonlinearity $2^{n-r}$.
We deduce $2^{n-1} - \frac{1}{2}\sqrt{2^{2n} - 2\sum_{a\in F_2^n} nl_{r-1}(D_af)} = 2^{n-1} - \frac{1}{2}\sqrt{2^{2n} - (2^{n+1} - 2^{n-r})2^{n-r}} = 2^{n-1} - \frac{1}{2}\sqrt{(2^n - 2^{n-r})^2} = 2^{n-r-1} = nl_r(f)$.

**Remark**. It is not clear to us whether the bound of Proposition 3 is always better (or equal) than that of proposition 2: for every function $h$ of algebraic degree at most $r$, we have the inequality

$$\begin{aligned}2^{n-1} - \frac{1}{2}\sqrt{\sum_{a\in F_2^n}\sum_{x\in F_2^n}(-1)^{D_af(x)+D_ah(x)}} &= \min(d_H(f,h), d_H(f,h+1))\\ &\geq \frac{1}{2}\max_{b\in F_2^n} d_H(D_bf, D_bh)\end{aligned}$$

but when upper bounding $\sum_{x\in F_2^n}(-1)^{D_af(x)+D_ah(x)}$ by $2^n - 2\,nl_{r-1}(D_af)$ and lower bounding $d_H(D_bf, D_bh)$ by $nl_{r-1}(D_bf)$, we cannot know whether this inequality will remain true. However, we could not find examples where the bound of Proposition 3 is worse than that of Proposition 2.

**Corollary 1** *Let $f$ be an $n$-variable function and $r$ a positive integer smaller than $n$. Assume that, for some non-negative integers $K$ and $k$, we have $nl_{r-1}(D_af) \geq 2^{n-1} - K\,2^k$ for every nonzero $a \in F_2^n$, then*

$$nl_r(f) \geq 2^{n-1} - \frac{1}{2}\sqrt{(2^n-1)K2^{k+1} + 2^n} \approx 2^{n-1} - \sqrt{K}\,2^{(n+k-1)/2}.$$

*Proof*: According to Proposition 3, we have

$$\begin{aligned} nl_r(f) &\geq 2^{n-1} - \frac{1}{2}\sqrt{2^{2n} - 2(2^n-1)(2^{n-1} - K\,2^k)} \\ &= 2^{n-1} - \frac{1}{2}\sqrt{(2^n-1)K\,2^{k+1} + 2^n}. \end{aligned}$$

□

**Remark**. Let $f$ have algebraic degree exactly 3. Proposition 2 implies only that $nl_2(f) \geq 2^{n-3}$. If we assume that all the derivatives $D_a f$, $a \neq 0$ have algebraic degree exactly 2, then Corollary 1 with $K = 1$ and $k = n-2$ implies that $nl_2(f) \geq 2^{n-1} - \frac{1}{2}\sqrt{(2^n-1)2^{n-1} + 2^n} = 2^{n-1} - \frac{1}{2}\sqrt{(2^n+1)2^{n-1}} \approx 2^{n-1} - 2^{n-3/2}$, which is slightly stronger. Note that $n$-variable cubic functions whose derivatives $D_a f$, $a \neq 0$ all have algebraic degree 2 do exist for $n \geq 5$, since the number of functions of algebraic degrees at most 3 equals $2^{\binom{n}{3}+\binom{n}{2}+n+1}$, the number of functions of algebraic degrees at most 3 having at least one affine derivative is upper bounded by $(2^n-1)\,2^{\binom{n-1}{3}+\binom{n-1}{2}+2n}$ (indeed, such function is an affine-type extension of a function of algebraic degree at most 3 on a linear hyperplane of $F_2^n$) and the difference between these two numbers is strictly positive for $n \geq 5$.

Applying two times Proposition 3, we obtain the bound

$$nl_r(f) \geq 2^{n-1} - \frac{1}{2}\sqrt{\sum_{a\in F_2^n}\sqrt{2^{2n} - 2\sum_{b\in F_2^n} nl_{r-2}(D_aD_bf)}}.$$

Applying it $\ell$ times, we get

$$nl_r(f) \geq 2^{n-1} - \frac{1}{2}\sqrt{\sum_{a_1\in F_2^n}\sqrt{\sum_{a_2\in F_2^n}\cdots\sqrt{2^{2n} - 2\sum_{a_\ell\in F_2^n} nl_{r-\ell}(D_{a_1}\cdots D_{a_\ell}f)}}}.$$

# 4 Functions with provably lower bounded second order nonlinearity

We study now the main classes of Boolean functions which are used in cryptography: the Maiorana-McFarland functions (which have led to many constructions of functions allowing good trade-off between several cryptographic criteria, such as nonlinearity, resiliency ...) and the Boolean functions $tr(\lambda F(x))$ where $F$ is a vectorial Boolean function over the field $F_{2^n}$ whose nonlinearity is provably high.

## 4.1 Maiorana-McFarland functions

Let $k$ be a positive integer smaller than $n$, let $g$ be a Boolean function on $F_2^{n-k}$ and let $\phi$ be a mapping from $F_2^{n-k}$ to $F_2^k$. Set:

$$f_{\phi,g}(x,y) = x\cdot\phi(y) + g(y),\ x \in F_2^k,\ y \in F_2^{n-k}$$

where "·" is the usual inner product in $F_2^k$.

