Improving the algebraic immunity of resilient and nonlinear functions

Claude Carlet*

Abstract

The currently known constructions of Boolean functions with high nonlinearities, high algebraic degrees and high resiliency orders do not seem to permit achieving sufficiently high algebraic immunities. We introduce a construction of Boolean functions, which builds a new function from three known ones. Assuming that the three functions have some resiliency order, nonlinearity and algebraic degree, as well as their sum modulo 2, the constructed function has the same resiliency order and can have the same nonlinearity, but has potentially better algebraic degree and algebraic immunity. The set of classical constructions together with this new one (and with a simpler derived one, having the same advantages) permit now to obtain functions achieving all necessary criteria for being used in the pseudo-random generators in stream ciphers.

Keywords : Algebraic attacks, Stream ciphers, Boolean Function, Algebraic Degree, Resiliency, nonlinearity.

1 Introduction

Boolean functions, that is, $\{0, 1\}$-valued functions defined on the set $F_2{}^n$ of all binary words of a given length n, are used in the pseudo-random generators of stream ciphers and play a central role in their security. The generation of the keystream consists, in many stream ciphers, of a linear part, producing a sequence with a large period, usually composed of one or several LFSR's, and a nonlinear combining or filtering function f that produces the output, given the state of the linear part.

*INRIA, Domaine de Voluceau, Rocquencourt, BP 105 - 78153, Le Chesnay Cedex, FRANCE; e-mail: claude.carlet@inria.fr; also member the University of Paris 8.

The main classical cryptographic criteria for designing such 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 counter linear synthesis by Berlekamp-Massey algorithm, a high order of correlation immunity (and more precisely, of resiliency, since the functions must be balanced) to counter correlation attacks (at least 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.

Since the introduction of these criteria, the problem of efficiently constructing highly resilient functions with high nonlinearities and algebraic degrees has received much attention. Few primary constructions are known, and secondary constructions are also necessary to obtain functions, on a sufficient number of variables, achieving or approaching the best possible cryptographic characteristics.

The recent algebraic attacks have dramatically complicated this situation. Algebraic attacks recover the secret key by solving an overdefined system of multivariate algebraic equations. The scenarios found in [8], under which low degree equations can be deduced from the knowledge of the nonlinear combining or filtering function, have led in [18] to a new parameter of the Boolean function: its algebraic immunity, which must be high.

No primary construction leading to functions with high algebraic immunity is known. The main known primary constructions of highly nonlinear resilient functions lead to functions with insufficient algebraic immunities. The known secondary constructions use functions on F_2^m, with $m < n$, to obtain functions on F_2^n, and they do not seem to permit achieving high algebraic immunity from functions with lower algebraic immunities. For instance, the 10-variable Boolean function used in the LILI keystream generator (a submission to NESSIE European call for cryptographic primitives) is built following [25] by using classical constructions; see [28]. It has algebraic immunity 4 and is responsible for the lack of resistance of LILI to algebraic attacks, as shown in [8]. Hence, we arrive now to a quite new situation, which is problematic: no satisfactory solution seems to exist for generating Boolean functions satisfying all necessary cryptographic criteria at sufficiently high levels!

As shown in [18], taking random balanced functions on sufficiently large numbers of variables could suffice to withstand algebraic attacks on the stream ciphers using them. As shown in [19], it would also permit to reach nonlinearities which would not be too far from the optimal ones.

But such solution is more or less a last resort, and it implies using functions on large numbers of variables, which reduces the efficiency of the corresponding stream ciphers. In any case, it does not permit to reach nonzero resiliency orders.
In this paper, we introduce a construction of functions on F_2^n from functions on F_2^n which, when combined with the classical primary and secondary constructions, leads to functions achieving high algebraic degrees, high nonlinearities and high resiliency orders, and also permits to attain potentially high algebraic immunity.
The paper is organized as follows. In Section 2, we recall the basic notions and properties. We also recall the known constructions of highly resilient functions and we explain why, in practice, they build functions whose algebraic immunity is too low. In Section 3, we introduce the new construction and a derived construction which is simpler, and we show why they lead potentially to functions with better algebraic degree and algebraic immunity.

2 Preliminaries

In this paper, we will deal in the same time with sums modulo 2 and with sums computed in $\mathbf{Z}$. We denote by $\oplus$ the addition in F_2 (but we denote by $+$ the addition in the field F_{2^n} and in the vectorspace F_2^n, since there will be no ambiguity) and by $+$ the addition in $\mathbf{Z}$. We denote by $\bigoplus_{i\in\ldots}$ (resp. $\sum_{i\in\ldots}$) the corresponding multiple sums. Let n be any positive integer. Any Boolean function f on n variables admits a unique algebraic normal form (A.N.F.):

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

where the a_I's are in F_2. We call *algebraic degree* of a Boolean function the degree of its algebraic normal form. Affine functions are those Boolean functions of degrees at most 1.
The *Hamming weight* $w_H(f)$ of a Boolean function f on n variables is the size of its support $\{x \in F_2^n;\ f(x) = 1\}$. The *Hamming distance* $d_H(f,g)$ between two Boolean functions f and g is the Hamming weight of their difference $f \oplus g$ (i.e. of their sum modulo 2). The *nonlinearity* of f is its minimum distance to all affine functions. Functions used in stream or block ciphers must have high nonlinearities to resist the attacks on these ciphers (correlation and linear attacks, see [2, 14, 15, 27]). The nonlinearity of f can be expressed by means of the discrete Fourier transform of the "sign" function $\chi_f(x) = (-1)^{f(x)}$, equal to $\widehat{\chi_f}(s) = \sum_{x\in F_2{}^n}(-1)^{f(x)\oplus x\cdot s}$

