On the higher order nonlinearities of algebraic immune functions

Claude Carlet

INRIA Projet CODES, BP 105, 78153 Le Chesnay Cedex, France;
e-mail: claude.carlet@inria.fr; also member of the University of Paris 8 (MAATICAH).

Abstract. One of the most basic requirements concerning Boolean functions used in cryptosystems is that they must have high algebraic degrees. This simple criterion is not always well adapted to the concrete situation in which Boolean functions are used in symmetric cryptography, since changing one or several output bits of a Boolean function considerably changes its algebraic degree while it may not change its robustness. The proper characteristic is the r-th order nonlinearity profile (which includes the first-order nonlinearity). However, studying it is difficult and almost no paper, in the literature, has ever been able to give general effective results on it. The values of the nonlinearity profile are known for very few functions and these functions have little cryptographic interest. A recent paper has given a lower bound on the nonlinearity profile of functions, given their algebraic immunity. We improve upon it, and we deduce that it is enough, for a Boolean function, to have high algebraic immunity, for having non-weak low order nonlinearity profile (even when it cannot be evaluated), except maybe for the first order.

Keywords: stream cipher, block cipher, algebraic attack, Boolean function, algebraic immunity, algebraic degree, higher order nonlinearity.

1 Introduction

Boolean functions, that is, F_2-valued functions defined over the vector space F_2^n of all binary vectors of a given length n, are used in the S-boxes of block ciphers and in the pseudo-random generators of stream ciphers. They play a central role in their security.

In stream ciphers, the main model for the generation of the keystream consists of a linear part, producing a sequence with a large period, usually composed of one or several LFSRs, and a nonlinear combining or filtering function f which produces the output, given the state of the linear part. In the nonlinear combiner sub-model, the outputs to several LFSRs are combined using a nonlinear Boolean function to produce the keystream. In the nonlinear filter sub-model, the content of some of the flip-flops in a single (longer) LFSR constitute the input to a nonlinear Boolean function which produces the keystream. These models which are very efficient, in particular in hardware, have undergone a lot of cryptanalysis and to resist those attacks, different design criteria have been proposed for both the LFSRs and the combining Boolean function. The main classical cryptographic criteria for designing the function f are balancedness (f is balanced if its Hamming weight equals 2^{n-1}) to prevent the system from leaking statistical information on the plaintext when the ciphertext is known, a high algebraic degree (that is, a high degree of the algebraic normal form of the function) to prevent the system from Massey's attack by the Berlekamp-Massey algorithm (cf. [31, 39], see also [34]), the non-existence of (non-zero) linear structure $a \in F_2^n$ (such that $f(x+a)+f(x)$ is constant) so that the function effectively depends on all its variables, a high order of correlation immunity (and more precisely, of resiliency, since the functions must be balanced - a function is t-resilient if each of its restrictions obtained by keeping constant t input bits is balanced) to counter correlation attacks (in the case of combining functions), and a high nonlinearity (that is, a large Hamming distance to affine functions) to withstand correlation attacks (again) and linear attacks.

The recent algebraic attacks [15] have led to further characteristics that a cryptographic Boolean function must have. These attacks cleverly use over-defined systems of multivariate nonlinear equations to recover the secret key (the idea of using such systems comes from C. Shannon [40], but the improvement in the efficiency of the method is recent). The core of the analysis in the standard algebraic attack is to find out low degree functions $g \neq 0$ and h such that $fg = h$. It has been shown in [37] that this is equivalent to the existence of a low degree nonzero annihilator of f or of $1+f$, that is, of a function g such that $fg = 0$ (i.e. whose support is disjoint of that of f) or $(1+f)g = 0$. The minimum degree of such g is called the (basic) algebraic immunity of f and must be as high as possible (the maximum being $\lceil \frac{n}{2} \rceil$). This condition is not sufficient, since a function can have sufficiently high algebraic immunity and be weak against fast algebraic attacks [16]. If one can find g of low degree and $h \neq 0$ such that $fg = h$, then a fast algebraic attack is feasible if the degree of h is not too high, see [16, 1, 24]. Since $fg = h$ implies $fh = ffg = fg = h$, we see that h is then an annihilator of $f+1$ and its degree is then at least equal to the algebraic immunity of f. This means that having a high algebraic immunity is not only a necessary condition for a resistance to standard algebraic attacks but also for a resistance to fast algebraic attacks.

Some of the criteria above play also important roles for S-boxes in block ciphers: the nonlinearity (cf. the linear attack by Matsui [32], see also [12]) and the algebraic degree (the complexity of the "higher order differential attack" on block ciphers due to Knudsen and Lai [26, 28] depends on the algebraic degrees of the Boolean functions involved in the system).

But these criteria must be considered in an extended way: suppose that, given a function which does not satisfy some criterion, it is possible, by changing one or a few bits in its output (that is, in its truth-table) to obtain a function which satisfies the criterion; then this criterion cannot have a general relevance to cryptography, since this change does not fundamentally modify the robustness of the system using this function (however, this situation is not quite the same according to whether the function is used in a synchronous stream cipher, a self-synchronizing stream cipher or a block cipher). Some papers in the literature have already addressed this problem for some criteria: see [38] for the criterion of non-existence of nonzero linear structure (Meier and Staffelbach considered the so-called distance to linear structures) and [29, 10] for the resiliency criterion. Some other criteria – for instance the nonlinearity – do not change much when a few bits of the output to the function are changed; hence, they do not need such extension. On the contrary, changing one single bit in the output to an n-variable function of algebraic degree at most $n-1$ moves its degree to n, and if, starting from a balanced function we want to keep balancedness, changing two bits moves it almost surely to $n-1$. A natural way of putting this right is to do as Meier and Staffelbach did for the linear structures, considering, for every $r < n$, the minimum Hamming distance to all functions of degrees at most r (whose set is the so-called r-th order Reed-Muller code and will be denoted by $RM(r)$). This distance is usually called the r-th order nonlinearity of the function (and more simply its nonlinearity in the first-order case).

We shall call the *nonlinearity profile* of a function the sequence whose r-th term, for $r = 1, \ldots, n-1$, equals the r-th order nonlinearity of the function. Note that the nonlinearity profile is extended-affine-invariant, in the sense that if ϕ is an affine automorphism of F_2^n and if ℓ is an affine Boolean function on F_2^n, then for every n-variable Boolean function f, the nonlinearity profile of the function $f \circ \phi + \ell$ equals that of f. The best known upper bound on the r-th order nonlinearity of general functions has been given in [11]. Several papers have shown the role played by this parameter in relation to some cryptanalyses (note that, contrary to the first order nonlinearity, it must have low value for allowing the attacks to be realistic) and studied it from an algorithmic viewpoint [14, 23, 25, 27, 33, 35]. However, very few have attempted to give constructions of functions with reasonably good nonlinearity profile or to show general properties of this parameter [25]. In fact, until recently, almost nothing relevant was known on this criterion (see Section 2). Fortunately, it has been shown lately that, if the algebraic immunity $AI(f)$ of a function f is known, then we can deduce a lower bound on its r-th order nonlinearity for every $r \leq AI(f) - 1$: for the first order, see [17] and the improvement of [30]; for any order, see [8] (see Section 2 for a recall of these bounds). This changes completely the situation with the nonlinearity profile. In this paper, we obtain a new bound which improves upon the bound of [8] for all values of $AI(f)$ when the number of variables is smaller than or equal to 12, and for most values of $AI(f)$ when the number of variables is smaller than or equal to 22 (which covers the practical situation of stream ciphers). It also improves asymptotically upon it.

