On Correlation-immune functions [1]

P. Camion, C. Carlet, P. Charpin & N. Sendrier

INRIA, Domaine de Voluceau, Rocquencourt,
BP 105, 78153 Le Chesnay Cedex, FRANCE
camion@hades.inria.fr

Abstract

We establish the link between *correlation-immune* functions and *orthogonal arrays*. We give a recursive definition of any correlation-immune function of maximal degree. We describe the set of quadratic balanced correlation-immune functions of maximal order. Some constructions are then deduced.

1 Introduction

In a general type of running-key generator, the output sequences of m Linear Feedback Shift Registers are taken as arguments of a single non linear combining function f. If the function f is not properly chosen, it can happen that the generator structure is not resistant to a *correlation attack*: there is a statistical dependence between any small subset of the m subgenerator sequences and the keystream sequence (cf. an example in [8], p. 116).

A function f which provides an immunity to a correlation attack is called *a correlation-immune function*. The kth-order correlation-immune functions (denoted k-CI functions) were introduced by T. Siegenthaler in [11]. X. Guo-zhen and J.L. Massey later gave an equivalent definition of the k-CI functions, using the Walsh transform of the boolean functions. It is their definition, recalled in Section 2, which is used in the present paper.

We wish to show that Algebraic Coding Theory provides an alternative point of view for the concept of correlation-immunity. We present two new definitions of the k-CI functions, related to coding theory, and deduce some constructions.

In Section 3 we point out that a k-CI function is an *orthogonal array of strength k*. We later give a recursive definition of any k-CI function of maximal degree. Using algebraic properties - which are in fact properties of Reed and Muller codes (RM-codes) and subcodes of RM-codes - we show that the recursive definition permits to obtain explicitly some k-CI functions.

In Section 4 we present some constructions. Using the recursive definition, we describe

[1] Research supported by CNET-FRANCE, n.885B016 007909245/PAA

J. Feigenbaum (Ed.): Advances in Cryptology - CRYPTO '91, LNCS 576, pp. 86-100, 1992.

a large class of 1-CI functions of maximal degree. We after give a full description of the set of the quadratric balanced correlation-immune function of maximal order. In the last paragraph we propose, in fact, an algorithm producing some balanced correlation-immune functions of maximal order.
The present paper is a shortened version of the scientific report [3]; the reader can find in [3] more explanations and examples.

2 Correlation-immune functions

Let $\mathbf{F} = GF(2)$ and $\mathbf{G} = \mathbf{F}^m$. An element x of $\mathbf{G}$ is an m-tuple $(x_1, \ldots, x_m)$ over $\mathbf{F}$. Let $x \in \mathbf{G}$ and $\lambda \in \mathbf{G}$, and define their dot product as: $x \cdot \lambda = x_1\lambda_1 + \ldots + x_m\lambda_m \in \mathbf{F}$.
Let f be a boolean function of m binary variables. The *Walsh transform* of $f(x)$ is the real-valued function over $\mathbf{G}$:

$$F(\lambda) = \sum_{x \in \mathbf{G}} f(x)\,(-1)^{x \cdot \lambda}\,. \tag{1}$$

The set of the elements $x \in \mathbf{G}$ such that $f(x) = 1$ is a binary array $M \times m$, where M is the weight of the value of f. This array is the *truth-table* of f.
In this paper, the weight of a binary vector u is always *the Hamming weight*, ie the number of nonzero components in u, and is denoted by $W(u)$.

Definition 2.1 [7] *Let $k \in [1, m-1]$. The function f is kth-order correlation immune (ie is a k-CI function) if and only if its Walsh transform satisfies:*

$$F(\lambda) = 0, \quad for \;\; 1 \le W(\lambda) \le k\,, \tag{2}$$

where $W(\lambda)$ denotes the Hamming weight of the binary m-tuple λ.

In the following, we denote by $v(f)$ the binary vector $\{f(x) \mid x \in \mathbf{G}\}$ and we say that $v(f)$ is *the value of f*. In general we shall suppose that the value of a k-CI function f is *balanced*, ie that $F(0) = 2^{m-1}$; we shall say that *the function f is balanced.*

Proposition 2.1 [11] *Let f be a k-CI function. Let $d(f)$ be the degree of f. Then $d(f) \le m-k$. Moreover if f is balanced then $d(f) < m-k$ unless $k = m-1$.*

Hence if f is a $(m-2)$-CI function and is balanced , f is an affine function. The only possible $(m-1)$-CI functions are [11]: $f(x) = x_1 + \ldots + x_m + c \quad , \; c \in \mathbf{F}$.
In [11], T. Siegenthaler showed how to construct by iteration a limited family of k-CI functions : a k-CI function is obtained from two linear functions of $m-(k+1)$ variables.

3 Others definitions

3.1 Orthogonal arrays

The characterization given by X. GUO-ZEN and J.L. MASSEY [7] for correlation-immune functions, concept introduced by T. SIEGENTHALER [11] was taken as Definition 2.1. That characterization corresponds precisely to the one by P. DELSARTE of orthogonal arrays, concept introduced by C.R. RAO [10] - as we point out in Theorem 3.1 -.