(and called the *Walsh transform*): the distance $d_H(f,l)$ between f and the affine function $l(x) = s \cdot x \oplus \epsilon$ ($s \in F_2^n$; $\epsilon \in F_2$) and the number $\widehat{\chi_f}(s)$ are related by:

$$\widehat{\chi_f}(s) = (-1)^\epsilon (2^n - 2d_H(f,l)) \qquad (1)$$

and the nonlinearity N_f of any Boolean function on F_2^n is therefore related to the Walsh spectrum of χ_f via the relation:

$$N_f = 2^{n-1} - \frac{1}{2} \max_{s \in F_2^n} |\widehat{\chi_f}(s)|. \qquad (2)$$

It is upper bounded by $2^{n-1} - 2^{n/2-1}$ because of the so-called Parseval's relation $\sum_{s \in F_2^n} \widehat{\chi_f}^2(s) = 2^{2n}$. A Boolean function is called *bent* if its nonlinearity equals $2^{n-1} - 2^{n/2-1}$, where n is necessarily even. Then, its distance to every affine function equals the maximum $2^{n-1} \pm 2^{n/2-1}$. But the function cannot be *balanced*, i.e. uniformly distributed. Hence, it cannot be used without modifications in the pseudo-random generator of a stream cipher, since this would leak statistical information on the plaintext, given the ciphertext.
If f is bent, then the *dual* Boolean function $\widetilde{f}$ defined on ${F_2}^n$ by $\widehat{\chi_f}(s) = 2^{\frac{n}{2}}(\chi_{\widetilde{f}})(s)$ is bent. The dual of $\widetilde{f}$ is f itself. Rothaus' inequality states that any bent function has algebraic degree at most $n/2$ [22].

A more important class of Boolean functions for cryptography is that of resilient functions. These functions play a central role in stream ciphers: in the standard model of these ciphers (cf. [26]), the outputs to n linear feedback shift registers are the input to a Boolean function. The output to the function produces the keystream, which is then bitwisely xored with the message to produce the cipher. Some devide-and-conquer attacks exist on this method of encryption (cf. [2, 14, 15, 27]) and lead to criteria the combining function must satisfy. Two main criteria are the following: the combining function must be balanced; it must also be such that the distribution probability of its output is unaltered when any m of its inputs are fixed [27], with m as large as possible. This property, called *m-th order correlation-immunity* [26], is characterized by the set of zero values in the Walsh spectrum [30]: f is m-th order correlation-immune if and only if $\widehat{\chi_f}(u) = 0$, for all $u \in F_2^n$ such that $1 \le w_H(u) \le m$, where $w_H(u)$ denotes the Hamming weight of the n-bit vector u, (the number of its nonzero components). Balanced m-th order correlation-immune functions are called *m-resilient* functions. They are characterized by the fact that $\widehat{\chi_f}(u) = 0$ for all $u \in F_2^n$ such that $0 \le w_H(u) \le m$.
Siegenthaler's inequality [26] states that any m-th order correlation immune function on n variables has degree at most $n - m$, that any m-resilient function ($0 \le m < n-1$) has algebraic degree smaller than or

equal to $n-m-1$ and that any $(n-1)$-resilient function has algebraic degree 1. We shall call *Siegenthaler's bound* this property.
Sarkar and Maitra have shown that the Hamming distance between any m-resilient function and any affine function is divisible by 2^{m+1}. This has led to an upper bound on the nonlinearity of m-resilient functions (also partly obtained by Tarannikov and by Zhang and Zheng): the nonlinearity of any m-resilient function is smaller than or equal to $2^{n-1}-2^{m+1}$ if $\frac{n}{2}-1<m+1$, to $2^{n-1}-2^{\frac{n}{2}-1}-2^{m+1}$ if n is even and $\frac{n}{2}-1 \geq m+1$ and to $2^{n-1}-2^{m+1}\left\lceil 2^{n/2-m-2}\right\rceil$ if n is odd and $\frac{n}{2}-1 \geq m+1$. We shall call this set of upper bounds *Sarkar et al.'s bound*. A similar bound exists for correlation immune functions, but we do not recall it since non-balanced correlation immune functions present small cryptographic interest.
Constructions providing resilient functions with degrees and nonlinearities approaching or achieving the known bounds are necessary for the design of stream ciphers. Two kinds of constructions can be identified. Some constructions give direct definitions of Boolean functions. There are few such *primary* constructions and new ideas for designing them are currently lacking. Moreover, the only known primary construction of resilient functions which leads to a wide class of such functions, the Maiorana-McFarland's construction, does not permit to design functions with optimum degrees and nonlinearities except for small values of the number of variables. Modifications and generalizations of this construction have been proposed (see e.g. [4, 20, 24]), but the generalizations lead to classes with the same properties as the original class and the number of the functions the modifications permit to construct is small (and they do not have good algebraic immunity, see below). Non-primary constructions use previously defined functions (that we shall call "building blocks" in the sequel) to build new ones and often lead in practice to recursive constructions. They are called *secondary* constructions. No simple secondary construction using, as building blocks, functions defined on the same space F_2^n as the constructed functions is known (we recall below the two known examples of such constructions; they need assumptions which are hard to satisfy). The purpose of the present paper is to introduce such general construction and to study its properties.

Until recently, a high algebraic degree, a high resiliency order and a high nonlinearity were the only requirements needed for the design of the function f used in a stream cipher as a combining function or as a filtering one. The recent algebraic attacks [8] have changed this situation by adding a new criterion of considerable importance to this list. Algebraic attacks recover the secret key by solving an overdefined system of multivariate algebraic equations. These attacks exploit multivariate relations involving key/state bits and output bits of f. If one such relation is found that is of low degree in the key/state bits, algebraic attacks are very efficient [7]. It is demonstrated in [8] that low degree relations and thus successful algebraic attacks exist for several well known constructions of stream ciphers that are immune to all previously known attacks. These low degree relations are obtained by producing low degree polynomial multiples of f, i.e., by multiplying the Boolean function f by a well chosen low degree function g, such that the product function $f * g$ (that is, the function which support equals the intersection of the supports of f and g) is again of low degree.