The paper is organized as follows. In Section 2 are given the necessary definitions and properties of the main cryptographic criteria on Boolean functions. In Section 3, we study the dimension of the vector space of annihilators with prescribed algebraic degrees, of a Boolean function with given algebraic degree. The results of this section are used in Section 4 to obtain the lower bound on the r-th order nonlinearity of a function of given algebraic immunity. Finally, in Section 5, we show that the bound of Section 4 simplifies the question of designing cryptographic Boolean functions meeting all necessary criteria.

2 Preliminaries

Let n be any positive integer. Any Boolean function f on n variables admits a unique 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 algebraic normal form. *Affine functions* are those Boolean functions of degrees at most 1.

A slightly different form for the ANF is $f(x) = \sum_{u \in F_2^n} a_u x^u$, where $a_u \in F_2$ and where $x^u = \prod_{i=1}^n x_i^{u_i}$. Then $d^\circ f$ equals $\max_{a_u \neq 0} wt(u)$, where $wt(u)$ denotes the Hamming weight $|\{i = 1, \ldots, n \,/\, u_i = 1\}|$ of u. Note that, for every $x \in F_2^n$, we have then $f(x) = \sum_{u \preceq x} a_u$, where $u \preceq x$ means that every coordinate of u is upper bounded by the corresponding coordinate of x, that is, that the support of u is included in the support of x.

The Hamming weight of a Boolean function is the Hamming weight of its list of values, that is, the size of its support $\{x \in F_2^n \,/\, f(x) = 1\}$. The Hamming distance between two Boolean functions is the Hamming weight of $f + g$, that is $d(f, g) = |\{x \in F_2^n \,/\, f(x) \neq g(x)\}|$.

Definition 1. *Let $f : F_2^n \to F_2$ be an n-variable Boolean function. Let r be a positive integer such that $r \leq n$. The* r-th order nonlinearity *of f is the minimum Hamming distance between f and all n-variable functions of algebraic degrees at most r.*

We shall denote the r-th order nonlinearity of f by $nl_r(f)$. The first-order nonlinearity of f is simply called the nonlinearity of f and denoted by $nl(f)$.

Clearly we have $nl_r(f) = 0$ if and only if f has degree at most r. So, the knowledge of the nonlinearity profile (i.e. of all the nonlinearities of orders $r \geq 1$) of a Boolean function includes the knowledge of its algebraic degree. It is in fact a much more complete cryptographic parameter than are the single (first-order) nonlinearity and the algebraic degree: the former is not sufficient for knowing the cryptographic behavior of a function (it does not allow to quantify for instance the resistance to Berlekamp-Massey attack) and the latter is still less sufficient, as explained in introduction.

As far as we know, the nonlinearity profile (or a great part of it) is known in general only for quadratic functions (the functions of algebraic degrees at most 2) and for their sums with functions of very small Hamming weights. Indeed, the first-order nonlinearities of quadratic functions are known (see [36]) and the r-th order nonlinearity of a quadratic function is obviously null for every $r \geq 2$. The first-order nonlinearities of the functions of degrees greater than or equal to 3 are unknown, except for some particular primary constructions of Boolean functions (such as the indicators of flats, or some concatenations of such indicators or of quadratic functions - including the Maiorana-McFarland functions) and for some secondary constructions; nothing is known on the second-order nonlinearities of functions of degrees at least 3 (except for functions of small weights). In the case of functions of small Hamming weights (e.g. the indicators of flats of small dimensions), the r-th order nonlinearity is equal, for sufficiently low values of r (namely, for $2^{n-r} > 2wt(f)$), to the weight $wt(f)$ itself.

The algebraic immunity [37] of a Boolean function f quantifies the resistance to the standard algebraic attack of the pseudo-random generators using it as a nonlinear function.

Definition 2. *Let $f : F_2^n \rightarrow F_2$ be an n-variable Boolean function. We call annihilator of f any n-variable function g whose product with f is null (i.e. whose support is included in the support of $f+1$, or in other words any function which is a multiple of $f+1$). The algebraic immunity of f is the minimum algebraic degree of all the nonzero annihilators of f or of $f+1$.*

We shall denote the algebraic immunity of f by $AI(f)$.
A very useful property is its affine invariance: for every affine automorphism ϕ of F_2^n, we have $AI(f \circ \phi) = AI(f)$. This comes from the affine invariance of the algebraic degree.

Clearly, since f is an annihilator of $f+1$ (and $f+1$ is an annihilator of f) we have $AI(f) \leq d^\circ f$.
As shown in [15], we always have $AI(f) \leq \lceil \frac{n}{2} \rceil$. This bound is tight (see below). Also, we know that almost all Boolean functions have algebraic immunities close to this optimum; more precisely, for all $a < 1$, AI(f) is almost surely greater than $\frac{n}{2} - \sqrt{\frac{n}{2} \ln\left(\frac{n}{a \ln 2}\right)}$ when n tends to infinity: see [21].

Even when restricting ourselves to functions with optimum algebraic immunity, the algebraic attacks oblige to use now functions on at least 13 variables, see [4, 8] (this number is a strict minimum and is in fact risky; a safer number of variables would better be near 20).

Very few functions are known (up to affine equivalence) with provably optimum algebraic immunities: the functions whose construction is introduced in [18] (see in [8] their further properties) and some functions which are symmetric (that is, whose outputs depend only on the Hamming weights of their inputs) [19, 3]. These functions have some drawbacks: all of them have insufficient nonlinearities and all but one are non-balanced. Moreover, the functions studied in [19, 3], and to a slightly smaller extent the functions introduced in [18], have not a good behavior against fast algebraic attacks, see [2, 20]. But the research in this domain is very active and it is probable that better examples of functions will be found in the future.

It was shown in [17] that the weight of a function f with given algebraic immunity satisfies: $\sum_{i=0}^{AI(f)-1} \binom{n}{i} \leq wt(f) \leq \sum_{i=0}^{n-AI(f)} \binom{n}{i}$. In particular, if n is odd and f has optimum algebraic immunity, then f is balanced.

The first lower bound on the (first-order) nonlinearity of functions with given algebraic immunity has been obtained in [17]: $nl(f) \geq \sum_{i=0}^{AI(f)-2} \binom{n}{i}$. In [30], M. Lobanov has improved upon this lower bound: $nl(f) \geq 2\sum_{i=0}^{AI(f)-2} \binom{n-1}{i}$. In [8], an easy generalization to the r-th order nonlinearity of the bound obtained in [17] has been given: $nl_r(f) \geq \sum_{i=0}^{AI(f)-r-1} \binom{n}{i}$. In the present paper, we extend Lobanov's bound into a bound which improves, asymptotically and in most cases of practical situations, upon the bound obtained in [8]. This bound is related to the dimension of the annihilators with prescribed algebraic degrees of Boolean functions with given algebraic degrees.