Definition 3.1 [6, Ch. 11] *An $M \times m$ matrix V with entries from a set of q elements is called an orthogonal array of size M, m constraints, q levels, strength k, and index μ if any set of k columns of V contains all q^k possible row vectors exactly μ times. Such an array is denoted by (M, m, q, k) . Clearly $M = \mu q^k$.*

Theorem 3.1 *A boolean function f on* G *is correlation immune of order k if and only if its truth table is an orthogonal array $(M, m, 2, k)$.*

Proof: Let f be a boolean function of m variables. Let M be the weight of $v(f)$; let T be the truth-table of f.
1. Suppose that T is an orthogonal array $(M, m, 2, k)$; let $\mu = 2^{-k}M$. Let $\lambda \in \mathbf{G}$ such that $W(\lambda) = k$. Then we have:

$$\begin{aligned} F(\lambda) &= \sum_{x \in \mathbf{G}} f(x)\,(-1)^{x.\lambda} = \mu \sum_{y \in \mathbf{F}^k} (-1)^{W(y)} \\ &= \mu\,(\,|\,\{\, y \in \mathbf{F}^k \;;\; W(y) \text{ is even } \}\,| - |\,\{\, y \in \mathbf{F}^k \;;\; W(y) \text{ is odd } \}\,|\,) = 0 \ . \end{aligned}$$

It is clear that an orthogonal array $(M, m, 2, k)$ is also an orthogonal array $(M, m, 2, i)$ for $i \in [1, k]$. Hence $F(\lambda) = 0$ for all λ such that $W(\lambda) \in [1, k]$. So f satisfies (2).
2. Suppose that f is a k-CI function. Let $\lambda \in \mathbf{G}$ such that $W(\lambda) = k$. Define the projection $\pi : \mathbf{G} \longmapsto \mathbf{F}^k$ as

$$\pi(x) = (x_i)_{i \in I} \quad , \quad I = \{\, i \in [1, m] \mid \lambda_i = 1 \,\} \ ,$$

and $\eta : \mathbf{F}^k \longmapsto \mathbf{Z}$ as

$$\eta(g) = |\, T_g \,| \quad , \quad T_g = \{\, x \in \mathbf{G} \mid f(x) = 1, \pi(x) = g \,\} \ .$$

We define the *support* $s(x)$ of any element $x \in \mathbf{F}^m$ by : $s(x) = \{\, i \mid x_i \neq 0 \,\}$. We now use the fact that T is an orthogonal array $(M, m, 2, k)$ if and only if for any $\lambda \in \mathbf{G}$ such that $W(\lambda) = k$, then the value $\eta(g)$ defined above is $\eta(g) = 2^{-k}M$, for all $g \in \mathbf{F}^k$. For any $\lambda' \in \mathbf{G}$ with $s(\lambda') \subset s(\lambda)$, we have:

$$F(\lambda') = \sum_{x \in T} (-1)^{\lambda' . x} = \sum_{g \in \mathbf{F}^k} \sum_{x \in T_g} (-1)^{\lambda' . x} = \sum_{g \in \mathbf{F}^k} \eta(g) \, (-1)^{\pi(\lambda') . g} \ .$$

On the other hand, denote by H the abelian group $(\mathbf{F}^k, +)$ and consider the abelian group algebra $\mathbf{C}H$. An element $z \in \mathbf{C}H$ is denoted by $\sum_{g \in H} z_g X^g$. The *Walsh-Fourier transform* $\hat{z}$ of z is given by

$$\hat{z}_h = \sum_{g \in H} z_g (-1)^{h . g}, \quad \text{for all } h \in H \ .$$

Thus for $z = \sum_{g \in H} \eta(g) X^g$, we observe that for every λ' such that $s(\lambda') \subset s(\lambda)$, then if $\pi(\lambda') = h$, we have that $F(\lambda') = \hat{z}_h$ and therefore $\hat{z}_h = 0$ for every nonzero h. Hence inversing the Fourier transform $z \rightarrow \hat{z}$, we obtain:

$$\eta(g) = 2^{-k} \sum_{g \in H} \hat{z}_h (-1)^{h . g} = 2^{-k} \hat{z}_0 = 2^{-k} \, M \quad . \, \square$$

Example 3.1 $m = 4$, $M = 2^{m-1} = 8$, $f(x) = x_1 + x_3 + x_4$. *Let V be the truth table of f; we study the transposed array :*

$$\tilde{V} = \begin{array}{c c} \left[\begin{array}{cccccccc} 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{array} \right] & \begin{array}{c} x_1 \\ x_2 \\ x_3 \\ x_4 \end{array} \end{array}$$

Each line of $\tilde{V}$ contains 4 times "0" and 4 times "1". Thus f is an 1-CI function. Moreover any set of two lines of $\tilde{V}$ contains all 2-dimensional vectors exactly twice. So f is an 2-CI function.

3.2 A recursive definition

The correlation-immune functions of maximal degree (for a fixed order), the value of which is balanced, are more interesting in applications; for this reason, we have chosen to present our results with these hypotheses. From now on we only consider balanced correlation-immune functions.

Definition 3.2 *A balanced k-CI function of m variables is said to be a $ci(k, m)$ function; such a function is said to have the maximal degree if and only if $d(f) = m - (k+1)$. By convention a $ci(0, m)$ function is a balanced function the degree of which equals $m - 1$.*

For any m, we denote by $\mathcal{F}^m$ the set of boolean functions of m variables $x_1, \ldots, x_m$. Let $f \in \mathcal{F}^m$ with degree $\leq m-1$. Using the polynomial form of f, it is easy to prove that, after possibly permuting the indices, then f can always be written as follows:

$$f = (x_m + 1)\, f_1 \; + \; x_m\, f_2 \qquad \begin{cases} f_1 \textit{ and } f_2 \in \mathcal{F}^{m-1} \\ d(f_1) = d(f_2) = d(f) \\ d(f_1 + f_2) < d(f) \end{cases} \tag{3}$$

Theorem 3.2 *Let f_1 and f_2 be functions derived from f by (3). Let F_1. and F_2 be respectively the Walsh transforms of f_1 and f_2. Then f is a $ci(k,m)$ function if and only if:*

(i) *f_1 and f_2 are $ci(k-1, m-1)$ functions*

(ii) *For all $\lambda' \in \mathbf{F}^{m-1}$ with $W(\lambda') = k$ we have:*

$$F_1(\lambda') \; + \; F_2(\lambda') = 0 \quad . \tag{4}$$

Moreover f has maximal degree if and only if f_1 and f_2 have maximal degree.