We have (see e.g. [2]):

$$nl_1(f_{\phi,g}) \geq 2^{n-1} - 2^{k-1}\max_{u\in F_2^k}|\phi^{-1}(u)|, \tag{1}$$

where $|\phi^{-1}(u)|$ denotes the size of $\phi^{-1}(u)$. Any derivative of such Maiorana-McFarland function is a Maiorana-McFarland function: for every $a \in F_2^k$ and every $b \in F_2^{n-k}$, we have $D_{a,b}\left(f_{\phi,g}(x,y+b)\right) = x \cdot D_b\,\phi(y) + a \cdot \phi(y) + D_b\,g(y) = f_{D_b\,\phi, a\cdot\phi + D_b g}(x,y)$. Note that for $b = 0$, we have $\max_{u \in F_2^k} |(D_b\phi)^{-1}(u)| = (D_b\phi)^{-1}(0)| = 2^{n-k}$. We deduce from Proposition 3 and from Relation (1) that

$$nl_2(f_{\phi,g}) \geq$$

$$2^{n-1} - \frac{1}{2}\sqrt{2^{2n} - 2\,(2^n - 2^k)(2^{n-1} - 2^{k-1} \max_{u \in F_2^k, b \in (F_2^{n-k})^*} |(D_b\,\phi)^{-1}(u)|)} =$$

$$2^{n-1} - \frac{1}{2}\sqrt{2^{n+k} + 2^k(2^n - 2^k) \max_{u \in F_2^k, b \in (F_2^{n-k})^*} |(D_b\,\phi)^{-1}(u)|}.$$

Similar bounds on $nl_r(f_{\phi,g})$ can also be given.

## 4.2 Functions of univariate degree $2^t - 1$

Let us identify $F_2^n$ with the field $F_{2^n}$ and denote by $tr$ the absolute trace function over $F_2^n$. Let $t \leq n$ be a positive integer and $g(x)$ a univariate polynomial of degree $2^t - 1$ over $F_2^n$. Let $f(x) = tr(g(x))$. Then every derivative $D_a f$, $a \neq 0$, is the trace of a univariate function of degree $2^t - 2$ and equals in fact the trace of a univariate function of degree at most $2^t - 3$, after reduction using the equality $tr(y^2) = tr(y)$. The term in $x^{2^t-3}$ can come from $x^{2^t-1} + (x+a)^{2^t-1}$ only, and thus cannot vanish. Hence, according to the Weil bound [19], its first-order nonlinearity is then at least $2^{n-1} - (2^t - 4)2^{n/2-1}$. Corollary 1 with $K = 2^t - 4$ and $k = n/2 - 1$ implies that $nl_2(f) \geq 2^{n-1} - \frac{1}{2}\sqrt{(2^n - 1)(2^t - 4)2^{n/2} + 2^n} \approx 2^{n-1} - 2^{3n/4+t/2-1}$.

## 4.3 The Welch function

The vectorial Welch function $x \to x^{2^t+3}$, where $t = \frac{n-1}{2}$, $n$ odd, is an AB function, i.e. has the best possible nonlinearity as a vectorial function [1]. It is a permutation. So all the Boolean functions $tr(\lambda x^{2^t+3})$, $\lambda \neq 0$, are affinely equivalent to each others (through the automorphisms $x \to \mu x$). We shall therefore study only the function $tr(x^{2^t+3})$ that we shall denote by $f_{welch}(x)$. The second order nonlinearity of this function is good, for all the values of $n$ for which it could be computed; we shall see at Subsection 4.5 that it is slightly better than that of the inverse function (for instance, for $n = 9$, it equals 184, according to [13, 16, 12]). Note however that this function cannot be used as a cryptographic function since its algebraic degree (which equals, for any such function, the 2-weight of the exponent, i.e., the number of 1's in its binary expansion, that is 3 here for the exponent $2^t + 3$) is too low and does not allow resistance to higher order differential cryptanalyses. Nevertheless, let us determine a lower bound on the first-order nonlinearities of its derivatives, in order to compare what we get thanks to Corollary 1 with the actual values of its second order nonlinearity obtained by running a computer.

**Lemma 1** *Any derivative, in a nonzero direction, of the function $f_{welch}(x) = tr(x^{2^t+3})$ has nonlinearity at least $2^{n-1} - 2^{\frac{n+3}{2}}$.*

*Proof*: A straightforward calculation (which was the starting point of Dobbertin's proof of the almost perfect nonlinearity of the Welch function [11]) gives for every nonzero $a \in F_2^n$ that, denoting $r = t + 1$, we have $D_a f_{welch}(ax) = tr(a^{2^t+3}[q(x + x^{2^t}) + 1])$, where $q(x) = x^{2^r+1} + x^3 + x$.
The function $g_a(x) = tr(a^{2^t+3}(q(x+x^{2^t}))$ is such that $g_a(x+1) = g_a(x)$. According to what we have seen at Section 2, this implies that $nl(D_a f_{welch})$ equals twice the nonlinearity of