3 The dimension of the vector space of prescribed degree annihilators of a function

The number of linearly independent low degree annihilators of a given Boolean function f and of the function $f+1$ is an important parameter for evaluating the complexity of algebraic attacks on the systems using this function. We shall see in the next section that it plays also an important role in relation to the r-th order nonlinearity.

Definition 3. *Let h be an n-variable Boolean function. We denote by $An_k(h)$ the vector space of those annihilators of degrees at most k of h and by $d_{k,h}$ the dimension of $An_k(h)$.*

Little is known on the behavior of $d_{k,h}$. For $k = n$, we have clearly $d_{n,h} = 2^n - wt(h)$ since $An_n(h)$ contains all functions whose supports are disjoint of that of h. It is also shown in [17, 8] that:
- for $k = AI(h)$, we have $d_{k,h} \leq \binom{n}{k}$,
- if h is balanced and has algebraic immunity $\frac{n}{2}$ (n even), then $d_{\frac{n}{2},h} \geq \frac{1}{2} \cdot \binom{n}{\frac{n}{2}}$,
- if h has algebraic immunity $\frac{n+1}{2}$ (n odd), then $d_{\frac{n+1}{2},h} = \binom{n}{\frac{n+1}{2}}$.
Also, Lobanov [30] showed that for every non-constant affine function h and every k, we have $d_{k,h} = \sum_{i=0}^{k-1} \binom{n-1}{i}$.

Before introducing an upper bound on $d_{k,h}$ which is valid for all functions, we generalize Lobanov's result by determining the values of $d_{k,h}$ for several classes of functions. This will be useful in the sequel.

Proposition 1. *Let h be any n-variable function of degree r, such that $0 \leq r \leq n$, and of weight 2^{n-r}. Then for every $k \geq 0$ we have $d_{k,h} = \sum_{i=0}^{k} \binom{n}{i} - \sum_{i=0}^{k} \binom{n-r}{i}$.*

Proof:
We know that h is the indicator of an $(n-r)$-dimensional flat (see e.g. [36]), and thanks to the affine invariance of the algebraic immunity, we may without loss of generality assume that it equals $(x_1+1)(x_2+1)\cdots(x_r+1)$. The system characterizing the elements of $An_k(h)$, that is, the system of all equations $\sum_{u \preceq x \mid wt(u) \leq k} a_u = 0$ where x ranges over the support $\{(0,\ldots,0)\} \times F_2^{n-r}$ of h,

does not involve any unknown a_u such that $(u_1,\dots,u_r)\neq(0,\dots,0)$. And when considering it as a system with unknowns a_u such that $(u_1,\dots,u_r)=(0,\dots,0)$, it is the system obtained when characterizing the $(n-r)$-variable annihilators of degrees at most k of the constant function 1. This last system has rank $\sum_{i=0}^{k}\binom{n-r}{i}$, since the function 1 admits the null function as only annihilator, and this implies that $d_{k,h}=\sum_{i=0}^{k}\binom{n}{i}-\sum_{i=0}^{k}\binom{n-r}{i}$. Note that, in the case $r=1$, this is the value given by Lobanov for non-constant affine functions. $\square$

Proposition 2. *Let h be any n-variable function of degree r, such that $0\leq r\leq n$, and of weight 2^n-2^{n-r}. Then for every $k\geq 0$ we have $d_{k,h}=\sum_{i=0}^{k-r}\binom{n-r}{i}$.*

Proof:
$h+1$ is the indicator of an $(n-r)$-dimensional flat, and we may without loss of generality assume that it equals $x_1x_2\cdots x_r$. Then the elements of $An_k(h)$ are the products of $x_1x_2\cdots x_r$ with those functions in the variables $x_{r+1},\dots,x_n$ whose degrees are at most $k-r$. Then $d_{k,h}=\sum_{i=0}^{k-r}\binom{n-r}{i}$. In the case $r=1$, this is also the value given by Lobanov. $\square$

Proposition 3. *Let t be an integer such that $t\leq n$ and let h be the symmetric function defined by $h(x)=1$ if and only if $wt(x)<t$. Then, for every k, we have $d_{k,h}=\sum_{i=t}^{k}\binom{n}{i}$.*

Proof:
The coefficients in the ANFs of the elements of $An_k(h)$ are the solutions of the system of equations $\displaystyle\sum_{u\preceq x\,|\,wt(u)\leq k} a_u=0$, where x ranges over the set of vectors of weights strictly smaller than t. If $k\geq t-1$, then these equations become $\sum_{u\preceq x}a_u=0$ and these ANFs are the polynomials such that $a_u=0$ if $wt(u)<t$ (and a_u is any element of F_2 if $t\leq wt(u)\leq k$). Otherwise, it is clear that $An_k(h)=\{0\}$. $\square$
If $t\leq\lceil\frac{n}{2}\rceil$, we have then $AI(h)=t$, since it is easy to show then that $An_k(h+1)$ does not contain nonzero functions of degrees strictly smaller than t.
Note that, denoting by f the majority function (i.e. the symmetric function of support $\{x\in F_2^n\,/\,wt(x)\geq\lceil\frac{n}{2}\rceil\}$), Proposition 3 with $t=\lceil\frac{n}{2}\rceil$ (resp. with $t=\lfloor\frac{n}{2}\rfloor+1$) gives the value of $d_{k,f+1}$ (resp. of $d_{k,f}$, thanks to affine invariance).

The functions studied in Propositions 1 and 2 are balanced when $r=1$ only, and those studied in Proposition 3 are balanced when $t=\frac{n+1}{2}$ (n odd). We study in the next proposition a more general case of balanced functions.

Proposition 4. *Let h be an n-variable function of weight 2^{n-r} ($1<r\leq n-1$) and ℓ a non-constant affine function such that $h+\ell$ is balanced. Then*

$$d_{k,h+\ell}=\sum_{i=0}^{k-1}\binom{n-1}{i}-\sum_{i=k-r+1}^{k-1}\binom{n-r-1}{i}+\sum_{i=0}^{k-r-1}\binom{n-r-1}{i}.$$