Proof: Let $\lambda \in \mathbf{F}^m$, $\lambda = (\lambda', \epsilon)$ with $\epsilon \in \mathbf{F}$ and $\lambda' \in \mathbf{F}^{m-1}$. Let $x' = (x_1, \ldots, x_{m-1})$. Then

$$F(\lambda) = \sum_{x \in \mathbf{G}} f(x)(-1)^{x'.\lambda' + x_m.\epsilon} = \sum_{x_m = 0} f_1(x')(-1)^{x'.\lambda'} \; + \; \sum_{x_m=1} f_2(x')(-1)^{x'.\lambda'+\epsilon};$$

$$F(\lambda) = F_1(\lambda') \; + \; (-1)^{\epsilon}\, F_2(\lambda') \quad . \tag{5}$$

1. Suppose that f satisfies (i) and (ii). In accordance with (5), we have:

$$F(0) = F_1(0) \; + \; F_2(0) = 2^{m-1} \quad ;$$

if λ is such that $0 < W(\lambda') < k$ then $F(\lambda) = 0$, from (i); if λ is such that $W(\lambda') = k$ and $\epsilon = 0$ then $F(\lambda) = 0$, from (ii). So f is a $ci(m,k)$ function.

2. Suppose now that f is a $ci(k,m)$ function. Then for all $\lambda = (\lambda', \epsilon)$ such that $W(\lambda) \in [1, k]$, formula (5) yields:

$$0 = F_1(\lambda') \; + \; (-1)^{\epsilon}\, F_2(\lambda') \quad . \tag{6}$$

For $\lambda = (\mathbf{0}, 1)$, we obtain $F_1(0) = F_2(0)$. Then f_1 and f_2 are balanced.
If $0 < W(\lambda') < k$, we obtain: $F_1(\lambda') = F_2(\lambda')$ for $\epsilon = 1$ and $F_1(\lambda') = -F_2(\lambda')$ for $\epsilon = 0$. Then $F_1(\lambda') = F_2(\lambda') = 0$ - ie (i) is satisfied. If $W(\lambda') = k$ and $\epsilon = 0$, we obtain (ii) immediately from (6).

3. A $ci(k,m)$ function has maximal degree $m-(k+1)$. A $ci(k-1, m-1)$ function has maximal degree $(m-1)-k$. Since by definition $d(f)$ equals $d(f_1)$ and $d(f_2)$, then f has maximal degree if and only if f_1 and f_2 have maximal degree. □

REMARK : Theorem 3.2 means that a $ci(k,m)$ function, has its truth-table T in the following form ($\tilde{T}$ is the transposed T):

$$\tilde{T} \ = \ \begin{array}{|c|c|}\hline \tilde{T}_1 & \tilde{T}_2 \\ \hline 0\,0\ \ldots 0\,0 & 1\,1\ \ldots 1\,1 \\ \hline \end{array} \quad \text{where}$$

- T_1 and T_2 are orthogonal arrays $(2^{m-2}, m-1, 2, k-1)$,
- let $\mu = 2^{m-k-1}$; let $u \in \mathbf{F}^k$ and a set of k rows containing a times u in $\tilde{T}_1$ and b times u in $\tilde{T}_2$; then $a+b=\mu$.

3.3 Correlation-immune functions and Reed-Muller codes

Recall that $\mathbf{G} = \mathbf{F}^m$. The Reed-Muller code of length 2^m and order r, denoted by $R(r,m)$, can be identified with the set of the boolean functions of m variables and of degree $\leq r$. The codewords are the values of the functions. Let $f \in \mathcal{F}^m$; the order of the correlation-immunity of f is obtained in studying the weights of the coset $f + R(1,m)$: for each $\lambda \in \mathbf{G}$, the function $h_\lambda \ : \ x \in \mathbf{G} \longmapsto x.\lambda$ is an element of $R(1,m)$; then

$$F(\lambda) = 0 \quad \Longleftrightarrow \quad W(\ v(f) + v(h_\lambda)\ ) = 2^{m-1} \quad . \tag{7}$$

Thus f is $ci(k,m)$ if and only if for all λ, with $W(\lambda) < k$, then $f + h_\lambda$ has weight 2^{m-1}. The writing of f as in (3) means that f_1 and f_2 are in $R(d(f), m-1)$ and in a same coset of the code $R(d(f)-1, m-1)$. It appears in Theorem 3.2 that it will be interesting to know the codewords g of weight 2^{m-2} of a coset $f_1 + R(d(f)-1, m-1)$ and, for such g's the weights of the codewords of the coset $g + R(1,m)$. So it seems difficult to obtain the overall description of the set of the $ci(k,m)$ functions, because this problem is related with open problems on Reed-Muller codes. However some well-known properties may be used:

1. In Corollaries 4.1 and 4.2 we use some transformations which preserve a given coset and carry a balanced word in another balanced word.

2. We are able to easily construct $ci(1,m)$ functions, because the set of the $ci(0,m')$ functions is well-known for any m'. So we can prove in the following Section that Theorem 3.2 expresses a constructive definition of k-CI functions with maximal degree.