the restriction of $g_a$ to the linear hyperplane $H = \{x \in F_2^n \,/\, tr(x) = 0\}$ (indeed, $H$ excludes 1 since $n$ is odd). Since the function $x \in H \to x + x^{2^t}$ is a linear automorphism of $H$, $nl(D_a f_{welch})$ therefore equals twice the $r$-th order nonlinearity of the restriction of $tr(a^{2^t+3} q(x))$ to $H$.
Let us denote $b = a^{2^t+3}$. The nonlinearity of the $n$-variable quadratic function $tr(bq(x))$ equals $2^{n-1} - 2^{\frac{n+k}{2}-1}$ where $k$ is the dimension of the vectorspace $\mathcal{E} = \{x \in F_2^n \,/\, \forall y \in F_2^n,\ tr(b\,q(x)) + tr(b\,q(y)) + tr(b\,q(x+y)) = 0\}$ and has same evenness as $n$ (see [20, 2]). We have $tr(b\,q(x)) + tr(b\,q(y)) + tr(b\,q(x+y)) = tr(b\,(x^{2^r} + x^2)y + b\,(y^{2^r} + y^2)x) = tr([b\,(x^{2^r} + x^2) + b^{2^t} x^{2^t} + b^{2^{n-1}} x^{2^{n-1}}]y)$, since $r + t = n$, $tr(u^2) = tr(u)$ and $u^{2^n} = u$, for every $u \in F_2^n$. We deduce that $\mathcal{E} = \{x \in F_2^n \,/\, b\,(x^{2^r} + x^2) + b^{2^t} x^{2^t} + b^{2^{n-1}} x^{2^{n-1}} = 0\} = \{x \in F_2^n \,/\, b^2(x^{2^{r+1}} + x^4) + b^{2^r} x^{2^r} + b\,x = 0\}$. We use now the multivariate method initiated in numerous papers by H. Dobbertin. Let us denote $y = x^{2^r}$ and $d = b^{2^r}$, then the equation becomes

$$E1: \quad b^2 y^2 + d\,y = b^2 x^4 + b\,x.$$

Squaring gives

$$E2: \quad b^4 y^4 + d^2 y^2 = b^4 x^8 + b^2 x^2$$

and raising $E1$ to the $2^r$ power gives

$$E3: \quad d^2 x^4 + b^2 x^2 = d^2 y^4 + d\,y \text{ , i.e.,} \quad d^2 y^4 + d\,y = d^2 x^4 + b^2\,x^2.$$

The square root (that is, the $2^{n-1}$-th power) of equation $E1 + E3$ is

$$E'1: \quad d\,y^2 + b\,y = b\,x^2 + (b\,x)^{2^{n-1}} + d\,x^2 + b\,x.$$

The equation $b^4 E3 + d^2 E2$ gives

$$E'2: \quad d^4 y^2 + b^4 d\,y = b^4 d^2 x^4 + b^6 x^2 + b^4 d^2 x^8 + b^2 d^2 x^2.$$

The equation $b^2 E'1 + d\,E1$ gives

$$E''1: \quad (b^3 + d^2)y = b^3 x^2 + b^2 (b\,x)^{2^{n-1}} + b^2 d\,x^2 + b^3 x + b^2 d\,x^4 + b\,d\,x$$

and the equation $d^4 E1 + b^2 E'2$ gives

$$E''2: \quad (d^5 + b^6 d)y = b^2 d^4 x^4 + b\,d^4 x + b^6 d^2 x^4 + b^8 x^2 + b^6 d^2 x^8 + b^4 d^2 x^2.$$

The square of the equation obtained by elimination of $y$ between the two equations $E''1$ and $E''2$ gives an equation of degree 16 in $x$. Hence, we have $k \leq 4$ and therefore $k \leq 3$ since $n$ is odd. Applying then Proposition 1, we deduce that the first-order nonlinearity of $D_a f_{welch}$ is at least $2\,(2^{n-1} - 2^{\frac{n+1}{2}} - 2^{n-2}) = 2^{n-1} - 2^{\frac{n+3}{2}}$. $\square$

Corollary 1 with $K = 1$ and $k = \frac{n+3}{2}$ gives then:

**Proposition 4** *Let $f_{welch}(x) = tr(x^{2^t+3})$, $t = \frac{n-1}{2}$. Then we have:*

$$nl_2(f_{welch}) \geq 2^{n-1} - \frac{1}{2}\sqrt{(2^n - 1)2^{\frac{n+5}{2}} + 2^n} \approx 2^{n-1} - 2^{(3n+1)/4}.$$

In Table 1, for $n$ ranging from 5 to 13, we indicate the values given by this bound, compared with the actual values, computed by running a computer, with an algorithm due to G. Kabatiansky and C. Tavernier and improved and implemented by Fourquet et al. [13, 16, 12]. For values of $n$ smaller than 5, the bound gives negative numbers and for values greater than 13, the algorithm is unable to produce results. Note that Proposition 4 gives an approximation of the actual value which is proportionally better and better when $n$ increases. Moreover, the difference between $2^{n-1}$ and our bound equals twice the difference between $2^{n-1}$ and the actual value, in average for $5 \leq n \leq 13$. In Table 2 we give, for $n = 15$ and 17, the values given by our bound, compared with upper bounds obtained by Fourquet et al. [13, 16, 12].

| $n$ | 5 | 7 | 9 | 11 | 13 |
|---|---|---|---|---|---|
| the bound | 0 | 19 | 128 | 662 | 3072 |
| the actual values | 6 | 36 | 184 | 848 | between 3360 and 3696 |
| % | 0 | 53 | 70 | 78 | between 83 and 91 |

Table 1: The values of the lower bound on $nl_2(f_{welch})$ given by Proposition 4, the actual values and the ratio

| $n$ | 15 | 17 |
|---|---|---|
| the lower bound | 13487 | 57343 |
| overestimation of the values | 15488 | 63680 |
| % | 87 | 90 |

Table 2: The values of the lower bound on $nl_2(f_{welch})$ given by Proposition 4, an overestimation of the actual values and the ratio