The scenarios found in [8], under which low degree multiples of a Boolean function may exist, have been simplified in [18] into two scenarios: (1) there exists a non-zero Boolean function g of low degree whose support is disjoint from the support of f (such a function g is called an annihilator of f); (2) there exists a non-zero Boolean function g of low degree whose support is included in the support of f (i.e. such that g is an annihilator of $f \oplus 1$). We write then: $g \preceq f$.

The *algebraic immunity* $AI(f)$ of a Boolean function f is the minimum value of d such that f or $f \oplus 1$ admits an annihilator of degree d. It should be high enough (at least equal to 6).

2.1 The known constructions of resilient functions and the corresponding degrees and nonlinearities

2.2 Primary constructions

2.2.1 Maiorana-McFarland constructions

In [1] is introduced a construction of resilient functions based on the idea of a construction of bent functions due to Maiorana and Mcfarland:

let m and $n = r + s$ be any positive integers ($r > m > 0$, $s > 0$), g any Boolean function on F_2^s and ϕ a mapping from F_2^s to F_2^r such that every element in $\phi(F_2^s)$ has Hamming weight strictly greater than m, then the function:

$$f(x, y) = x \cdot \phi(y) \oplus g(y),\ x \in F_2^r,\ y \in F_2^s \tag{3}$$

is m-resilient, since we have $\widehat{\chi_f}(a, b) = 2^r \sum_{y \in \phi^{-1}(a)} (-1)^{g(y) \oplus b \cdot y}$.

The functions of the form (3), for $\frac{n}{2} - 1 < m + 1$, can have high nonlinearities. However, optimum or nearly optimal functions could be obtained with this construction only with functions in which r was large and s was small. The functions being then concatenations of affine functions on a pretty large number of variables, their algebraic immunity can hardly achieve high values. This can be checked by computer experiment. In the case $\frac{n}{2}-1 \geq m+1$, no function belonging to Maiorana-McFarland's class and having nearly optimal nonlinearity could be constructed, except in the limit case $\frac{n}{2}-1 = m+1$. Generalizations of Maiorana-McFarland's construction exist (see e.g. [4]) but they have more or less the same behavior as the original construction. Modifications have also been proposed (see e.g. [21], in which some affine functions, at least one, are replaced by suitably chosen nonlinear functions) but it is shown in [18] that the algebraic immunities of these functions are often low.

2.2.2 Effective partial-spreads constructions

An idea due to Dillon for constructing bent functions is used in [3] to obtain a construction of correlation-immune functions:
Let s and r be two positive integers and $n = r + s$, g a function from F_{2^r} to F_2, ϕ a linear mapping from F_2^s to F_{2^r} and u an element of F_{2^r} such that $u{+}\phi(y) \neq 0$, $\forall y \in F_2^s$.
Let f be the function from $F_{2^r} \times F_2^s \sim F_2^n$ to F_2 defined by:

$$f(x,y) = g\left(\frac{x}{u{+}\phi(y)}\right) \oplus b \cdot y, \tag{4}$$

where $v \in F_2^s$. If, for every z in F_{2^r}, $\phi^*(z) \oplus v$ has weight greater than m, where $\phi^* : F_{2^r} \mapsto F_2^s$ is the adjoint of ϕ, then f is m-resilient.
The same observations as for Maiorana-McFarland's construction on the ability of these functions to have nonlinearities near Sarkar-Maitra's bound can be made. However, these functions have potentially higher algebraic immunities, as can be checked by computer experiment. So this class of functions, which has, until now, not been much used to construct functions, may present more interest now. Nevertheless, this class of functions is small and gives little opportunity to satisfy additional conditions needed in practice (because of the implementation, ...).

2.3 Secondary constructions

We shall call constructions with extension of the number of variables those constructions using functions on F_2^m, with $m < n$, to obtain functions on F_2^n.

2.3.1 General constructions with extension of the number of variables

Direct sums of functions: if f is an r-variable t-resilient function and if g is an s-variable m-resilient function, then the function:

$$h(x_1,\dots,x_r,x_{r+1},\dots,x_{r+s}) = f(x_1,\dots,x_r)\oplus g(x_{r+1},\dots,x_{r+s})$$

is $(t+m+1)$-resilient. This comes from the easily provable relation $\widehat{\chi_h}(a,b)=\widehat{\chi_f}(a)\times\widehat{\chi_g}(b)$, $a\in F_2^r$, $b\in F_2^s$. We have also $d^\circ h=\max(d^\circ f,d^\circ g)$ and, thanks to Relation (2), $\mathcal{N}_h = 2^{r+s-1}-\frac{1}{2}(2^r-2\mathcal{N}_f)(2^s-2\mathcal{N}_g) = 2^r\mathcal{N}_g+2^s\mathcal{N}_f-2\mathcal{N}_f\mathcal{N}_g$. But such function has low algebraic immunity, since we have $AI(h)\le\min(AI(f),AI(g))$.