A $ci(0,m')$ function is a function of degree $m'-1$ and of m' variables, the value of which has weight $2^{m'-1}$. The class of the functions $ci(0,m')$ was studied in [2] and [5];

it is very simple to construct such a function : let $V(g)$ be the truth-table of a function g of m' variables and of weight $2^{m'-1}$. Then the degree of g is $m'-1$ if and only if there exists an hyperplane H of $\mathbf{F}^{m'}$ such that the size of the set $V(g) \cap H$ is odd (see the construction of f_1 in Example 4.1). The number N of such functions g can be calculated with formulae given in [2]; for small values of m', we obtain:

$m' = 3 \Rightarrow N = 56$; $m' = 4 \Rightarrow N = 12000$;
$m' = 5 \Rightarrow N = 582\ 284\ 160$;
$m' = 6$, $\Rightarrow N = 1.803.989.388.148.674.048$.

4 Construction of correlation-immune functions

4.1 Extending an orthogonal array to a stronger one

In accordance with Theorem 3.1 and 3.2, we shall construct a k-CI function by extending the truth-table of a $(k-1)$-CI function to an orthogonal array of strength k. So we define two simple applications on $\mathcal{F}^m$ which preserves the zeros of the Walsh transform.

Proposition 4.1 *Let $f \in \mathcal{F}^m$ and denote by ν the weight of the value of f; let us define the applications from $\mathcal{F}^m$ onto itself:*

$$\Lambda \ : \ f \longmapsto 1 + f \ ; \tag{8}$$

for $a \in \mathbf{G}$ and $\tau_a \ : \ x \in \mathbf{G} \ \rightarrow \ x + a$,

$$\Omega_a \ : \ f \longmapsto f \circ \tau_a \quad . \tag{9}$$

Let F, F' and F'' be the Walsh transforms of f, $\Lambda(f)$ and $\Omega_a(f)$ respectively. Then for all λ in $\mathbf{G}$, $\lambda \neq 0$, we have:

(i) $F'(\lambda) = -F(\lambda)$ (ii) $F''(\lambda) = (-1)^{\lambda.a}\ F(\lambda)$.

Moreover $F'(0) = 2^m - \nu$ *and* $F''(0) = F(0) = \nu$.

Proof: By definition we have $F'(\lambda)+F(\lambda) = \sum_{x \in \mathbf{G} f(x)=0} (-1)^{x.\lambda} + \sum_{x \in \mathbf{G} f(x)=1} (-1)^{x.\lambda}$; this sum equals 0 if $\lambda \neq 0$ and equals 2^m otherwise. Formula (ii) and the value of $F''(0)$ become from:

$$F''(\lambda) = \sum_{x \in \mathbf{G}} f(x+a)\ (-1)^{(x+a).\lambda + a.\lambda} = (-1)^{\lambda.a}\ F(\lambda)\ . \ \square$$

Then from (i) of Proposition 4.1 and from Theorem 3.2 we immediatly have:

Corollary 4.1 *Let $f_1 \in \mathcal{F}^{m-1}$. Then the function of m variables $(x_1, \dots, x_m)$*

$$f = (x_m + 1)\ f_1\ +\ x_m\ \Lambda(f_1) = f_1\ +\ x_m \qquad (10)$$

is a $ci(k,m)$ function if and only if f_1 is a $ci(k-1, m-1)$ function; moreover f has maximal degree if and only if f_1 has maximal degree.

Corollary 4.2 *Let k be an odd integer. Let $a \in \mathbf{F}^{m-1}$, $a = (1, 1, \dots, 1)$; τ_a and Ω_a are denoted by τ_1 and Ω_1. Let $f_1 \in \mathcal{F}^{m-1}$. Then the function of m variables $(x_1, \dots, x_m)$*

$$f = (x_m + 1)\ f_1\ +\ x_m\ \Omega_1(f_1) = f_1\ +\ x_m\ (f_1 + f_1 \circ \tau_1) \qquad (11)$$

is a $ci(k,m)$ function if and only if f_1 is a $ci(k-1, m-1)$ function; moreover f has maximal degree if and only if f_1 has maximal degree.

Proof: Since $a = (1, 1, \dots, 1)$, formula (ii) of Proposition 4.1 becomes $F''(\lambda) = (-1)^{W(\lambda)} F(\lambda$ When k is odd we can apply Theorem 3.2 with $f_2 = \Omega_1(f_1)$.□

REMARK : If k is even we can also apply Theorem 3.2 with $f_2 = \Lambda(\Omega_1(f_1))$; indeed calculating the Walsh transform of f_2 with F'' given above, we then see by (i) of Proposition 4.1, that formula (ii) of Theorem 3.2 is satisfied.

REMARK : Starting from functions given by (10) and (11), the construction of SIEGENTHALER permits to obtain other correlation-immune functions. Indeed his algorithm is based on this result: *let f be a boolean function defined by (3); if f_1 and f_2 are $ci(k, m-1)$ functions, then f is a $ci(k,m)$ function.* Note that the order is not increased in that construction.

REMARK : In fact, Corollary 4.1 is obvious. It consists in the addition of a variable; that is an addition of a binary symmetric channel having capacity zero [11].