### 4.4 A power function with better second-order nonlinearity

We study now a function which is similar to the Welch function, but whose second order nonlinearity computed in [13, 16, 12] gives better results than for the Welch function. The Boolean function $f_{welch'}(x) = tr(x^{2^r+3})$, $r = \frac{n+1}{2}$ (that we shall call the modified-Welch function) has derivatives $D_a f_{welch'}(x) = tr(a\,x^{2^r+2} + a^2x^{2^r+1} + a^{2^r}x^3) + \ell(x)$ where $\ell$ is affine. The nonlinearity of this quadratic function equals $2^{n-1} - 2^{\frac{n+k}{2}-1}$ where $k$ is the dimension of the vectorspace $\mathcal{E} = \{x \in F_2^n \,/\, a^{2^{n-1}}x^{2^{r-1}} + a^{2^{r-1}}x^{2^r} + a^2x^{2^r} + a^{2^r}x^{2^{r-1}} + a^{2^r}x^2 + a^{2^{r-1}}x^{2^{n-1}} = 0\}$. We denote $y = x^{2^r}$ and $b = a^{2^r}$. The square of the equation above becomes:

$$E1: \quad (a+b^2)y + (b+a^4)y^2 + b^2x^4 + bx = 0.$$

The square of $E1$ is:

$$E2: \quad (a^2+b^4)y^2 + (b^2+a^8)y^4 + b^4x^8 + b^2x^2 = 0$$

and its $2^r$-th power is:

$$E3: \quad (b+a^4)x^2 + (a^2+b^4)x^4 + a^4y^4 + a^2y = 0.$$

Eliminating $y^4$ from equations $E2$ and $E3$ gives the following equation $E'1$:

$$(a^6+a^4b^4)y^2 + (a^{10}+a^2b^2)y + a^4b^4x^8 + (a^2+b^4)(b^2+a^8)x^4 + (a^4b^2+(b+a^4)^3)x^2 = 0.$$

Eliminating $y$ from $E1$ and $E3$ and taking the square root of the resulting equation gives

$$E'2: \quad (a^{5\cdot 2^{n-1}} + a^2b)y^2 + (ab^{2^{n-1}} + a^3)y + (ab + (a+b^2)^{3\cdot 2^{n-1}})x^2 + (a^{2^{n-1}}+b)(b^{2^{n-1}}+a^2)x + ab^{2^{n-1}}x^{2^{n-1}} = 0.$$

Eliminating then $y^2$ from $E'1$ and $E'2$ gives an equation $E''1$ in $y$, $x^8$, $x^4$, $x^2$, $x$ and $x^{2^{n-1}}$. Eliminating $y^2$ from equations $E1$ and $E'1$ gives and equation $E''2$ in $y$, $x^8$, $x^4$, $x^2$ and $x$. Eliminating $y$ from $E''1$ and $E''2$ and squaring the resulting equation gives an equation of degree 16 in $x$. This shows that $k \leq 3$. We deduce that the nonlinearity of $D_a f_{welch'}$ is at least $2^{n-1} - 2^{\frac{n+1}{2}}$ and Corollary 1 with $K = 1$ and $k = \frac{n+1}{2}$ gives then:

**Proposition 5** *Let $f_{welch'}(x) = tr(x^{2^r+3})$, $r = \frac{n+1}{2}$. Then we have:*

$$nl_2(f_{welch'}) \geq 2^{n-1} - \frac{1}{2}\sqrt{(2^n-1)2^{\frac{n+3}{2}} + 2^n} \approx 2^{n-1} - 2^{(3n-1)/4}.$$

**Remark**. The methods we used for lower bounding the second order nonlinearities of the Welch functions and of the modified-Welch functions are not exactly the same. In fact, the method used to prove Proposition 4 is slightly more complex than that used for Proposition 5. This is because the method of Proposition 5 gives worse results in the case of the Welch function. In the case of the modified-Welch function, both methods give the same result and we presented the simplest one.

The bound of Proposition 5 is better than for the Welch function. And actually, for $n = 9$, we can see in Table 3 below that the value of $nl_2(f_{welch'})$ is 188 as shown in [13, 16, 12], which is better than for the Welch function (that is, 184). Note at the last line of Table 3 that our bound is better than in the case of the Welch function. The difference between $2^{n-1}$ and our bound is in average 1.5 times the difference between $2^{n-1}$ and the actual value (for these values of $n$). Finally, note that our bound gives a lower bound for $n = 13$ which is better than what could give the algorithm.
In Table 4 we give, for $n = 15$ and 17, the values given by our bound, compared with upper bounds obtained by Fourquet et al. [13, 16, 12].

| $n$ | 5 | 7 | 9 | 11 | 13 |
|---|---|---|---|---|---|
| the bound | 5 | 32 | 165 | 768 | 3371 |
| the actual values | 6 | 36 | 188 | 848 | between 3300 and 3696 |
| % | 83 | 89 | 88 | 90 | between 91 and 100 |

Table 3: The values of the lower bound on $nl_2(f_{welch'})$ given by Proposition 5, the actual values and the ratio

| $n$ | 15 | 17 |
|---|---|---|
| the lower bound | 14335 | 59741 |
| overestimation of the values | 15504 | 63648 |
| % | 92 | 94 |

Table 4: The values of the lower bound on $nl_2(f_{welch'})$ given by Proposition 6, an overestimation of the actual values and the ratio

## 4.5 The inverse function

Let us consider the so-called inverse functions $f_\lambda(x) = tr(\lambda x^{2^n-2})$, where $\lambda$ is any element of $F_{2^n}^*$ and where $n$ is any positive integer. Here again, all the Boolean functions $f_\lambda$, $\lambda \neq 0$, are affinely equivalent to each others. We shall write $f_{inv}$ for $f_1$. But we shall need however the notation $f_\lambda$ in the calculations below. We have $f_\lambda(x) = tr\left(\frac{\lambda}{x}\right)$, with the convention that $\frac{\lambda}{0} = 0$ (we shall always assume this kind of convention in the sequel). Recall that the component functions of the S-boxes of the AES are of this type (with $n = 8$).
We shall be able to obtain a lower bound for the whole nonlinearity profile of $f_{inv}$.