Siegenthaler's construction: Let f and g be two Boolean functions on F_2^r. Consider the function

$$h(x_1,\dots,x_r,x_{r+1}) = (x_{r+1}\oplus 1)f(x_1,\dots,x_r)\oplus x_{r+1}g(x_1,\dots,x_r)$$

on F_2^{r+1}. Then:

$$\widehat{\chi_h}(a_1,\dots,a_r,a_{r+1}) = \widehat{\chi_f}(a_1,\dots,a_r)+(-1)^{a_{r+1}}\,\widehat{\chi_g}(a_1,\dots,a_r).$$

Thus, if f and g are m-resilient, then h is m-resilient; moreover, if for every $a\in F_2^r$ of Hamming weight $m+1$, we have $\widehat{\chi_f}(a)+\widehat{\chi_g}(a)=0$, then h is $(m+1)$-resilient. And we have: $\mathcal{N}_h\ge\mathcal{N}_f+\mathcal{N}_g$. If f and g achieve maximum possible nonlinearity $2^{r-1}-2^{m+1}$ and if h is $(m+1)$-resilient, then the nonlinearity 2^r-2^{m+2} of h is the best possible. If the supports of the Walsh transforms of f and g are disjoint, then we have $\mathcal{N}_h=2^{r-1}+\min(\mathcal{N}_f,\mathcal{N}_g)$; thus, if f and g achieve maximum possible nonlinearity $2^{r-1}-2^{m+1}$, then h achieves best possible nonlinearity 2^r-2^{m+1}. But we could not obtain good algebraic immunity with such functions. The reason is the following: we have $(x_{r+1}\oplus 1)f'\oplus x_{r+1}g'\preceq (x_{r+1}\oplus 1)f\oplus x_{r+1}g\oplus 1$ if and only if $f'\preceq f\oplus 1$ and $g'\preceq g\oplus 1$, and the degree of $(x_{r+1}\oplus 1)f'\oplus x_{r+1}g'$ is upper bounded by the degree of $f'\oplus g'$ plus 1; it seems difficult to avoid the existence of annihilators f' and g' of f and g such that $f'\oplus g'$ has low degree, and to achieve high resiliency order and/or high nonlinearity with h.

Tarannikov's construction: Let g be any Boolean function on ${F_2}^r$. Define the Boolean function h on ${F_2}^{r+1}$ by $h(x_1,\dots,x_r,x_{r+1}) = x_{r+1}\oplus g(x_1,\dots,x_{r-1},x_r\oplus x_{r+1})$. The Walsh transform $\widehat{\chi_h}(a_1,\dots,a_{r+1})$ is equal to $\displaystyle\sum_{x_1,\dots,x_{r+1}\in F_2}(-1)^{a\cdot x\oplus g(x_1,\dots,x_r)\oplus a_r x_r\oplus(a_r\oplus a_{r+1}\oplus 1)x_{r+1}}$, where we write $a=$

$(a_1, \ldots, a_{r-1})$ and $x = (x_1, \ldots, x_{r-1})$; it is null if $a_{r+1} = a_r$ and it equals $2\ \widehat{\chi_g}(a_1, \ldots, a_{r-1}, a_r)$ if $a_r = a_{r+1} \oplus 1$. Thus: $\mathcal{N}_h = 2\ \mathcal{N}_g$; If g is m-resilient, then h is m-resilient. If, additionally, $\widehat{\chi_g}(a_1, \ldots, a_{r-1}, 1)$ is null for every vector $(a_1, \ldots, a_{r-1})$ of weight at most m, then h is $(m+1)$-resilient.
Tarannikov in [29], and after him, Pasalic et al. in [23] used this construction to design a more complex one, that we call *Tarannikov et al.'s construction*, which permitted to achieve maximum tradeoff between resiliency, algebraic degree and nonlinearity. It uses (see [23]) two $(n-1)$-variable m-resilient functions f_1 and f_2 achieving Siegenthaler's and Sarkar et al.'s bounds to design an $(n+3)$-variable $(m+2)$-resilient function h also achieving these bounds, assuming that $f_1 + f_2$ has same degree as f_1 and f_2 and that the supports of the Walsh transforms of f_1 and f_2 are disjoint. The two restrictions $h_1(x_1, \ldots, x_{n+2}) = h(x_1, \ldots, x_{n+2}, 0)$ and $h_2(x_1, \ldots, x_{n+2}) = h(x_1, \ldots, x_{n+2}, 1)$ have then also disjoint Walsh supports, and these two functions can then be used in the places of f_1 and f_2. This permits to generate functions achieving Sarkar et al.'s and Siegenthaler's bounds on sufficiently high numbers of variables. But Tarannikov et al.'s construction does not seem either to permit to achieve high algebraic immunities. It only permits to keep the same algebraic immunity for h as for the building blocks f, g, ...

Another construction: There exists a secondary construction of resilient functions from bent functions (see [3]): let r be a positive integer, m a positive even integer and f a function such that, for any element x', the function: $f_{x'} : x \to f(x, x')$ is bent. If, for every element u of Hamming weight at most t, the function $\varphi_u : x' \to \widetilde{f_{x'}}(u)$ is $(t - w_H(u))$-resilient, then f is t-resilient (the converse is true). This construction does not seem able to produce functions with higher algebraic immunities than the functions used as building blocks.

2.3.2 Constructions without extension of the number of variables

Such constructions, by modifying the support of highly nonlinear resilient functions without decreasing their characteristics, seem more appropriate for trying to increase the algebraic immunities of such functions, previously obtained by classical constructions. There exist, in the literature, two such constructions.

Modifying a function on a subspace: A construction due to Dillon for bent functions has been adapted to correlation-immune functions in

[3]: let t, m and n any positive integers and f a t-th order correlation-immune function from F_2^n to F_2^m; assume there exists a subspace E of F_2^n, whose minimum nonzero weight is greater than t and such that the restriction of f to the orthogonal of E (i.e. the subspace of F_2^n: $E^{\perp} = \{u \in F_2^n \,|\, \forall x \in E,\, u \cdot x = 1\}$) is constant. Then f remains t-th order correlation-immune if we change its constant value on $E^{\perp}$ into any other one.

Composing with a permutation: An idea of construction has been given in [13] for bent functions and can be adapted to resilient functions: if $d_H(f, \sum_{i=1}^{n} a_i\, \sigma_i) = 2^{n-1}$ for every $a \in {F_2}^n$ of weight at most k, then $f \circ \sigma^{-1}$ is k-resilient.

But these two secondary constructions without extension of the number of variables above need strong hypothesis on the functions used as buiding blocks to produce resilient functions. Hence, they seem inefficient to construct classes of new functions.