Example 4.1 A construction of a $ci(k,m)$ functions using Corollaries 4.1 and 4.2. *We explained in Section 3.3 how we can construct a $ci(0, m-1)$ function; from Corollaries 4.1 and 4.2 that yields the construction of at least one $ci(1,m)$ function. First assume that $m = 5$ and $f_1 \in \mathcal{F}^4$. Let H be the hyperplane generated by the basis $\{x_1, x_2, x_3\}$. The following array shows the values of f_1, $\Lambda(f_1)$ and $\Omega_1(f_1)$. The value of f_1 has weight* 8. *The transposed truth-table $\tilde{V}$ of f_1 is such that the set $\tilde{V} \cap H$ has* 5 *elements. Then f_1 is a $ci(0,4)$ function - ie a balanced function of degree 3 -.*

$$\begin{array}{c|cccccccccccccccc}
x_1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\
x_2 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\
x_3 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
x_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
f_1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
\Lambda(f_1) & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\
\Omega_1(f_1) & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0
\end{array}$$

According to Corollary 4.2, the function $f' = f_1 + x_5(f_1 + \Omega_1(f_1))$ is a $ci(1,5)$ function with degree 3. The polynomial forms of f_1 and f' are

$$\begin{aligned} f_1(x_1,\ldots,x_4) &= x_1+x_2+x_3+x_1x_2+x_1x_4+x_2x_4+x_3x_4 \\ &\quad +x_1x_2x_3+x_1x_2x_4+x_1x_3x_4+x_2x_3x_4 \ . \end{aligned}$$

$$\begin{aligned} f'(x_1,\ldots,x_5) &= f_1(x_1,\ldots,x_4)+x_5(f_1(x_1,\ldots,x_4)+f_1(x_1+1,\ldots,x_4+1)) \\ &= x_1+x_2+x_3+x_5+x_1x_2+x_1x_4+x_1x_5+x_2x_4+x_2x_5 \\ &\quad +x_3x_4+x_1x_2x_3+x_1x_2x_4+x_1x_3x_4+x_2x_3x_4 \ . \end{aligned}$$

The truth-table T of f' is an orthogonal array $(16,5,2,1)$:

$$\tilde{T} = \begin{array}{c|cccccccccccccccc}
x_1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\
x_2 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\
x_3 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
x_4 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\
x_5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}$$

4.2 Quadratic balanced correlation-immune functions of maximal order

Let f be a boolean function on G, and define $\forall\ \lambda \in G\ :\ f_\lambda(x) = f(x) + x.\lambda$. According to Definition 3.2, the function f is a $ci(k,m)$ function, $k > 0$, if and only if for any λ such that $0 < W(\lambda) \le k$ the function f_λ is balanced. Indeed, for any $\lambda \neq 0$, we have:

$$\sum_{x\in G}(-1)^{f(x)+x.\lambda} = \sum_{x\in G}(-1)^{x.\lambda} - 2\sum_{x\in G} f(x)(-1)^{x.\lambda} = -2F(\lambda) \ ,$$

where $\sum_{x\in G}(-1)^{f_\lambda(x)}$ equals zero if and only if f_λ is balanced. We say that f is *quadratic if and only if its degree equals exactly* 2 (i.e. $f \in R(2,m)\backslash R(1,m)$). If f is quadratic, then f_λ is also quadratic. Then we can determine whether f_λ is balanced or not (for instance see [4]):

Lemma 4.1 *Let g be a quadratic boolean function. Let the symplectic form associated with g:*

$$\phi_g \ : \ (x,y) \ \longmapsto \ g(0)+g(x)+g(y)+g(x+y) \ ,$$

Recall that the kernel of ϕ_g is the subspace of G: $E_g = \{ \ x \in G \mid \forall y \in G, \ \phi_g(x,y)=0 \ \}$ (of even codimension). Then g is balanced if and only if its restriction to E_g is not constant.

Theorem 4.1 *A $ci(m-3,m)$ quadratic function takes one of the polynomial forms given in (12), (13), (14) or (15).*
Moreover, we can obtain all the $ci(m-3,m)$ quadratic functions by applying several times Corollaries 4.1 and 4.2 to $ci(0,3)$ quadratic functions.

Proof: For any function f of degree 2 on G, define: $A_f = \{ \ \lambda \in G \mid f_\lambda$ is not balanced $\}$. By definition, the immunity order of f is equal to the smallest weight of the elements of A_f minus 1. In accordance with Lemma 4.1, A_f is the set of all $\lambda \in G$ such that f_λ is constant on E_f, or, equivalently:

$$A_f = \{ \ \lambda \in G \mid \forall x, \ x \in E_f \ \ : \ f(x)+f(0) = \lambda.x \ \} \ .$$

The function $f+f(0)$ is linear on E_f, and is therefore the restriction to E_f of at least one linear form on G. As the linear forms on G all are of the type $x \longrightarrow \lambda.x$, A_f has at least one element. Let $\lambda_0 \in A_f$; then we have: $A_f = \{ \ \lambda \in G \mid \forall x \in E_f, \ \lambda.x = \lambda_0.x \ \}$. We denote by $E_f^\perp$ the linear space: $E_f^\perp = \{ \ \lambda \in G \mid \forall x \in E_f, \ \lambda.x = 0\}$. Then λ is in A_f if and only if $\lambda+\lambda_0$ is in $E_f^\perp$; A_f is an affine subspace of G, of direction $E_f^\perp$, and therefore of even dimension. We will now determine such subspaces, and deduce the corresponding functions f.

Let A be an affine subspace of G, of even dimension, and whose elements have weights at least equal to $m-2$. Let $\{e_1,\ldots,e_m\}$ be the natural basis of G, and the space $A' = A+e_0$, where $e_0 = e_1+\ldots+e_m$. So A' is an affine-subspace, of even dimension, whose elements have weights at most equal to 2; it is clear that we can determine equivalently A or A'. If A' contains 0, then it is the linear space equal to:

$$A^{(1)} = \{0,e_i,e_j,e_i+e_j\} \quad \text{or} \quad A^{(2)} = \{0,e_i+e_j,e_i+e_k,e_j+e_k\} \ ,$$

where i, j and k are distincts elements of $[1,m]$.
If A' contains at last one element of weight 1, then it is equal to:

$$A^{(3)} = \{e_i,e_j,e_i+e_k,e_j+e_k\} \ .$$

If A' contains only elements of weight 2, then it is equal to:

$$A^{(4)} = \{e_i+e_j,e_i+e_k,e_l+e_j,e_l+e_k\} \ ,$$

where i, j, k and l are distincts elements of $[1,m]$.