Proof:
We may without loss of generality assume that $h(x)$ equals $(x_1+1)(x_2+1)\cdots(x_r+1)$ and that $l(x)$ is the function x_{r+1}. The annihilators of $h+\ell$ are then the multiples of $(x_1+1)(x_2+1)\cdots(x_r+1)+(x_{r+1}+1)$.
Let $\left((x_1+1)(x_2+1)\cdots(x_r+1)+(x_{r+1}+1)\right)\left(\sum_{u\in F_2^n} a_u x^u\right)$ be such a multiple.
1. For every u such that $u_{r+1}=1$ and for which there exists $i\le r$ such that $u_i=1$, we have $\left((x_1+1)(x_2+1)\cdots(x_r+1)+(x_{r+1}+1)\right)x^u=0$. Hence, we must not take this case into account when quantifying the dimension.
2. For every u such that $u_{r+1}=0$ and for which there exists $i\le r$ such that $u_i=1$, the corresponding multiple $\left((x_1+1)(x_2+1)\cdots(x_r+1)+(x_{r+1}+1)\right)x^u$ equals $(x_{r+1}+1)\prod_{i=1}^n x_i^{u_i}$. Its degree equals $1+wt(u_1,\dots,u_n)$.
3. For every u such that $u_1=\cdots=u_r=0$ and $u_{r+1}=1$, the corresponding multiple equals $\left((x_1+1)(x_2+1)\cdots(x_r+1)\right)\prod_{i=r+1}^n x_i^{u_i}$. Its degree equals $r+1+wt(u_{r+2},\dots,u_n)$.
4. For every u such that $u_1=\cdots=u_r=u_{r+1}=0$, the corresponding multiple equals $\left((x_1+1)(x_2+1)\cdots(x_r+1)+(x_{r+1}+1)\right)\prod_{i=r+2}^n x_i^{u_i}$. Its degree equals $r+wt(u_{r+2},\dots,u_n)$.
The functions of cases 2, 3 and 4 are linearly independent. Then $d_{k,h+\ell}$ equals $\sum_{i=0}^{k-1}\binom{n-1}{i}-\sum_{i=0}^{k-1}\binom{n-r-1}{i}+\sum_{i=0}^{k-r-1}\binom{n-r-1}{i}+\sum_{i=0}^{k-r}\binom{n-r-1}{i}$. $\square$

Remark: The knowledge of $d_{k,h}$ for some function h gives information on $d_{k,h'}$ for some other functions h':
1. For every n-variable functions h and h' and every positive integer k, we have $|d_{k,h}-d_{k,h'}|\le\max(wt(h(h'+1)),wt((h+1)h'))\le wt(h+h')$, since the ranks of the systems characterizing $An_k(h)$ and $An_k(h')$ satisfy the same inequality (indeed, adding equations to a system increases its rank by at most the number of added equations, and the system characterizing the ANFs of the annihilators of h' can be obtained from the system characterizing the ANFs of those of h by adding $wt(h'(h+1))$ equations and suppressing $wt(h(h'+1))$ equations).
2. Let h be an n-variable function and let h' be the $(n+1)$-variable function $h'(x_1,\dots,x_{n+1})=h(x_1,\dots,x_n)$. Let k be an integer. The ANFs of the elements of $An_k(h)$ are the solutions of the system of equations $\sum_{u\in F_2^n\,|\,wt(u)\le k} a_u x^u=0$, where x ranges over $supp(h)$ and the ANFs of the elements of $An_k(h')$ are the solutions of the system of equations $\sum_{v\in F_2^{n+1}\,|\,wt(v)\le k} b_v y^v=0$, where y ranges over $supp(h')=supp(h)\times F_2$. This last equation is equal to $\sum_{u\in F_2^n\,|\,wt(u)\le k} b_{u,0}x^u=0$ if $y=(x,0)$ and to $\sum_{u\in F_2^n\,|\,wt(u)\le k} b_{u,0}x^u+\sum_{u\in F_2^n\,|\,wt(u)\le k-1} b_{u,1}x^u=0$ if $y=(x,1)$. Hence, $d_{k,h'}=d_{k,h}+d_{k-1,h}$.

In the next lemma, we extend to all Boolean functions the result from [30], recalled above, which dealt only with affine functions.

Lemma 1. *Let n be a positive integer. Let r and k be positive integers smaller than or equal to n. Let h be any n-variable Boolean function of algebraic degree r. Then*

$$d_{k,h} \le \min\left(\sum_{i=AI(h)}^{k}\binom{n}{i}, \sum_{i=0}^{k}\binom{n}{i} - \sum_{i=0}^{k}\binom{n-r}{i}\right).$$

Proof:
We first prove that $d_{k,h} \le \sum_{i=AI(h)}^{k}\binom{n}{i}$. If two elements of $An_k(h)$ have the same degree k part $\sum_{u\in F_2^n \mid wt(u)=k} x^u$, then their sum belongs to $An_{k-1}(h)$. We deduce that $d_{k,h}$ is smaller than or equal to the sum of $d_{k-1,h}$ and of the dimension of $RM(k)/RM(k-1)$, where $RM(k)$ is the Reed-Muller code of order k. This proves $d_{k,h} \le d_{k-1,h} + \binom{n}{k}$ and we deduce the relation $d_{k,h} \le \sum_{i=AI(h)}^{k}\binom{n}{i}$ by induction on k, since, by definition, $d_{AI(h)-1,h} = 0$.
We prove now that $d_{k,h} \le \sum_{i=0}^{k}\binom{n}{i} - \sum_{i=0}^{k}\binom{n-r}{i}$. Since h has degree r and since the dimension of $An_k(h)$ is invariant under affine equivalence, we can assume without loss of generality that $h(x) = x_1x_2\cdots x_r + h'(x)$, where h' has degree at most r and where the term $x_1x_2\cdots x_r$ has null coefficient in its ANF. For any choice of $n-r$ bits $u_{r+1},\dots,u_n$, the restriction $h_{u_{r+1},\dots,u_n}$ of h obtained by fixing the variables $x_{r+1},\dots,x_n$ to the values $u_{r+1},\dots,u_n$ (respectively) has then degree r, and has therefore odd weight, since r is the number of its variables. Hence it has weight at least 1. For every $(u_{r+1},\dots,u_n) \in F_2^{n-r}$, let us denote by $x_{u_{r+1},\dots,u_n}$ a vector x such that $(x_{r+1},\dots,x_n) = (u_{r+1},\dots,u_n)$ and $h(x) = 1$.
A function $g(x) = \sum_{u\in F_2^n \mid wt(u)\le k} a_u x^u$ is an annihilator of h if and only if, for every $x \in F_2^n$ such that $h(x) = 1$, we have $g(x) = 0$. A necessary condition is that $g(x) = 0$ for every $x = x_{u_{r+1},\dots,u_n}$. If, in each of the resulting equations, we transfer all unknowns a_u such that $(u_1,\dots,u_r) \ne (0,\dots,0)$ to the right hand side, we obtain a system S' in the unknowns a_u such that $(u_1,\dots,u_r) = (0,\dots,0)$. Replacing the right hand sides of the resulting equations by 0 (i.e. considering the corresponding homogeneous system S'_0) gives the system that we obtain when we characterize the $(n-r)$-variable annihilators of degrees at most k of the constant function 1, considered as a function in the variables $x_{r+1},\dots,x_n$. Since the constant function 1 admits only the null function as annihilator, this means that the matrix of S'_0 has full rank $\sum_{i=0}^{k}\binom{n-r}{i}$. Hence the rank of the whole system of equations $\sum_{u\in F_2^n \mid wt(u)\le k} a_u x^u = 0$, where x ranges over the support of h, is at least $\sum_{i=0}^{k}\binom{n-r}{i}$ and the dimension of $A_k(h)$ is at most $\sum_{i=0}^{k}\binom{n}{i} - \sum_{i=0}^{k}\binom{n-r}{i}$.$\square$

Note that, for $k = n$, Lemma 1 gives $wt(h) \ge \max(\sum_{i=0}^{AI(h)-1}\binom{n}{i}, 2^{n-r})$ since $d_{n,h} = 2^n - wt(h)$, but we already knew that $wt(h) \ge 2^{n-r}$, see [36].