3 A new secondary construction of Boolean functions

Given three Boolean functions f_1, f_2 and f_3, there is a nice relationship between their Walsh transforms and the Walsh transforms of two of their elementary symmetric related functions:

Lemma 1 *Let f_1, f_2 and f_3 be three Boolean functions on F_2^n. Denote by σ_1 the Boolean function equal to $f_1 \oplus f_2 \oplus f_3$ and by σ_2 the Boolean function equal to $f_1 f_2 \oplus f_1 f_3 \oplus f_2 f_3$. Then we have $f_1 + f_2 + f_3 = \sigma_1 + 2\sigma_2$. This implies*

$$\widehat{\chi_{f_1}} + \widehat{\chi_{f_2}} + \widehat{\chi_{f_3}} = \widehat{\chi_{\sigma_1}} + 2\widehat{\chi_{\sigma_2}}. \tag{5}$$

Proof. The fact that $f_1 + f_2 + f_3 = \sigma_1 + 2\sigma_2$ (recall that the sums are computed in $\mathbf{Z}$ and not mod 2) can be checked easily and directly implies $\chi_{f_1} + \chi_{f_2} + \chi_{f_3} = \chi_{\sigma_1} + 2\chi_{\sigma_2}$, thanks to the equality $\chi_f = 1 - 2f$ (valid for every Boolean function). The linearity of the Fourier transform with respect to the addition in $\mathbf{Z}$ implies then relation (5). ◇

We use now this observation to derive constructions of resilient functions with high nonlinearities. In the following theorem, saying that a function f is 0-order correlation immune does not impose any condition on f and saying it is 0-resilient means it is balanced.

Theorem 1 *Let n be any positive integer and k any non-negative integer such that $k \leq n$. Let f_1, f_2 and f_3 be three k-th order correlation immune (resp. k-resilient) functions. Then the function $\sigma_1 = f_1 \oplus f_2 \oplus f_3$ is k-th order correlation immune (resp. k-resilient) if and only if the function $\sigma_2 = f_1 f_2 \oplus f_1 f_3 \oplus f_2 f_3$ is k-th order correlation immune (resp. k-resilient). Moreover:*

$$N_{\sigma_2} \geq \frac{1}{2}\left(N_{\sigma_1} + \sum_{i=1}^{3} N_{f_i}\right) - 2^{n-1} \tag{6}$$

and if the Walsh supports of f_1, f_2 and f_3 are pairwise disjoint (that is, if at most one value $\widehat{\chi_{f_i}}(s)$, $i = 1, 2, 3$ is nonzero, for every vector s), then

$$N_{\sigma_2} \geq \frac{1}{2}\left(N_{\sigma_1} + \min_{1 \leq i \leq 3} N_{f_i}\right). \tag{7}$$

Proof. Relation (5) and the fact that for every non-zero vector a of weight at most k we have $\widehat{\chi_{f_i}}(a) = 0$ for $i = 1, 2, 3$ imply that $\widehat{\chi_{\sigma_1}}(a) = 0$ if and only if $\widehat{\chi_{\sigma_2}}(a) = 0$. Same property occurs for $a = 0$ in the case f_1, f_2 and f_3 are resilient. Relation (5) implies the relation $\max_{s \in F_2^n} |\widehat{\chi_{\sigma_2}}(s)| \leq \frac{1}{2}\left(\sum_{i=1}^{3}\left(\max_{s \in F_2^n} |\widehat{\chi_{f_i}}(s)|\right) + \max_{s \in F_2^n} |\widehat{\chi_{\sigma_1}}(s)|\right)$ and Relation (2) implies then Relation (6). If the Walsh supports of f_1, f_2 and f_3 are pairwise disjoint, then Relation (5) implies the relation

$$\max_{s \in F_2^n} |\widehat{\chi_{\sigma_2}}(s)| \leq \frac{1}{2}\left(\max_{1 \leq i \leq 3}\left(\max_{s \in F_2^n} |\widehat{\chi_{f_i}}(s)|\right) + \max_{s \in F_2^n} |\widehat{\chi_{\sigma_1}}(s)|\right)$$

and Relation (2) implies then Relation (7). ◇

We show in Appendix a way of applying Theorem 1.
Remark: The same transformation can be used to construct bent functions. This will be the subject of a forthcoming paper.

We use now the invariance of the notion of correlation-immune (resp. resilient) function under translation to deduce a more practical (but less general) result.

Proposition 1 *Let n be any positive integer and k any non-negative integer such that $k \leq n$. Let f and g be two k-th order correlation immune (resp. k-resilient) functions on F_2^n. Assume that there exist $a, b \in F_2^n$ such that $D_a f \oplus D_b g$ is constant. Then the function $h(x) = f(x) \oplus D_a f(x)(f(x) \oplus g(x))$, that is, $h(x) = \begin{cases} f(x) \text{ if } D_a f(x) = 0 \\ g(x) \text{ if } D_a f(x) = 1 \end{cases}$ is k-th order correlation immune (resp. k-resilient). Moreover:*

$$N_h \geq N_f + N_g - 2^{n-1} \tag{8}$$

and if the Walsh support of f is disjoint of that of g, then

$$N_h \geq \min(N_f, N_g). \tag{9}$$

Note that finding hihgly nonlinear resilient functions with disjoint supports is easy, by using Tarannikov et al.'s construction.