Now we are able to examine the four possible definitions of A_f:

$$\begin{array}{llll}\text{(i)} & A_f = A^{(1)} + e_0 & \text{(ii)} & A_f = A^{(2)} + e_0 \\ \text{(iii)} & A_f = A^{(3)} + e_0 & \text{(iv)} & A_f = A^{(4)} + e_0 \ .\end{array}$$

(i) $A_f = \{e_0, e_0 + e_i, e_0 + e_j, e_0 + e_i + e_j\}$. Since A_f is a coset of $E_f^\perp$, we have:

$$E_f^\perp = \{0, e_i, e_j, e_i + e_j\} \quad \text{and then} \quad E_f = \{x \in G \mid x_i = x_j = 0\}.$$

Let the function $g \ : \ x \in G \ \rightarrow \ x_i x_j$. Clearly $E_g = E_f$. There exists only one symplectic form admitting E_f as kernel, since E_f has codimension 2 and since there exists only one non-zero symplectic form on a linear space of dimension 2. So f belongs to the same coset of the code $R(1,m)$ as g. Hence there exists $\lambda \in G$ and $\epsilon \in \mathbf{F}$ such that: $f(x) = g(x) + \lambda.x + \epsilon$. Since g is not balanced, then λ must be in A_f. So we obtain the algebraic normal form of f:

$$f = x_i x_j + \sum_{t \in [1,m]-\{i,j\}} x_t + \epsilon_i x_i + \epsilon_j x_j + \epsilon \ , \ \epsilon_i \in \mathbf{F}, \ \epsilon_j \in \mathbf{F}, \ \epsilon \in \mathbf{F} \ . \tag{12}$$

For m=3 the following functions are clearly balanced:

$$x_1x_2 + x_1 + x_2 + x_3 \ , \ x_1x_2 + x_1 + x_3 \ , \ x_1x_2 + x_2 + x_3 \ , \ x_1x_2 + x_3 \ .$$

Indeed their expressions all contain a linear function which is linearly independant from x_1 and x_2 and so they are $ci(0,3)$ functions. By using Corollary 4.1, $m-3$ times, we obtain the following $ci(m-3,m)$ functions:

$$\begin{array}{cc} x_1x_2 + x_1 + x_2 + x_3 + \ldots + x_m & x_1x_2 + x_1 + x_3 + \ldots + x_m \\ x_1x_2 + x_2 + x_3 + \ldots + x_m & x_1x_2 + x_3 + \ldots + x_m \ . \end{array}$$

By permuting the variables we can check that all $ci(m-3,m)$ functions described by (12) are then obtained.

(ii) $A_f = \{e_0, e_0 + e_i + e_j, e_0 + e_i + e_k, e_0 + e_j + e_k\}$. Then

$$E_f^\perp = \{0, e_i + e_j, e_i + e_k, e_j + e_k\} \quad , \quad E_f = \{x \in G \mid x_i + x_j = x_i + x_k = 0\},$$

$$\text{and} \quad f(x) = (x_i + x_j)(x_i + x_k) + \lambda.x + \epsilon \quad \text{with} \quad \lambda \in A_f \ , \ \text{where}$$

$$\lambda.x = \sum_{t=1}^{m} x_t \ \text{ or } \sum_{t \in [1,m]-\{i,j\}} x_t \ \text{ or } \sum_{t \in [1,m]-\{i,k\}} x_t \ \text{ or } \sum_{t \in [1,m]-\{j,k\}} x_t . \tag{13}$$

Now the following functions are balanced, since their expressions all contain a linear function which is linearly independant from $x_1 + x_2$ and $x_1 + x_3$; so they are $ci(0,3)$ functions:

$$(x_1+x_2)(x_1+x_3)+x_1+x_2+x_3 \quad , \quad (x_1+x_2)(x_1+x_3)+x_2+x_3 \ ,$$
$$(x_1+x_2)(x_1+x_3)+x_1+x_3 \quad , \quad (x_1+x_2)(x_1+x_3)+x_1+x_2 \ .$$

We then obtain all the functions described by (13), for the same reasons as in (i).

(iii) $A_f = \{e_0+e_i, e_0+e_j, e_0+e_i+e_k, e_0+e_j+e_k\}$. Then

$$E_f^\perp = \{0, e_i+e_j, e_k, e_i+e_j+e_k\} \quad , \quad E_f = \{x \in G \mid x_i + x_j = x_k = 0\},$$

$$\text{and} \quad f(x) = (x_i+x_j)x_k + \lambda.x + \epsilon \quad \text{with} \quad \lambda \in A_f \ , \ \text{where}$$

$$\lambda.x = \sum_{t\in[1,m]-\{i\}} x_t \ \text{ or } \sum_{t\in[1,m]-\{j\}} x_t \ \text{ or } \sum_{t\in[1,m]-\{i,k\}} x_t \ \text{ or } \sum_{t\in[1,m]-\{j,k\}} x_t. \qquad (14)$$

Now the following functions are balanced, since their expressions all contain a linear function which is linearly independant from x_1+x_2 and x_3; so they are $ci(0,3)$ functions:

$$(x_1+x_2)(x_1+x_3)+x_1+x_2+x_3 \quad , \quad (x_1+x_2)(x_1+x_3)+x_2+x_3 \ ,$$
$$(x_1+x_2)(x_1+x_3)+x_1+x_3 \quad , \quad (x_1+x_2)(x_1+x_3)+x_1+x_2 \ .$$

We then obtain all the functions described by (14), for the same reasons as in (i).