Remarks:
1. The bound of Lemma 1, which generalizes and improves upon the bound $d_{AI(h),h} \le \binom{n}{AI(h)}$ of [8], is tight for every value of $AI(h)$ (upper bounded by $\lceil \frac{n}{2} \rceil$) and for every $k \le n$. When $AI(h) = 1$, it is achieved at least by all functions of weight 2^{n-r}, according to Proposition 1 (all of these functions have algebraic

immunity 1). When $AI(h) = t > 1$, it is achieved by the function of Proposition 3. Note that in this case, the value of r is large. We do not assert that the bound is also tight for every r.
2. We show in Appendix that the bound of Lemma 1 can be improved in some cases. We do not know whether the stronger inequality $d_{k,h} \leq \sum_{i=AI(h)}^{k} \binom{n}{i} - \sum_{i=AI(h)}^{k} \binom{n-r}{i}$ can be true for every function of degree r and for some values of k (depending on the value of $AI(h)$). The functions of Proposition 1 are not counter-examples since they have algebraic immunity 1, and those of Proposition 3 neither since they have degree $r > n - t$. $\square$

4 The lower bound on the r-th order nonlinearity

We need a preliminary result before stating our main result.

Proposition 5. *Let f be a Boolean function in n variables and let r, r' be non-negative integers such that $r' \leq r$ and $AI(f) - r - 1 \geq 0$. For every n-variable function h of degree r and of algebraic immunity r', we have*

$$wt(fh) \geq \max\left(\sum_{i=0}^{r'-1} \binom{n}{i}, \sum_{i=0}^{AI(f)-r-1} \binom{n-r}{i}\right) \text{ if } r' \leq AI(f) - r - 1,$$
$$\geq \sum_{i=0}^{AI(f)-r-1} \binom{n}{i} \text{ if } r' > AI(f) - r - 1.$$

In all cases, we have:

$$wt(fh) \geq \sum_{i=0}^{AI(f)-r-1} \binom{n-r}{i}.$$

Proof:
Let k be any non-negative integer. A Boolean function of degree at most k belongs to $An_k(fh)$ if and only if the coefficients in its ANF satisfy a system of $wt(fh)$ equations in $\sum_{i=0}^{k} \binom{n}{i}$ variables. Hence we have: $\dim(An_k(fh)) \geq \sum_{i=0}^{k} \binom{n}{i} - wt(fh)$.
According to Lemma 1: $\dim(An_k(h)) \leq \min\left(\sum_{i=r'}^{k} \binom{n}{i}, \sum_{i=0}^{k} \binom{n}{i} - \sum_{i=0}^{k} \binom{n-r}{i}\right)$.
If $\dim(An_k(fh)) > \dim(An_k(h))$, then there exists an annihilator g of degree at most k of fh which is not an annihilator of h. Then, gh is a nonzero annihilator of f and has degree at most $k+r$. Thus, if $k = AI(f) - r - 1 \geq 0$, we arrive to a contradiction. We deduce that $\dim(An_{AI(f)-r-1}(fh)) \leq \dim(An_{AI(f)-r-1}(h))$. This implies that $\sum_{i=0}^{AI(f)-r-1} \binom{n}{i} - wt(fh)$ is upper bounded by:

$$\min\left(\sum_{i=r'}^{AI(f)-r-1} \binom{n}{i}, \sum_{i=0}^{AI(f)-r-1} \binom{n}{i} - \sum_{i=0}^{AI(f)-r-1} \binom{n-r}{i}\right),$$

that is:

$$wt(fh) \geq \max\left(\sum_{i=0}^{r'-1}\binom{n}{i}, \sum_{i=0}^{AI(f)-r-1}\binom{n-r}{i}\right) \text{ if } r' \leq AI(f)-r-1,$$

and

$$wt(fh) \geq \sum_{i=0}^{AI(f)-r-1}\binom{n}{i} \text{ if } r' > AI(f)-r-1. \qquad \square$$

Theorem 1. *Let f be any Boolean function in n variables and let r be any nonnegative integer such that $AI(f)-r-1 \geq 0$. Then $nl_r(f) \geq 2\sum_{i=0}^{AI(f)-r-1}\binom{n-r}{i}$. More precisely, we have $nl_r(f) \geq \max_{r' \leq n}(\min(\lambda_{r'}, \mu_{r'}))$, where:*

$$\begin{aligned}
\lambda_{r'} &= 2\ \max\left(\sum_{i=0}^{r'-1}\binom{n}{i}, \sum_{i=0}^{AI(f)-r-1}\binom{n-r}{i}\right) \text{ if } r' \leq AI(f)-r-1,\\
&= 2\sum_{i=0}^{AI(f)-r-1}\binom{n}{i} \text{ if } r' > AI(f)-r-1,\\
\mu_{r'} &= \sum_{i=0}^{AI(f)-r-1}\binom{n-r}{i} + \sum_{i=0}^{AI(f)-r'}\binom{n-r'+1}{i}.
\end{aligned}$$

Proof:
Let h be a function of degree at most r such that $nl_r(f) = wt(f+h) = wt(f(h+1)) + wt((f+1)h)$. Proposition 5 implies $nl_r(f) \geq 2\sum_{i=0}^{AI(f)-r-1}\binom{n-r}{i}$. Let r' be any nonnegative integer. If $AI(h) \geq r'$, then Proposition 5 applied to the functions f and $h+1$ and to the functions $f+1$ and h shows that $wt(f(h+1))$ and $wt((f+1)h)$ are lower bounded by:

$$\max\left(\sum_{i=0}^{r'-1}\binom{n}{i}, \sum_{i=0}^{AI(f)-r-1}\binom{n-r}{i}\right) \text{ if } r' \leq AI(f)-r-1,$$
$$\sum_{i=0}^{AI(f)-r-1}\binom{n}{i} \text{ if } r' > AI(f)-r-1.$$

If $AI(h) < r'$, then there exists $g \neq 0$ such that either $g \in An_{r'-1}(h+1)$, and therefore $supp(g) \subseteq supp(h)$, or $g \in An_{r'-1}(h)$, and therefore $supp(g) \subseteq supp(h+1)$. If $supp(g) \subseteq supp(h)$, then we apply Proposition 5 (last sentence of) to the functions f and $h+1$ and to the functions $f+1$ and g. We obtain: $wt(f(h+1)) \geq \sum_{i=0}^{AI(f)-r-1}\binom{n-r}{i}$ and $wt((f+1)h) \geq wt((f+1)g) \geq \sum_{i=0}^{AI(f)-r'}\binom{n-r'+1}{i}$. The case where $supp(g) \subseteq supp(h+1)$ is similar. $\square$