Proof. Let $D_a f \oplus D_b g = \epsilon$. Taking $f_1(x) = f(x)$, $f_2(x) = f(x+a)$ and $f_3(x) = g(x)$, the hypothesis of Theorem 1 is satisfied, since $\sigma_1(x) = D_a f(x) \oplus g(x) = D_b g(x) \oplus \epsilon \oplus g(x) = g(x+b) \oplus \epsilon$ is k-th order correlation immune (resp. k-resilient). Hence, $h(x) = f(x) \oplus D_a f(x)(f(x) \oplus g(x))$ is k-th order correlation immune (resp. k-resilient). Relation (8) is a direct consequence of Relation (6). Note that the Walsh support of f_2 equals that of $f_1 = f$, since we have $\widehat{\chi_{f_2}}(s) = (-1)^{a \cdot s} \widehat{\chi_f}(s)$ and that the Walsh support of σ_1 equals that of $f_3 = g$. Hence, if the Walsh support of f is disjoint of that of g, then Relation (5) implies the relation

$$\max_{s \in F_2^n} |\widehat{\chi_h}(s)| \leq \max\left(\max_{s \in F_2^n} |\widehat{\chi_f}(s)|, \max_{s \in F_2^n} |\widehat{\chi_g}(s)|\right)$$

and Relation (2) implies then Relation (9). ◇

Example: choose f and g in Maiorana-McFarland's class; that is, $f(x,y) = x \cdot \phi(y) \oplus h(y),\ g(x,y) = x \cdot \psi(y) \oplus k(y),\ x \in F_2^r,\ y \in F_2^s$, where every element in $\phi(F_2^s)$ and in $\psi(F_2^s)$ has Hamming weight greater than m. For every $a, b \in F_2^r$ and $c, d \in F_2^s$, we have: $D_{(a,c)} f(x,y) = x \cdot D_c \phi(y) \oplus a \cdot \phi(y+c) \oplus D_c h(y)$ and $D_{(b,d)} g(x,y) = x \cdot D_d \psi(y) \oplus b \cdot \psi(y+d) \oplus D_d k(y)$. Hence, if there exist c and d such that $D_c \phi = D_d \psi$, and a and b such that $a \cdot \phi(y+c) \oplus D_c h(y) \oplus b \cdot \psi(y+d) \oplus D_d k(y)$ is constant, then the function h equal to $f(x,y)$ if $D_{(a,c)} f(x,y) = 0$ and to $g(x,y)$ if $D_{(a,c)} f(x,y) = 1$ is m-resilient. If the sets $\phi(F_2^s)$ and $\psi(F_2^s)$ are disjoint, then we have $N_h \geq \min(N_f, N_g)$. Note that, in general, h does not belong to Maiorana-McFarland's class.

Remark: The notion of resilient function being also invariant under any permutation of the input coordinates $x_1, \ldots, x_n$, Proposition 1 is also valid if we replace $D_a f$ by $f(x_1, \ldots, x_n) \oplus f(x_{\tau(1)}, \ldots, x_{\tau(n)})$ and $D_b g$ by $g(x_1, \ldots, x_n) \oplus g(x_{\tau'(1)}, \ldots, x_{\tau'(n)})$, where τ and τ' are two permutations of $\{1, \ldots, n\}$.

Computer experiment shows that the secondary construction of Theorem 1 and its particular case given in Proposition 1 permit to increase the algebraic immunity, while keeping the same resiliency order and the same *nonlinearity. The reason is in the fact that the support of σ_2 (resp. h) is, in general, more complex than those of f_1, f_2 and f_3 (resp. f and g).* This was not the case with the previously known secondary constructions.

We give in Appendix a primary construction of resilient functions deduced from Theorem 1 and a generalization of Lemma 1.

References

[1] P. Camion, C. Carlet, P. Charpin, N. Sendrier, On correlation-immune functions, *Advances in Cryptology: Crypto '91, Proceedings, Lecture Notes in Computer Science*, V. 576 (1991), pp. 86–100.

[2] Canteaut, A. and Trabbia, M. Improved fast correlation attacks using parity-check equations of weight 4 and 5, *Advanced in Cryptology-EUROCRYPT 2000. Lecture notes in computer science* 1807 (2000), pp. 573-588.

[3] Carlet, C. (1997). More correlation-immune and resilient functions over Galois fields and Galois rings. *Advances in Cryptology, EUROCRYPT' 97, Lecture Notes in Computer Science* 1233, 422-433, Springer Verlag.

[4] C. Carlet. A larger Class of Cryptographic Boolean Functions via a Study of the Maiorana-McFarland Construction. *Advances in Cryptology - CRYPT0 2002, no. 2442 in Lecture Notes in Computer Science*, pp. 549-564, 2002.

[5] C. Carlet. On the secondary constructions of resilient and bent functions. Proceedings of the 2003 Workshop on Coding, Cryptography and Combinatorics. BirkHauser Verlag, 2004.

[6] Chee, S., Lee, S., Kim, K. and Kim, D. Correlation immune functions with controlable nonlinearity. *ETRI Journal*, vol 19, no 4, pp. 389-401 (1997).

[7] N. Courtois. Fast Algebraic Attacks on Stream Ciphers with Linear Feedback. *Advances in cryptology–CRYPTO 2003, Lecture Notes in Computer Science* 2729, pp. 177-194, Springer, 2003.

[8] N. Courtois and W. Meier. Algebraic Attacks on Stream Ciphers with Linear Feedback. *Advances in cryptology–EUROCRYPT 2003, Lecture Notes in Computer Science* 2656, pp. 346-359, Springer, 2002.

[9] N. Courtois and J. Pieprzyk. Cryptanalysis of block ciphers with overdefined systems of equations. *Advances in cryptology–ASIACRYPT 2002, Lecture Notes in Computer Science* 2501, pp. 267-287, Springer, 2003.

[10] Dillon, J. F. (1974). *Elementary Hadamard Difference sets*. Ph. D. Thesis, Univ. of Maryland.

[11] Dillon, J. F. (1975). Elementary Hadamard Difference sets, *Proc. Sixth S-E Conf. Comb. Graph Theory and Comp.*, F. Hoffman et al. (Eds), Winnipeg Utilitas Math, 237-249.

[12] J.-C. Faugère and G. Ars. An Algebraic Cryptanalysis of Nonlinear Filter Generators using Gröbner bases. *Rapport de Recherche INRIA* 4739, 2003.