For every nonzero $a \in F_{2^n}$, we have $(D_a f_\lambda)(ax) = tr\left(\frac{\lambda}{ax} + \frac{\lambda}{ax+a}\right) = tr\left(\frac{\lambda/a}{x^2+x}\right) = f_{\lambda/a}(x^2+x)$ if $x \notin F_2$ and $(D_a f_\lambda)(ax) = tr(\lambda/a)$ if $x \in F_2$. We deduce that, for every $r$, we have $nl_r(D_a f_\lambda) \geq nl_r(g_{\lambda/a}) - 2$, where $g_{\lambda/a}(x) = f_{\lambda/a}(x^2 + x)$ is such that $g_{\lambda/a}(x + 1) = g_{\lambda/a}(x)$. We have seen at Section 2 that this implies that $nl_r(g_{\lambda/a})$ equals twice the $r$-th order nonlinearity of the restriction of $g_{\lambda/a}$ to any linear hyperplane $H$ excluding 1. Since the function $x \to x^2 + x$ is a linear isomorphism from $H$ to the hyperplane $\{x \in F_{2^n} \,/\, tr(x) = 0\}$, we see that $nl_r(g_{\lambda/a})$ equals twice the $r$-th order nonlinearity of the restriction of $f_{\lambda/a}$ to this hyperplane. Applying then Proposition 1, we deduce that

$$nl_r(D_a f_\lambda) \geq 2\, nl_r(f_{\lambda/a}) - 2^{n-1} - 2. \tag{2}$$

The first order nonlinearity of the inverse function is lower bounded by $2^{n-1} - 2^{n/2}$ (it equals this value if $n$ is even). It has been more precisely proven in [18] that the character sums $\sum_{x\in F_{2^n}} (-1)^{f_\lambda(x)+tr(ax)}$, called Kloosterman sums, can take any value divisible by 4 in the range $[-2^{n/2+1}+1, 2^{n/2+1}+1]$.
We deduce:

**Lemma 2** *Every derivative (in a nonzero direction) of any inverse Boolean function has first-order nonlinearity at least* $2^{n-1} - 2^{n/2+1} - 2$.

**Remark**.
In [8] is proven that, when $a$ ranges over $F_{2^n}^*$, the values of the sums $\sum_{x\in F_{2^n}} (-1)^{D_a f_{inv}(x)}$ are all integers divisible by 8 in the range $[-2^{n/2+1} - 3, 2^{n/2+1} + 1]$. Nothing is proven for the sums $\sum_{x\in F_{2^n}} (-1)^{D_a f_{inv}(x)+tr(bx)}$. This property of the former sums cannot be extended to all of the latter, since the derivatives of the inverse Boolean function would then have nonlinearities at least $2^{n-1} - 2^{n/2} - 1$ and this would lead, thanks to Corollary 1, to a lower bound on the second order nonlinearity of this function which is in contradiction with the actual values given at Table 5. Is it possible to prove that some of the derivatives of $f_{inv}$ have nonlinearities at least $2^{n-1} - 2^{n/2} - 1$? The nice idea of [8] for proving the result in the case $b = 0$ does not seem to work for $b \neq 0$: denoting $y = x^{2^n-2}$ and observing that $(D_a f_\lambda)(ax) = tr\left(\frac{\lambda\, y^2}{a(y+1)}\right) = tr\left(\frac{\lambda}{a}(y+1) + \frac{\lambda}{a(y+1)}\right)$ brings back to Kloosterman sums when $b = 0$, but when $b \neq 0$, we have $(D_a f_\lambda)(ax) + tr(bx) = tr\left(\lambda(y+1) + \frac{\lambda}{a(y+1)} + \frac{b}{y}\right)$ and this leads to a sum which is more complex than a Kloosterman sum.

Applying Corollary 1 with $r = 2$, $K = 2^{n/2+1} + 2$ and $k = 0$, we deduce:

**Proposition 6** *Let* $f_{inv}(x) = tr(x^{2^n-2})$. *Then we have:*

$$nl_2(f_{inv}) \geq 2^{n-1} - \frac{1}{2}\sqrt{(2^n-1)(2^{n/2+2}+4) + 2^n} \approx 2^{n-1} - 2^{3n/4}.$$

In Table 5, for $n$ ranging from 4 to 12 (for smaller values of $n$, the bound gives negative numbers), we indicate the values given by this bound, compared with the actual values computed by Fourquet et al. [13, 16, 12]. Note that, here again, Proposition 6 gives an approximation of the actual value which is proportionally better and better when $n$ increases. In fact, the approximation is better than for the Welch function. The difference between $2^{n-1}$ and our bound is in average 1.5 times the difference between $2^{n-1}$ and the actual value (for these values of $n$).
In Table 6 we give, for $n = 13, 14$ and 15, the values given by our bound, compared with upper bounds obtained by Fourquet et al. [13, 16, 12].

| $n$ | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|
| the bound | 0 | 2 | 8 | 25 | 62 | 146 | 328 | 716 | 1532 |
| the values | 2 | 6 | 14 | 36 | 82 | 182 | 392 | 842 | between 1720 and 1776 |
| % | 0 | 33 | 52 | 69 | 76 | 80 | 84 | 85 | between 86 and 89 |