Remarks:
1. Let δ be the algebraic degree of f. We assume that $r < \delta$ since, otherwise, $nl_r(f)$ is null. According to McEliece's theorem (see [36] or [7]), $nl_r(f)$ is divisible by $2^{\lceil \frac{n}{\delta} \rceil -1} = 2^{\lfloor \frac{n-1}{\delta} \rfloor}$.
2. The bound of Theorem 1 improves upon the bound $nl_r(f) \geq \sum_{i=0}^{AI(f)-r-1} \binom{n}{i}$ of [8] for $r = 1$ (this is Lobanov's result). For $r > 1$, it improves upon it for every $n \leq 12$ and if we go up to $n \leq 22$, it improves upon it for every $AI(f)$ and most of the possible values of r (all of them, unless $AI(f)$ is very large). We give in Appendix the table of the values of the bounds of Theorem 1 and of [8] when $AI(f) = \lceil \frac{n}{2} \rceil$ (i.e. when $AI(f)$ is optimal - this is the worst case for the bound of Theorem 1), for all values of n from 13 to 22 and for all values of r from 1 to $AI(f) - 1$.
3. Both bounds of Theorem 1 and [8] show that, for most values of r, functions with high algebraic immunities are (much better than) "r-th order bent", in the sense of the paper [25]. In particular, they give more robust functions than those constructed in this paper.
4. In the case of optimal $AI(f) = \lceil \frac{n}{2} \rceil$, the bound of Theorem 1 gives $nl_r(f) \geq 2\left(2^{n-r-1} - \sum_{i=n/2-r}^{(n-r)/2-1} \binom{n-r}{i} - \frac{1}{2}\binom{n-r}{(n-r)/2}\right)$ when n and r are even; it gives $nl_r(f) \geq 2\left(2^{n-r-1} - \sum_{i=n/2-r}^{(n-r-1)/2} \binom{n-r}{i}\right)$ when n is even and r is odd, $nl_r(f) \geq 2\left(2^{n-r-1} - \sum_{i=(n+1)/2-r}^{(n-r)/2-1} \binom{n-r}{i} - \frac{1}{2}\binom{n-r}{(n-r)/2}\right)$ when n and r are odd and finally $nl_r(f) \geq 2\left(2^{n-r-1} - \sum_{i=(n+1)/2-r}^{(n-r-1)/2} \binom{n-r}{i}\right)$ when n is odd and r is even. This is in all cases asymptotically greater than $2^{n-r} - (r+1)\binom{n-r}{\lfloor \frac{n-r}{2} \rfloor} \approx 2^{n-r}\left(1 - \frac{(r+1)\sqrt{2}}{\sqrt{\pi n}}\right)$ when n tends to ∞ and r is fixed.
5. For $r \geq 2$, the asymptotic lower bound $2^{n-r}\left(1 - \frac{(r+1)\sqrt{2}}{\sqrt{\pi n}}\right)$ given by Theorem 1 is far from the asymptotic lower bound $2^{n-1} - 2^{n/2}\, n^{r/2}\sqrt{\frac{\ln 2}{2r!}}$ given in [13]. But this last bound, which can be shown by a simple argument (counting the number of those functions whose r-th order nonlinearities are smaller than this number and showing that it is negligible compared to the number of all functions), is only a proof of existence and gives absolutely no idea of what can be explicitly a function with greater r-th order nonlinearity. Before our bound, as far as we know, obtaining for every low $r > 1$ an *effective* way of designing functions with provably non-weak r-th order nonlinearities was open. Moreover (and most importantly), the functions with high algebraic immunities satisfy our bound *for every r of reasonably low value.*

5 The consequences on the necessary criteria for the use of Boolean functions in symmetric ciphers

The impact of our result may be greater for block ciphers and for self-synchronizing stream ciphers than for synchronous stream ciphers. Indeed, the role of the r-th order nonlinearity relatively to the currently known attacks has been more

clearly shown for the former than for the latter (see [14, 23, 25, 27, 33, 35]). As far as we know, and except in extreme situations, no explicit attack is known, that can use the approximation by a low degree function of the combining or filtering function in the pseudo-random generator of a stream cipher (except that an approximate pseudo-random sequence can be generated with a lower complexity, which may allow predicting bits after the observation of a part of the sequence). However, such attack may be found in the future and we know now that a high algebraic immunity will help in this matter. Note that a high algebraic immunity does not necessarily prevent the system from fast algebraic attacks, see e.g. [2], but these attacks were not the subject of the present paper.
The requirements concerning the Boolean functions used in symmetric ciphers are going towards a simplification. We know that the functions with optimum algebraic immunity are balanced, for n odd. A very high (first-order) nonlinearity takes care of the distance to linear structures, since we know that if a function has a nonzero linear structure then its nonlinearity is upper bounded by $2^{n-1} - 2^{\frac{n-1}{2}}$ and, therefore that, if its distance to linear structures is d then its nonlinearity is upper bounded by $2^{n-1} - 2^{\frac{n-1}{2}} + d$. Lobanov's bound does not guarantee a resistance to attacks using approximations by affine functions, since such resistance needs (see e.g. [22, 5]) a high nonlinearity, but Theorem 1 (with $r \geq 2$) shows that having a high algebraic immunity helps protecting against attacks by approximations by non-affine functions of low degrees, since the complexity of such attacks increases fastly with the degree.
However, the problem of determining functions meeting all the criteria needed for the combining or the filtering model of stream ciphers is still wide open.