[13] Hou, X.-D. and P. Langevin (1997) Results on bent functions, *Journal of Combinatorial Theory, Series A*, 80, 232-246.

[14] Johansson, T. and Jönsson, F. Improved fast correlation attack on stream ciphers via convolutional codes. *Advances in Cryptology - EUROCRYPT'99, number 1592 in Lecture Notes in Computer Science* (1999), pp. 347–362.

[15] Johansson, T. and Jönsson, F. Fast correlation attacks based on turbo code techniques. *Advances in Cryptology - CRYPTO'99, number 1666 in Lecture Notes in Computer Science* (1999), pp. 181–197.

[16] Mac Williams, F. J. and N. J. Sloane (1977). *The theory of error-correcting codes*, Amsterdam, North Holland.

[17] Meier, W. and O. Staffelbach (1990). Nonlinearity Criteria for Cryptographic Functions, *Advances in Cryptology, EUROCRYPT' 89, Lecture Notes in Computer Science 434, 549-562, Springer Verlag.*

[18] W. Meier, E. Pasalic and C. Carlet. Algebraic attacks and decomposition of Boolean functions. *Advances in Cryptology, EUROCRYPT 2004, Lecture Notes in Computer Science, Springer Verlag* 3027, pp. 474-491, 2004.

[19] D. Olejár and M. Stanek. "On cryptographic properties of random Boolean functions." Journal of Universal Computer Science, vol. 4, No.8, pp. 705-717, 1998.

[20] E. Pasalic and S. Maitra. A Maiorana-McFarland type construction for resilient Boolean functions on n variables (n even) with nonlinearity $> 2^{n-1} - 2^{n/2} + 2^{n/2-2}$. *Proceedings of the Workshop on Coding and Cryptography 2003*, pp. 365-374, 2003.

[21] E. Pasalic. Degree optimized resilient Boolean functions from Maiorana-McFarland class. In *9th IMA Conference on Cryptography and Coding*, 2003.

[22] Rothaus, O. S. (1976). On bent functions, *J. Comb. Theory*, 20A, 300-305.

[23] E. Pasalic, T. Johansson, S. Maitra and P. Sarkar. New constructions of resilient and correlation immune Boolean functions achieving upper bounds on nonlinearity. *Proceedings of the Workshop on Coding and Cryptography* 2001, published by *Electronic Notes in Discrete Mathematics*, Elsevier, vo. 6, pp. 425-434, 2001.

[24] P. Sarkar and S. Maitra. Construction of nonlinear Boolean functions with important cryptographic properties. *Advances in Cryptology - EUROCRYPT 2000*, no. 1807 in Lecture Notes in Computer Science, Springer Verlag, pp. 485-506, 2000.

[25] Sarkar, P. and Maitra, S. Nonlinearity Bounds and Constructions of Resilient Boolean Functions. *CRYPTO 2000, LNCS* Vol. 1880, ed. Mihir Bellare, pp. 515-532 (2000).

[26] Siegenthaler, T. Correlation-immunity of nonlinear combining functions for cryptographic applications. *IEEE Transactions on Information theory*, V. IT-30, No 5 (1984), pp. 776-780.

[27] Siegenthaler, T. Decrypting a Class of Stream Ciphers Using Ciphertext Only. *IEEE Transactions on Computer, V. C-34*, No 1, pp. 81-85, 1985.

[28] L. Simpson, E. Dawson, J. Golic and W. Millan. LILI Keystream generator, *P roceedings of SAC'2000, Lecture Notes in Computer Science* 1807, Springer, pp. 248-261, 2001; cf. `www.isrc.qut.edu.au/lili/`.

[29] Y. V. Tarannikov. On resilient Boolean functions with maximum possible nonlinearity. *Proceedings of INDOCRYPT 2000, Lecture Notes in Computer Science* 1977, pp. 19-30, 2000.

[30] Xiao Guo-Zhen and Massey, J. L. A Spectral Characterization of Correlation-Immune Combining Functions. *IEEE Trans. Inf. Theory*, Vol IT 34, n° 3 (1988), pp. 569-571.

4 Appendix

4.1 Deducing a primary construction of resilient functions from Theorem 1

The following primary constructions are direct consequences of Theorem 1, where the functions f_1, f_2 and f_3 are chosen in Maiorana McFarland's classes.

Proposition 2 *let t and $n = r + s$ be any positive integers ($r > t > 0$, $s > 0$). Let g_1, g_2 and g_3 be any boolean functions on F_2^s and ϕ_1, ϕ_2 and ϕ_3 any mappings from F_2^s to F_2^r such that for every element y in F_2^s, the vectors $\phi_1(y)$, $\phi_2(y)$, $\phi_3(y)$ and $\phi_1(y) \oplus \phi_2(y) \oplus \phi_3(y)$ have Hamming weights greater than t. Then the function:*

$$f(x, y) = [x \cdot \phi_1(y) \ \oplus \ g_1(y)] \, [x \cdot \phi_2(y) \ \oplus \ g_2(y)] \oplus$$

$$[x \cdot \phi_1(y) \oplus g_1(y)] \, [x \cdot \phi_3(y) \oplus g_3(y)] \oplus [x \cdot \phi_2(y) \oplus g_2(y)] \, [x \cdot \phi_3(y) \oplus g_3(y)]$$

is t-resilient.

Note that, according to Theorem 1 and because of the property of the Walsh transform of Maiorana-McFarland's functions recalled after Relation (3), if the sets $\phi_1(F_2^s)$, $\phi_2(F_2^s)$, and $\phi_3(F_2^s)$ are disjoint, then the nonlinearity of f is at least equal to the mean of:
- the nonlinearity of the Maiorana-McFarland's function $x \cdot \phi(y) \ \oplus \ g(y)$, where $\phi = \phi_1 + \phi_2 + \phi_3$ and $g = g_1 \oplus g_2 \oplus g_3$,
- the minimum of the nonlinearities of the functions $x \cdot \phi_i(y) \ \oplus \ g_i(y)$, $i = 1, 2, 3$. Hence, f can be nearly optimum with respect to Siegenthaler's and Sarkar et al.'s bounds; and its algebraic immunity may be higher than those of Maiorana-Mcfarland's nearly optimum functions.