Table 5: The values of the lower bound on $nl_2(f_{inv})$ given by Proposition 6, the actual values and the ratio

| $n$ | 13 | 14 | 15 |
|---|---|---|---|
| the lower bound | 3230 | 6737 | 13941 |
| overestimation of the values | 3696 | 7580 | 15506 |
| % | 87 | 89 | 90 |

Table 6: THE VALUES OF THE LOWER BOUND ON $nl_2(f_{inv})$ GIVEN BY PROPOSITION 6, AN OVERESTIMATION OF THE ACTUAL VALUES AND THE RATIO

### 4.6 Remark on the Kasami function

Determining or efficiently lower bounding the first-order nonlinearities of the derivatives of Kasami functions is an open problem, and we could not obtain a lower bound on its nonlinearity profile by using Corollary 1. When $n$ is odd, an obvious observation is that, for every Boolean function $g$ of algebraic degree strictly less than the algebraic degree $k+1$ of the Kasami function $f(x) = tr(ax^{2^{2k}-2^k+1})$, $gcd(k,n) = 1$, the Hamming distance between the functions $f$ and $g$ is equal to the Hamming weight of the function $tr(ax^{2^{3k}+1} + g(x^{2^k+1}))$. Indeed, the mapping $x \to x^{2^k+1}$ is a permutation and $f(x^{2^k+1}) = tr(ax^{2^{3k}+1})$. This function has algebraic degree at most $2r$ when the algebraic degree of $g$ is at most $r \leq k$. Since the function $f+g$ has algebraic degree $k+1$ under this same condition, we deduce that $nl_r(f) \geq \max(2^{n-2r}, 2^{n-k-1})$, for every $r \leq k$. The second order nonlinearities of Kasami functions seem worse than those of the Welch, modified-Welch and inverse functions, according to [13, 16, 12], but they seem much better than what gives this observation for $r = 2$.

## 5 A bound for the whole nonlinearity profile of the inverse function

Thanks to Proposition 6 and to Relation (2), we can now apply Corollary 1 with $r = 3$, $K = \sqrt{(2^n-1)(2^{n/2+2}+4)+2^n}+2$ and $k = 0$, we deduce:

**Proposition 7** *Let $f_{inv}(x) = tr(x^{2^n-2})$. Then we have:*

$$nl_3(f_{inv}) \geq 2^{n-1} - \frac{1}{2}\sqrt{(2^n-1)\left(2\sqrt{(2^n-1)(2^{n/2+2}+4)+2^n}+4\right)+2^n} \approx 2^{n-1} - 2^{7n/8}.$$

Unfortunately, we cannot produce a table to compare this bound and the actual values, as we did for the second order, because for small values of $n$ (precisely, for $n \leq 8$), the bound gives negative numbers, and for greater values, the algorithm is unable to produce results.

The same process can be iteratively applied, giving a lower bound on the $r$-th order nonlinearity of the inverse functions for $r \geq 4$. The expression of this lower bound is:

$$nl_r(f_{inv}) \geq 2^{n-1} - l_r,$$

where, according to Relation (2) and to Corollary 1, the sequence $l_r$ is defined by $l_1 = 2^{n/2}$ and $l_r = \sqrt{(2^n-1)(l_{r-1}+1)+2^{n-2}}$. The expression of $l_r$ is more and more complex when $r$ increases. Its value is approximately equal to $k_r$, where $k_1 = n/2$ and $k_r = \frac{n+k_{r-1}}{2}$, and therefore $k_r = (1-2^{-r})\,n$. Hence, $nl_r(f_{inv})$ is approximately lower bounded by $2^{n-1} - 2^{(1-2^{-r})\,n}$ and asymptotically equivalent to $2^{n-1}$.

**Acknowledgement**

This work was initiated while the author was visiting Professor Harald Niederreiter in Singapore. The author wishes to take this opportunity for thanking him for very stimulating discussions on several subjects, including the nonlinearity profile.

# References

[1] A. Canteaut, P. Charpin, and H. Dobbertin. Binary $m$-sequences with three-valued crosscorrelation: A proof of Welch's conjecture. *IEEE Trans. Inform. Theory*, 46 (1), pp. 4-8, 2000.

[2] C. Carlet. Boolean Functions for Cryptography and Error Correcting Codes. Chapter of the monography *Boolean Methods and Models*, Y. Crama and P. Hammer eds, Cambridge University Press, to appear. Preliminary version available at http://www-rocq.inria.fr/codes/Claude.Carlet/pubs.html

[3] C. Carlet. Vectorial (multi-output) Boolean Functions for Cryptography. Chapter of the monography *Boolean Methods and Models*, Y. Crama and P. Hammer eds, Cambridge University Press, to appear. Preliminary version available at http://www-rocq.inria.fr/codes/Claude.Carlet/pubs.html

[4] C. Carlet. The complexity of Boolean functions from cryptographic viewpoint. Dagstuhl Seminar "Complexity of Boolean Functions", 2006. Paper available at URL http://drops.dagstuhl.de/portals/06111/

[5] C. Carlet. On the higher order nonlinearities of algebraic immune functions. CRYPTO 2006. Lecture Notes in Computer Science 4117, pp. 584-601, 2006.

[6] C. Carlet, D. Dalai, K. Gupta and S. Maitra. Algebraic Immunity for Cryptographically Significant Boolean Functions: Analysis and Construction. IEEE Transactions on Information Theory, vol. 52, no. 7, pp. 3105-3121, July 2006.