References

1. F. Armknecht. Improving Fast Algebraic Attacks. FSE 2004, number 3017 in Lecture Notes in Computer Science, pp. 65–82. Springer Verlag, 2004.
2. F. Armknecht, C. Carlet, P. Gaborit, S. Knzli, W. Meier and O. Ruatta. Efficient computation of algebraic immunity for algebraic and fast algebraic attacks. Advances in Cryptology, EUROCRYPT 2006, Lecture Notes in Computer Science 4004 , pp. 147-164, 2006.
3. A. Braeken and B. Preneel. On the Algebraic Immunity of Symmetric Boolean Functions. Indocrypt 2005, LNCS 3797, pp. 35–48, 2005. Some false results of this reference have been corrected in Braeken's PhD thesis entitled "Cryptographic properties of Boolean functions and S-boxes" and available at URL http://homes.esat.kuleuven.be/ abraeken/thesisAn.pdf.
4. A. Canteaut. Open problems related to algebraic attacks on stream ciphers. Proceedings of WCC 2005, pp. 1-10, 2005.
5. A. Canteaut and M. Trabbia. Improved fast correlation attacks using parity-check equations of weight 4 and 5. EUROCRYPT 2000, number 1807 in Lecture Notes in Computer Science, pp. 573–588. Springer Verlag, 2000.
6. C. Carlet. On bent and highly nonlinear balanced/resilient functions and their algebraic immunities. Proceedings of AAECC 16, LNCS 3857, pp. 1-28, 2006.
7. 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 in 2006. Preliminary version available at http://www-rocq.inria.fr/codes/Claude.Carlet/pubs.html
8. C. Carlet, D. Dalai, K. Gupta and S. Maitra. Algebraic Immunity for Cryptographically Significant Boolean Functions: Analysis and Construction. To appear in IEEE Transactions on Information Theory, vol. 52, no. 7, July 2006.
9. C. Carlet and P. Gaborit. On the construction of balanced Boolean functions with a good algebraic immunity. Proceedings of BFCA (First Workshop on Boolean Functions: Cryptography and Applications), Rouen, France, March 2005, pp. 1-14.
10. C. Carlet, P Guillot and S. Mesnager. On immunity profile of Boolean functions. Proceedings of SETA'06 (International Conference on Sequences and their Applications). To appear in Lecture Notes in Computer Science.
11. C. Carlet and S. Mesnager. Improving the upper bounds on the covering radii of binary Reed-Muller codes. C. Carlet et S. Mesnager. To appear in IEEE Transactions on Information Theory, 2006.
12. F. Chabaud and S. Vaudenay. Links between Differential and Linear Cryptanalysis. EUROCRYPT'94, Advances in Cryptology, Lecture Notes in Computer Science 950, Springer Verlag, pp. 356-365, 1995.
13. G. Cohen, I. Honkala, S. Litsyn and A. Lobstein. *Covering codes*. North-Holland, 1997.
14. N. Courtois. Higher order correlation attacks, XL algorithm and cryptanalysis of Toyocrypt. Proceedings of ICISC 2002, LNCS 2587, pp. 182-199.
15. N. Courtois and W. Meier. Algebraic attacks on stream ciphers with linear feedback. Advances in Cryptology - EUROCRYPT 2003, number 2656 in Lecture Notes in Computer Science, pp. 345–359. Springer Verlag, 2003.
16. N. Courtois. Fast algebraic attacks on stream ciphers with linear feedback. CRYPTO 2003, number 2729 in Lecture Notes in Computer Science, pp. 176–194. Springer Verlag, 2003.
17. D. K. Dalai, K. C. Gupta and S. Maitra. Results on Algebraic Immunity for Cryptographically Significant Boolean Functions. Indocrypt 2004, Chennai, India, December 20–22, pp. 92–106, number 3348 in Lecture Notes in Computer Science, Springer Verlag, 2004
18. D. K. Dalai, K. C. Gupta and S. Maitra. Cryptographically Significant Boolean functions: Construction and Analysis in terms of Algebraic Immunity. Workshop on Fast Software Encryption, FSE 2005, pages 98–111, number 3557, Lecture Notes in Computer Science, Springer-Verlag.
19. D. K. Dalai, S. Maitra and S. Sarkar. Basic Theory in Construction of Boolean Functions with Maximum Possible Annihilator Immunity. Cryptology ePrint Archive, http://eprint.iacr.org/, No. 2005/229, 15 July, 2005. To be published in Designs, Codes and Cryptography.
20. D. K. Dalai, K. C. Gupta and S. Maitra. Notion of algebraic immunity and its evaluation related to fast algebraic attacks. Paper 2006/018 in http://eprint.iacr.org/
21. F. Didier. A new upper bound on the block error probability after decoding over the erasure channel. Preprint available at http://www-rocq.inria.fr/codes/Frederic.Didier/
A revised version will appear in IEEE Transactions on Information Theory, 2006.
22. R. Forré. A fast correlation attack on nonlinearly feedforward filtered shift register sequences. EUROCRYPT '89, Lecture Notes in Comput. Sci. 434, pp. 586-595, Springer, 1990.
23. J. Golic. Fast low order approximation of cryptographic functions. Proceedings of EUROCRYPT'96, LNCS 1070, pp. 268-282, 1996.

24. P. Hawkes and G. G. Rose. Rewriting Variables: The Complexity of Fast Algebraic Attacks on Stream Ciphers. CRYPTO 2004, LNCS 3152, pp. 390–406. Springer Verlag, 2004.
25. 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.
26. L.R. Knudsen. Truncated and higher order differentials. Fast Software Encryption, Second International Workshop, Lecture Notes in Computer Science, n 1008. pp. 196–211. – Springer-Verlag, 1995.
27. L.R. Knudsen and M. J. B. Robshaw. Non-linear approximations in linear cryptanalysis. Proceedings of EUROCRYPT'96, LNCS 1070, pp. 224-236, 1996.
28. X. Lai. Higher order derivatives and differential cryptanalysis. Proc. "Symposium on Communication, Coding and Cryptography", in honor of J. L. Massey on the occasion of his 60'th birthday. 1994.
29. K. Kurosawa, T. Johansson and D. Stinson. Almost k-wise independent sample spaces and their applications. J. of Cryptology, vol. 14, no. 4, pp. 231-253, 2001.
30. M. Lobanov. Tight bound between nonlinearity and algebraic immunity. Paper 2005/441 in http://eprint.iacr.org/
31. J.L. Massey. Shift-register synthesis and BCH decoding. IEEE Transactions on Information Theory, vol. 15, pp. 122–127, 1969.
32. M. Matsui. Linear cryptanalysis method for DES cipher. Advances in Cryptology - EUROCRYPT'93, number 765 in Lecture Notes in Computer Science. Springer-Verlag, pp. 386-397, 1994.
33. U. M. Maurer. New approaches to the design of self-synchronizing stream ciphers. Proceedings of EUROCRYPT'91. LNCS 547, pp. 458-471, 1991.
34. A. Menezes, P. van Oorschot, and S. Vanstone. *Handbook of Applied Cryptography.* CRC Press Series on Discrete Mathematics and Its Applications, 1996.
35. W. Millan. Low order approximation of cipher functions. Cryptographic Policy and Algorithms. LNCS 1029, pp. 144-155, 1996.
36. F. J. MacWilliams and N. J. Sloane. *The theory of error-correcting codes*, Amsterdam, North Holland. 1977.
37. W. Meier, E. Pasalic and C. Carlet. Algebraic attacks and decomposition of Boolean functions. EUROCRYPT 2004, number 3027 in Lecture Notes in Computer Science, pp. 474–491. Springer Verlag, 2004.
38. W. Meier and O. Staffelbach. Nonlinearity Criteria for Cryptographic Functions. EUROCRYPT' 89, Lecture Notes in Computer Science 434, Springer Verlag, pp. 549-562, 1990.
39. R. A. Rueppel *Analysis and design of stream ciphers* Com. and Contr. Eng. Series, Berlin, Heidelberg, NY, London, Paris, Tokyo, 1986
40. C.E. Shannon. Communication theory of secrecy systems. Bell system technical journal, 28, pp. 656-715, 1949.