We can also apply this property to the class of resilient functions derived from the $\mathcal{PS}_{ap}$ construction: Let n and m be two positive integers, g_1, g_2 and g_3 three functions from F_{2^m} to F_2, ϕ a linear mapping from F_2^n to F_{2^m} and a an element of F_{2^m} such that $a \oplus \phi(y) \neq 0$, $\forall y \in F_2^n$. Let b_1, b_2 and $b_3 \in F_2^n$ such that, for every z in F_{2^m}, $\phi^*(z) \oplus b_i$, $i = 1, 2, 3$ and $\phi^*(z) \oplus b_1 \oplus b_2 \oplus b_3$ have weight greater than t, where ϕ^* is the adjoint of ϕ, then the function

$$f(x, y) =$$

$$\left(g_1\left(\frac{x}{a \oplus \phi(y)}\right) \oplus b_1 \cdot y\right)\left(g_2\left(\frac{x}{a \oplus \phi(y)}\right) \oplus b_2 \cdot y\right) \oplus$$
$$\left(g_1\left(\frac{x}{a \oplus \phi(y)}\right) \oplus b_1 \cdot y\right)\left(g_3\left(\frac{x}{a \oplus \phi(y)}\right) \oplus b_3 \cdot y\right) \oplus$$
$$\left(g_2\left(\frac{x}{a \oplus \phi(y)}\right) \oplus b_2 \cdot y\right)\left(g_3\left(\frac{x}{a \oplus \phi(y)}\right) \oplus b_3 \cdot y\right)$$

is t-resilient. The complexity of the support of this function may permit getting a good algebraic immunity.

4.2 A generalization of Lemma 1

Proposition 3 can be generalized to more than 3 functions. This leads to further methods of constructions.

Proposition 3 *Let $f_1, \ldots, f_m$ be Boolean functions on F_2^n. For every positive integer l, let σ_l be the Boolean function defined by*

$$\sigma_l = \bigoplus_{1 \leq i_1 < \ldots < i_l \leq m} \prod_{j=1}^{l} f_{i_j} \quad \text{if } l \leq m \text{ and } \sigma_l = 0 \text{ otherwise.}$$

Then we have $f_1 + \ldots + f_m = \sum_{i \geq 0} 2^i \sigma_{2^i}$. Denoting by $\widehat{f}$ the Fourier transform of f, that is, $\widehat{f}(s) = \sum_{x \in F_2{}^n} f(x)(-1)^{x \cdot s}$, this implies $\widehat{f_1} + \ldots + \widehat{f_m} = \sum_{i \geq 0} 2^i \widehat{\sigma_{2^i}}$. Moreover, if $m+1$ is a power of 2, say $m+1 = 2^r$, then

$$\widehat{\chi_{f_1}} + \ldots + \widehat{\chi_{f_m}} = \sum_{i=0}^{r-1} 2^i \widehat{\chi_{\sigma_{2^i}}}. \tag{10}$$

Proof. Let x be any vector of F_2^n and $j = \sum_{k=1}^{m} f_k(x)$. According to Lucas' Theorem (cf. [16]), the binary expansion of j is $\sum_{i \geq 0} 2^i \left(\binom{j}{2^i} [\text{mod } 2]\right)$. It is a simple matter to check that $\binom{j}{2^i}$ [mod 2] $= \sigma_{2^i}(x)$. Thus, $f_1 + \ldots + f_m = \sum_{i \geq 0} 2^i \sigma_{2^i}$. This implies $\widehat{f_1} + \ldots + \widehat{f_m} = \sum_{i \geq 0} 2^i \widehat{\sigma_{2^i}}$.
The linearity of the Walsh transform with respect to the addition in **Z** implies then directly $\widehat{f_1} + \ldots + \widehat{f_m} = \sum_{i \geq 0} 2^i \widehat{\sigma_{2^i}}$.
If $m + 1 = 2^r$, then we have $m = \sum_{i=0}^{r-1} 2^i$. Thus, we deduce $\chi_{f_1} + \ldots + \chi_{f_m} = \sum_{i=0}^{r-1} 2^i \chi_{\sigma_{2^i}}$ from $f_1 + \ldots + f_m = \sum_{i=0}^{r-1} 2^i \sigma_{2^i}$. The linearity of the Walsh transform implies then relation (10). $\diamond$

Corollary 1 *Let n be any positive integer and k any non-negative integer such that $k \leq n$. Let $f_1, \ldots, f_7$ be seven k-th order correlation* *immune (resp. k-resilient) functions. Assume that the function* $\sigma_4 = \bigoplus_{1\le i_1<\ldots<i_4\le 7} \prod_{j=1}^{l} f_{i_j}$ *is k-th order correlation immune (resp. k-resilient). Then the function* $\sigma_1 = f_1 \oplus \ldots \oplus f_7$ *is k-th order correlation immune (resp. k-resilient) if and only if the function* $\sigma_2 = f_1 f_2 \oplus f_1 f_3 \oplus \ldots \oplus f_6 f_7$ *is k-th order correlation immune (resp. k-resilient).*

Proof. Relation (10) and the fact that for every (non-zero) vector a of weight at most k we have $\widehat{\chi_{f_i}}(a) = 0$ for $i = 1, \ldots, 7$ and $\widehat{\chi_{\sigma_4}}(a) = 0$ imply that $\widehat{\chi_{\sigma_1}}(a) = 0$ if and only if $\widehat{\chi_{\sigma_2}}(a) = 0$. $\diamond$