[7] C. Carlet and S. Mesnager. Improving the upper bounds on the covering radii of binary Reed-Muller codes. To appear in IEEE Transactions on Information Theory, 2006.

[8] P. Charpin, T. Helleseth and V. Zinoviev. Propagation characteristics of $x \to x^{-1}$ and Kloosterman sums. To appear in Finite Fields and their Applications.

[9] G. Cohen, I. Honkala, S. Litsyn and A. Lobstein. *Covering codes.* North-Holland, 1997.

[10] N. Courtois. Higher order correlation attacks, XL algorithm and cryptanalysis of Toyocrypt. Proceedings of ICISC 2002, LNCS 2587, pp. 182-199.

[11] H. Dobbertin, Almost perfect nonlinear power functions on $GF(2^n)$: The Welch case, *IEEE Trans. Inform. Theory* 45, pp. 1271-1275, 1999.

[12] I. Dumer, G. Kabatiansky and C. Tavernier. List decoding of Reed-Muller codes up to the Johnson bound with almost linear complexity. Proceedings of ISIT 2006. Seattle, USA.

[13] R. Fourquet. Une FFT adaptée au décodage par liste dans les codes de Reed-Muller d'ordres 1 et 2. Master-thesis of the University of Paris VIII, Thales communication, Bois Colombes, 2006.

[14] J. Golic. Fast low order approximation of cryptographic functions. Proceedings of EUROCRYPT'96, LNCS 1070, pp. 268-282, 1996.

[15] T. Iwata and K. Kurosawa. Probabilistic higher order differential attack and higher order bent functions. Proceedings of ASIACRYPT'99, LNCS 1716, pp. 62-74, 1999.

[16] G. Kabatiansky and C. Tavernier. List decoding of second order Reed-Muller codes. In Proc. $8^{th}$ Intern. Simp. Comm. Theory and Applications. Ambelside, UK, july 2005.

[17] L.R. Knudsen and M. J. B. Robshaw. Non-linear approximations in linear cryptanalysis. Proceedings of EUROCRYPT'96, LNCS 1070, pp. 224-236, 1996.

[18] G. Lachaud and J. Wolfmann. The Weights of the Orthogonals of the Extended Quadratic Binary Goppa Codes. *IEEE Trans. Inform. Theory*, vol. 36, pp. 686-692, 1990.

[19] R. Lidl and H. Niederreiter, *Finite Fields*, Encyclopedia of Mathematics and its Applications, vol. 20, Addison-Wesley, Reading, Massachussetts (1983)

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

[21] U. M. Maurer. New approaches to the design of self-synchronizing stream ciphers. Proceedings of EUROCRYPT'91. LNCS 547, pp. 458-471, 1991.

[22] W. Millan. Low order approximation of cipher functions. Cryptographic Policy and Algorithms. LNCS 1029, pp. 144-155, 1996.

# 6 Appendix

In the next proposition, we generalize Proposition 2 and its iteration.

**Proposition 8** *Let $f$ be an $n$-variable function and $r$ a positive integer smaller than $n$. Let $E$ and $F$ be two subspaces of $F_2^n$ whose direct sum equals $F_2^n$. We denote the dimension of $E$ by $m$. For every $u \in E$, let us denote by $f_u$ the restriction of $f$ to the flat $u + F$ ($f_u$ can be viewed as an $(n-m)$-variable function). Let $k$ be a positive integer such that $k \leq m$ and let $\{A_1, \cdots, A_{2^{m-k}}\}$ be a partition of $E$ by $k$-dimensional flats. For every $i = 1, \cdots, 2^{m-k}$, let us denote by $f_i$ the function $\sum_{u \in A_i} f_u$. Then, for every $r \geq k$, we have:*

$$nl_r(f) \geq \sum_{i=1}^{2^{m-k}} nl_{r-k}(f_i).$$

*Proof*: Let $h$ be a function of algebraic degree at most $r$ such that $nl_r(f) = w_H(f+h)$. We have $w_H(f+h) = \sum_{u \in E} w_H(f_u + h_u) = \sum_{i=1}^{2^{m-k}} \left(\sum_{u \in A_i} w_H(f_u + h_u)\right)$. For every $i = 1, \cdots, 2^{m-k}$, we have $\sum_{u \in A_i} w_H(f_u + h_u) \geq w_H(f_i + h_i) = d_H(f_i, h_i)$. It is a simple matter to prove by induction on $k$ that, $h$ having algebraic degree at most $r$, function $h_i$ has algebraic degree at most $r-k$. This completes the proof. $\square$

The bound $nl_r(f) \geq \frac{1}{2} nl_{r-1}(D_a f)$, $a \neq 0$, corresponds in Proposition 8 to the case where $E = \{0, a\}$ (and $F$ is any vector space not containing $a$) and $k = 1$. The bound $nl_r(f) \geq \frac{1}{2^i} nl_{r-i}(D_{a_1} \cdots D_{a_i} f)$, $a \neq 0$, corresponds to the case where $E$ is an $i$-th dimensional vector space and $k = i$.
Note that for $k = r$, Proposition 8 gives

$$nl_r(f) \geq \sum_{i=1}^{2^{m-r}} \min(w_H(f_i), w_H(f_i + 1)).$$

If every function $f_i$ is balanced then this gives $nl_r(f) \geq 2^{m-r} \cdot 2^{n-m-1} = 2^{n-r-1}$. This is the same bound as for a function of algebraic degree $r+1$, while we do not have to assume here that $f$ has algebraic degree $r+1$. But we see clearly the limitation of this approach since we do not get a bound which is equivalent to $2^{n-1}$.