(iv) $A_f = \{e_0+e_i+e_j, e_0+e_i+e_k, e_0+e_l+e_j, e_0+e_l+e_k\}$. Then

$$E_f^\perp = \{0, e_j+e_k, e_i+e_l, e_i+e_j+e_k+e_l\} \quad , \quad E_f = \{x \in G \mid x_j + x_k = x_i + x_l = 0\},$$

$$\text{and} \quad f(x) = (x_j+x_k)(x_i+x_l) + \lambda.x + \epsilon \quad \text{with} \quad \lambda \in A_f \ , \ \text{where}$$

$$\lambda.x = \sum_{t\in[1,m]-\{i,j\}} x_t \ \text{ or } \sum_{t\in[1,m]-\{i,k\}} x_t \ \text{ or } \sum_{t\in[1,m]-\{l,j\}} x_t \ \text{ or } \sum_{t\in[1,m]-\{l,k\}} x_t. \qquad (15)$$

Let us consider once again the $ci(0,3)$ functions using in **(iii)**. Using Corollary 4.2 with $m = 4$, we obtain the $ci(1,4)$ functions:

$$(x_1+x_2)(x_3+x_4)+x_2+x_3 \quad , \quad (x_1+x_2)(x_3+x_4)+x_1+x_3 \ ,$$
$$(x_1+x_2)(x_3+x_4)+x_2 \quad , \quad (x_1+x_2)(x_3+x_4)+x_1 \ .$$

So we **can use** Corollary 4.1 and obtain all the functions described by (15). □

Remark : It is easy to check that all the $ci(0,3)$ functions we use in the proof, are equivalent; that is natural since their symplectic forms have the same rank and since they all are balanced.

In accordance with (12), (13), (14) and (15), we can state the number of the $ci(m-3,m)$ quadratic functions:

Corollary 4.3 *Let Q_m be the number of $ci(m-3, m)$ quadratic functions. Then:*

$$Q_m = \frac{1}{3}m(m-1)(3m-2)(m+1) \ . \tag{16}$$

Proof: Counting the functions described by (12), (13), (14) and (15), we obtain respectively:

$$8\binom{m}{2} \ , \ 8\binom{m}{3} \ , \ 8\binom{m}{2}(m-2) \ \text{ and } \ 4\binom{m}{2}\binom{m-2}{2} \ . \ \square$$

Corollary 4.4 *If an orthogonal array $(2^{m-1}, m, 2, m-3)$ (with index 4) is the truth table of a quadratic boolean function, that function is given by (12), (13), (14) or (15). The number of such orthogonal arrays is given by (16).*

There are $NQ_m = (2^u - 1)2^{m+1}$, $u = \binom{m}{2}$, quadratic functions. One can see, with the following array, that few of them are $ci(m-3, m)$ quadratic functions. We denote by BQ_m the number of balanced quadratic functions.

m	NQ_m	BQ_m	Q_m
3	112	56	56
4	2016	840	200
5	65472	36456	520
6	4194176	1828008	1120
7	536870656	300503336	2128
8	137438952960	60273666600	3696

4.3 More balanced correlation-immune functions of maximal orders

Proposition 4.2 *Let $r \in [1, m[$, g a boolean function on $\mathbf{F}^{m-r}$, and ϕ a mapping from $\mathbf{F}^{m-r}$ to $\mathbf{F}^r$. Let f be the boolean function f such defined:*

$$\mathbf{D} = \mathbf{F}^r \times \mathbf{F}^{m-r} \ , \ \ \forall \ (x, y) \in \mathbf{D} \ : \ f(x,y) = < x, \phi(y) >_r + g(y) \ , \tag{17}$$

where $<, >_r$ denotes the usual dots product on $\mathbf{F}^r$. Then f is a $ci(k, m)$ function, with

$$k \geq \ inf \ \{ \ W(\phi(y)) \mid y \in \mathbf{F}^{m-r} \ \} \ - \ 1 \ .$$

Proof: Let $a \in \mathbf{F}^r$ and $b \in \mathbf{F}^{m-r}$. We have

$$L = \sum_{(x,y)\in\mathbf{D}} (-1)^{f(x,y)+<a,x>_r+<b,y>_{m-r}} = \sum_{y\in\mathbf{F}^{m-r}} (-1)^{g(y)+<b,y>_{m-r}} \sum_{x\in\mathbf{F}^r} (-1)^{<a+\phi(y),x>_r} \ .$$

The function $x \rightarrow < a + \phi(y), x >_r$ is a linear form on $\mathbf{F^r}$, and therefore is either null or balanced. Then

$$L = 2^r \sum_{y \in \phi^{-1}(a)} (-1)^{g(y) + < b, y >_{m-r}} \; .$$

Hence: $a \notin \phi(\mathrm{F}^{m-r}) \;\; \Rightarrow \;\; L = 0$. Let $\mu = inf \; \{ \; W(\phi(y)) \mid y \in \mathbf{F}^{m-r} \; \} - 1$. If $W(a) + W(b) \leq \mu$ then a does not belong to $\phi(\mathbf{F}^{m-r})$ and $L = 0$. That means : for all $\lambda = (a, b)$, $\lambda \in \mathrm{D}$, such that $W(\lambda) \leq \mu$, the function f_λ is balanced. Then f is $ci(k, m)$, with $k \geq \mu$. □