6 Appendix

6.1 Remark on Lemma 1

For every choice of an element x in the support of h, let us denote by E_x the corresponding equation $\sum_{u\in F_2^n \mid wt(u)\leq k} a_u x^u = 0$. We obtain an equivalent system by replacing every equation E_x such that $x \notin \{x_{v_{r+1},\dots,v_n}; (v_{r+1},\dots,v_n) \in F_2^{n-r}\}$ (see the proof of Lemma 1) by the equation $E_x + E_{x_{v_{r+1},\dots,v_n}}$, where $v_{r+1},\dots,v_n$

are the $n-r$ last coordinates of x. Note that, in this equation, every a_u such that $(u_1,\dots,u_r)=(0,\dots,0)$ vanishes. The resulting system contains the 2^{n-r} equations $E_{x_{v_{r+1},\dots,v_n}}$ (in the $\sum_{i=0}^{k}\binom{n}{i}$ unknowns a_u such that $wt(u)\le k$) and $wt(h)-2^{n-r}$ equations in the $\sum_{i=0}^{k}\binom{n}{i}-\sum_{i=0}^{k}\binom{n-r}{i}$ unknowns a_u such that $wt(u)\le k$ and $(u_1,\dots,u_r)\ne(0,\dots,0)$. Let us denote these two sub-systems by S_k^1 and S_k^2, respectively. Moving in each of the equations $E_{x_{v_{r+1},\dots,v_n}}$ of S_k^1 all unknowns such that $(u_1,\dots,u_r)\ne(0,\dots,0)$ to the right hand side, we obtain a sub-system expressing the a_u's such that $(u_1,\dots,u_r)=(0,\dots,0)$ by means of the a_u's such that $(u_1,\dots,u_r)\ne(0,\dots,0)$, if such a_u's can exist (note that the number of equations in S_k^1 is greater than the number of the a_u's such that $(u_1,\dots,u_r)=(0,\dots,0)$). Indeed, we have seen that this system has full rank $\sum_{i=0}^{k}\binom{n-r}{i}$. Hence, a solution of S_k^2 may allow zero or one solution of S_k^1. Let $d'_{k,h}$ be the dimension of the vector space of solutions of S_k^2. Then we have $d_{k,h}\le d'_{k,h}$ and $d'_{k,h}$ equals $\sum_{i=0}^{k}\binom{n}{i}-\sum_{i=0}^{k}\binom{n-r}{i}$ minus the rank of the system S_k^2. Similarly, $d'_{k-1,h}$ equals $\sum_{i=0}^{k-1}\binom{n}{i}-\sum_{i=0}^{k-1}\binom{n-r}{i}$ minus the rank of the system S_{k-1}^2. We deduce that $d'_{k,h}\le d'_{k-1,h}+\binom{n}{k}-\binom{n-r}{k}$, since the rank of S_{k-1}^2 is upper bounded by the rank of S_k^2 (the system S_{k-1}^2 can be obtained from S_k^2 by erasing the unknowns a_u such that $wt(u)=k$). We deduce by induction on k that $d_{k,h}\le\sum_{i=AI'(h)}^{k}\binom{n}{i}-\sum_{i=AI'(h)}^{k}\binom{n-r}{i}$, where $AI'(h)$ equals the minimum value of k such that S_k^2 has non-trivial solutions. Note that $AI'(h)\le AI(h)$ and that $AI'(h)$ may be strictly smaller than $AI(h)$, since S_k^2 may have non-trivial solutions when $S_k^1\cup S_k^2$ has none. Hence, we cannot deduce that $d_{k,h}\le\sum_{i=AI(h)}^{k}\binom{n}{i}-\sum_{i=AI(h)}^{k}\binom{n-r}{i}$. But the bound $d_{k,h}\le\sum_{i=AI'(h)}^{k}\binom{n}{i}-\sum_{i=AI'(h)}^{k}\binom{n-r}{i}$ *may be better, in many concrete situations, than the bound of Lemma 1.*
In particular, for $k=n$, it implies $wt(h)\ge\sum_{i=0}^{AI'(h)-1}\binom{n}{i}+\sum_{i=AI'(h)}^{n}\binom{n-r}{i}$, and therefore (applying it also to $h+1$): $\sum_{i=0}^{AI'(h)-1}\binom{n}{i}+\sum_{i=AI'(h)}^{n}\binom{n-r}{i}\le wt(h)\le\sum_{i=0}^{n-AI'(h)}\binom{n}{i}+\sum_{i=AI'(h)}^{n}\binom{n-r}{i}$ and, when $AI(h)$ is not large and $AI'(h)$ is not too small, this is better than the double inequality $\max(\sum_{i=0}^{AI(h)-1}\binom{n}{i},2^{n-r})\le wt(h)\le\min(\sum_{i=0}^{n-AI(h)}\binom{n}{i},2^n-2^{n-r})$ implied by Lemma 1 for $k=n$.

6.2 Table

We give in the next table the values of the lower bounds of Theorem 1 and of [8], for n ranging from 13 to 22 (this covers all practical cases, currently for stream ciphers), for optimum algebraic immunity $\lceil\frac{n}{2}\rceil$ (note that this is the most unfavorable case for the bound of Theorem 1) and for r ranging from 1 to $AI(f)-1$.

n	$AI(f)$	r	The bound of Th. 1	The bound of [8]
13	7	1	3172	2380
		2	1124	1093
		3	352	378
		4	184	92
		5	28	14
		6	2	1
14	7	1	4760	3473
		2	1588	1471
		3	464	470
		4	212	106
		5	30	15
		6	2	1
15	8	1	12952	9949
		2	4760	4944
		3	1588	1941
		4	1026	576
		5	242	121
		6	32	16
		7	2	1
16	8	1	19898	14893
		2	6946	6885
		3	2186	2517
		4	1392	697
		5	274	137
		6	34	17
		7	2	1
17	9	1	52666	41226
		2	19898	21778
		3	6946	9402
		4	2186	3214
		5	1668	834
		6	308	154
		7	36	18
		8	2	1
18	9	1	82452	63004
		2	29786	31180
		3	9888	12616
		4	2942	4048
		5	1976	988
		6	344	172
		7	38	19
		8	2	1

n	$AI(f)$	r	The bound of Th. 1	The bound of [8]
19	10	1	213524	169766
		2	82452	94184
		3	29786	43796
		4	9888	16664
		5	6415	5036
		6	2320	1160
		7	382	191
		8	40	20
		9	2	1
20	10	1	339532	263950
		2	126008	137980
		3	43556	60460
		4	13770	21700
		5	8826	6196
		6	2702	1351
		7	422	211
		8	42	21
		9	2	1
21	11	1	863820	695860
		2	339532	401930
		3	126008	198440
		4	43556	82160
		5	15094	27896
		6	15094	7547
		7	3124	1562
		8	464	232
		9	44	22
		10	2	1
22	11	1	1391720	1097790
		2	527900	600370
		3	188368	280600
		4	62360	110056
		5	18804	35443
		6	18218	9109
		7	3588	1794
		8	508	254
		9	46	23
		10	2	1

Table 1. The lower bounds on $nl_r(f)$ given by Theorem 1 and by [8], for $13 \leq n \leq 22$, $r \leq AI(f) - 1$ and $AI(f)$ optimum