**Remark on Proposition 3**: we could try considering similar sums as in the proof of this proposition, but with exponents greater than 2. However, this leads to considering sets of the form $\{x_1, \cdots, x_\ell\}$ instead of $\{x, y\}$ and such set being not a flat when $\ell > 2$, we would not be led to derivatives of $h$; the algebraic degree would then remain $r$ instead of being reduced. A possible solution to this problem is to sum up sums of this form corresponding to $h$ ranging over a coset of the Reed-Muller code of order $\ell - 1$. For every $n$-variable function $h$ of algebraic degree at most $r$, we have

$$\sum_{g \in RM(\ell-1,n)} \left( \sum_{x \in F_2^n} (-1)^{f(x)+h(x)+g(x)} \right)^{2^\ell} =$$

$$\sum_{x_1, \cdots, x_{2^\ell} \in F_2^n} (-1)^{\sum_{i=1}^{2^\ell} (f+h)(x_i)} \left( \sum_{g \in RM(\ell-1,n)} (-1)^{\sum_{i=1}^{2^\ell} g(x_i)} \right).$$

The sum $\sum_{g \in RM(\ell-1,n)} (-1)^{\sum_{i=1}^{2^\ell} g(x_i)}$ is the character sum over a vector space of a linear form; hence, it is null for every tuple $(x_1, \cdots x_{2^\ell})$ such that the expression $\sum_{i=1}^{2^\ell} g(x_i)$ is not identically null. Since $\sum_{i=1}^{2^\ell} g(x_i)$ equals the inner product between $g \in RM(\ell-1,n)$ and the function equal to $1_{x_1} + \cdots + 1_{x_{2^\ell}}$, where $1_{x_i}$ denotes the indicator of the singleton $\{x_i\}$, this means that the sum is nonzero if and only if the function $1_{x_1} + \cdots + 1_{x_{2^\ell}}$ belongs to the dual of $RM(\ell-1,n)$, that is, $RM(n-\ell,n)$. We know (see [20, 2]) that the elements of weights at most $2^\ell$ of the code $RM(n-\ell,n)$ are the null element and the indicators of $\ell$-dimensional flats. Denoting by $\mathcal{E}_\ell$ the set of those $2^\ell$-tuples $(x_1, \cdots, x_{2^\ell})$ of vectors $x_i$ of $F_2^n$ which are pairwise equal to each others and by $\mathcal{A}_\ell$ the set of those $2^\ell$-tuples $(x_1, \cdots, x_{2^\ell})$ of vectors $x_i$ of $F_2^n$ such that $\{x_1, \cdots, x_{2^\ell}\}$ is an $\ell$-dimensional flat, we deduce (since we know that the maximum is achieved at least by two functions $g$):

$$\max_{g \in RM(\ell-1,n)} \left( \sum_{x \in F_2^n} (-1)^{f(x)+h(x)+g(x)} \right)^{2^\ell} \leq$$

$$\frac{1}{2} \sum_{g \in RM(\ell-1,n)} \left( \sum_{x \in F_2^n} (-1)^{f(x)+h(x)+g(x)} \right)^{2^\ell} =$$

$$\frac{1}{2}|RM(\ell-1,n)| \left( \sum_{(x_1,\cdots,x_{2^\ell}) \in \mathcal{E}_\ell} (-1)^{\sum_{i=1}^{2^\ell} (f+h)(x_i)} + \sum_{(x_1,\cdots,x_{2^\ell}) \in \mathcal{A}_\ell} (-1)^{\sum_{i=1}^{2^\ell} (f+h)(x_i)} \right) =$$

$$\frac{1}{2}|RM(\ell-1,n)| \left( |\mathcal{E}_\ell| + \sum_{\substack{a_1,\cdots,a_\ell \in F_2^n \\ a_1,\cdots,a_\ell \text{ linearly independent}}} \sum_{x \in F_2^n} (-1)^{D_{a_1} \cdots D_{a_\ell} f(x) + D_{a_1} \cdots D_{a_\ell} h(x)} \right).$$

Note that $D_{a_1} \cdots D_{a_\ell} h$ has algebraic degree at most $r - \ell$. Then we deduce

$$nl_r(f) \geq 2^{n-1} - \frac{1}{2} \sqrt[2^\ell]{\Delta},$$

where $\Delta$ equals

$$\frac{1}{2}|RM(\ell-1,n)|\left(|\mathcal{E}_\ell|+2^n(2^n-1)\cdots(2^n-2^{\ell-1})-2\sum_{a_1,\cdots,a_\ell\in F_2^n} nl_{r-\ell}(D_{a_1}\cdots D_{a_\ell}f)\right).$$

But this bound is interesting only when $\sum_{a_1,\cdots,a_\ell\in F_2^n} nl_{r-\ell}(D_{a_1}\cdots D_{a_\ell}f)$ has very small value. Similarly to Corollary 1, we deduce that if, for some non-negative integers $K$ and $k$, we have $nl_{r-2}(D_{a_1}\cdots D_{a_\ell}f)\geq 2^{n-1}-K\,2^k$ for every linearly independent $a_1,\cdots,a_\ell\in F_2^n$, then

$$nl_r(f) \quad\geq\quad 2^{n-1}-\frac{1}{2}\sqrt[2^\ell]{\Lambda}$$

where $\Lambda=\frac{1}{2}|RM(\ell-1,n)|\left(|\mathcal{E}_\ell|+K\,2^{k+1}(2^n-1)(2^n-2)\cdots(2^n-\ell)\right)$. But this can be useful only if $K\,2^k$ is very small. $\square$