Corollary 4.5 *There exist functions of type (17), which have degree $m - r + 1$ and are $ci(r-2, m)$ – i.e. they have a maximum immunity order for their degree.*

Proof: Let $r > 2$; then $m - r \leq m - 2$. We get $\phi(y) = (\phi_1(y) + 1, \ldots, \phi_r(y) + 1)$:

$$\phi_i \; : \; \mathbf{F}^{m-r} \longrightarrow \mathbf{F} \;\; , \;\; V(\phi_i) \cap V(\phi_j) = \emptyset \;\; , \;\; \exists i \; : \; |V(\phi_i)| \text{ is odd} \qquad (18)$$

($V(\phi_i)$ is the truth-table of ϕ_i). Then , for any $y \in \mathbf{F}^{m-r}$, the weight of $\phi(y)$ is at least equal to $r - 1$. Therefore f is a $ci(r-2, m)$ function; moreover f has degree $m - r + 1$, since one of the function ϕ_i has degree $m - r$. □

Example 4.2 A construction of a $ci(2, 7)$ function using corollary 4.5. *We get $m = 7$ and $r = 4$. Using Corollary 4.5, we are able to construct a $ci(2, 7)$ function f of maximal degree $d(f) = 4$. We choose the ϕ_i's, satisfying (18):*

$$\begin{array}{c|cccccccc|}
x_5 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\
x_6 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\
x_7 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
\hline
\phi_1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\
\phi_2 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
\phi_3 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
\phi_4 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0
\end{array}$$

The polynomial form of the ϕ_i's:

$$\begin{array}{lcllcl}
\phi_1(x') & = & x_5x_6 + x_6 + x_5x_7 + x_7 + x_6x_5x_7 & \phi_2(x') & = & x_5x_6 + x_5x_7 + x_5 + x_6 + x_7 + \\
\phi_3(x') & = & x_5x_7 + x_5 & \phi_4(x') & = & x_6x_5x_7 + x_5x_7
\end{array}$$

where $x' = (x_5, x_6, x_7)$. Then the function:

$$f(x) = x_1(\phi_1(x') + 1) + x_2(\phi_2(x') + 1) + x_3(\phi_3(x') + 1) + x_4(\phi_4(x') + 1),$$

where $x = (x_1, x_2, x_3, x_4, x_5, x_6, x_7)$, is a $ci(2, 7)$ function. Its polynomial form is :

$$\begin{array}{lcl}
f(x) & = & x_1x_5x_6x_7 + x_5x_6x_7x_4 + x_1x_5x_6 + x_1x_5x_7 + x_5x_6x_2 + x_5x_7x_2 + x_5x_7x_3 \\
& & + x_5x_7x_4 + x_1x_6 + x_1x_7 + x_5x_2 + x_5x_3 + x_6x_2 + x_7x_2 + x_1 + x_3 + x_4
\end{array}$$

In accordance with Theorem 3.2, we can construct two $ci(1,6)$ *functions:*

$$f(x) = (x_1 + 1)\ f_1(x_2, \ldots, x_7)\ +\ x_1\ f_2(x_2, \ldots, x_7) \qquad \textit{with}$$

$$f_1 = x_5x_6x_7x_4 + x_5x_6x_2 + x_5x_7x_2 + x_5x_7x_3 + x_5x_7x_4 + x_5x_2 + x_5x_3 + x_6x_2 + x_7x_2 + x_3 + x_4$$

$$\begin{aligned} f_2 \ = \ & x_5x_6x_7x_4 + x_5x_6x_7 + x_5x_6x_2 + x_5x_7x_2 + x_5x_7x_3 + x_5x_7x_4 + x_5x_6 \\ & + x_5x_7 + x_5x_2 + x_5x_3 + x_6x_2 + x_7x_2 + x_6 + x_7 + x_3 + x_4 + 1. \end{aligned}$$

From Theorem 3.2, f_1 *and* f_2 *are* $ci(1,6)$ *functions with maximal degree.*

References

[1] R.C. Bose & K.A. Bush *Orthogonal arrays of strength two and three*, Am. Math. Stat., 23(1952) 508-524.

[2] P. Camion *Etude de codes binaires abéliens modulaires autoduaux de petites longueurs*, Revue du CETHEDEC, NS 79-2 (1979) 3-24.

[3] P. Camion, C. Carlet, P. Charpin & N. Sendrier *Definition and construction of correlation-immune functions*, to appear as INRIA report.

[4] C. Carlet *Codes de Reed et Muller, codes de Kerdock et de Preparata*, Thèse de l'Université PARIS 6, LITP 90-59.

[5] P. Charpin *Etude sur la valuation des H-codes binaires*, Cahiers du B.U.R.O. , n. 41, Univ. P. et M. Curie , Paris 1983.

[6] P. Delsarte *An algebraic approach to the association schemes of coding theory*, Thesis, Université Catholique de Louvain, June 1973.

[7] X. Guo-zhen & J.L. Massey *A spectral characterisation of Correlation-immune Combining functions*, IEEE, vol. 34, n.3, May 88.

[8] R.A. Rueppel *Analysis and Design of stream ciphers*, Communications and Control Engineering Series, Springer-Verlag Berlin Heidelberg 1986.

[9] F.J. MacWilliams & N.J.A. Sloane *The theory of Error Correcting Codes*, North-Holland 1986.

[10] C.R. Rao *Factorial experiments derivable from combinatorial arrangements of arrays*, J. Roy. statist. Soc. 9, 128-139.

[11] T. Siegenthaler *Correlation-Immunity of nonlinear combining fonctions for Cryptographics Applications*, IEEE on Inf. Theory, vol IT-30, n.5, Sept. 84.