Construction of correlation immune Boolean functions*

Chuan-Kun Wu

School of Computing and Information Technology,
University of Western Sydney (Nepean)
PO Box 10, Kingswood, NSW 2747, Australia.
c.wu@uws.edu.au

Ed Dawson

Information Security Research Centre,
Queensland University of Technology
GPO Box 2434, Brisbane QLD 4001, Australia.
dawson@fit.qut.edu.au

Abstract

It is shown in this paper that every correlation immune Boolean function of n variables can be written as $f(x) = g(xG^T)$, where g is an algebraic non-degenerate Boolean function of k ($k \leq n$) variables and G is a generating matrix of an $[n, k, d]$ linear code. In this expression the correlation immunity of $f(x)$ must be at least $d - 1$. In this paper we further prove when the correlation immunity exceeds this lower bound. A method which can theoretically search all possible correlation immune functions exhaustively is proposed. Constructions of higher order correlation immune functions as well as algebraic non-degenerate correlation immune functions are discussed in particular. It is also shown that many cryptographic properties of g can be inherited by the correlation immune function $f(x) = g(xG^T)$ which enables us to construct correlation immune functions with other cryptographic properties.

1 Introduction

Correlation immune functions were introduced by Siegenthaler [20] in order to protect some shift register based stream ciphers against correlation attacks. Further cryp-

*Part of the content has been published in the proceedings of International Conference on Information and Communications Security (ICICS97).

tographic applications of correlation immune functions can be found in for example [1, 8, 9]. It is obvious that constructions of such functions are important, especially in the case where the constructed functions can be controlled to have other cryptographic properties. Enumeration of Boolean functions having correlation immunity and other cryptographic properties were studied in [17] and [13]. There have been alternative ways for constructing correlation immune functions (see for example [2, 3, 6, 19, 20, 22, 24]). However, the correlation immunity of the constructed functions from the methods known so far is mainly measured in terms of lower bounds. Apart from the correlation immunity, other cryptographic properties have been less considered in those constructions. In this paper we investigate the inherent structure of correlation immune functions in terms of algebraic degeneration and subsequently the constructions of functions with concrete correlation immunity are investigated. Additionally, it is shown that other cryptographic properties of the constructed functions can easily be controlled while the designed correlation immunity remains.

Denote by $F_2 = \{0, 1\}$ the binary field. A function $f : F_2^n \longrightarrow F_2$ is called a *Boolean function* of n variables. We write it as $f(x_1, ..., x_n)$ or simply $f(x)$. The *truth table* of $f(x)$ is a binary vector of length 2^n generated by $f(x)$ when x, treated as a binary integer, runs through 0 to $2^n - 1$. The *Hamming weight* of $f(x)$, denoted by $W_H(f)$, is the number of ones in its truth table. A function $f(x)$ is called *balanced* if $W_H(f) = 2^{n-1}$. The function $f(x)$ is called an *affine* function if there exist $a_0, a_1, ..., a_n \in F_2$ such that $f(x) = a_0 \oplus a_1x_1 \oplus \cdots \oplus a_nx_n$, where $\oplus$ means the modulo 2 addition. In particular, if $a_0 = 0$, $f(x)$ is also called a *linear* function. We will denote by $\mathcal{F}_n$, the set of all Boolean functions of n variables and by $\mathcal{L}_n$, the set of affine ones.

For x and y in F_2^n, we will denote by $\langle x, y\rangle = x_1y_1 \oplus x_2y_2 \oplus \cdots \oplus x_ny_n$ the inner product of x and y. It is noticed that when one of them is a constant and the other is a vector of n variables, the inner product then yields a new variable. The inner product can also be written as $x \cdot y^T$, where y^T is the transpose of y. Some concepts from the theory of error-correcting codes [14] are included here which will be used in the forthcoming discussion. An $[n, k, d]$ linear code C is a subspace of F_2^n of dimension k and with minimum distance d, i.e., the minimum Hamming weight of its code words is d. A generating matrix G of C is a $k \times n$ matrix of which the row vectors form a basis of C. For any matrix D we will denote by C_D the linear code linearly spanned by the row vectors of D.

2 Algebraic degeneration

Let $f(x) \in \mathcal{F}_n$. Then there are up to n variables which contribute to the output of the function $f(x)$. However there are cases where some variables do not contribute to the output of the function. For example, for $n = 3$, $f(x) = x_1 \oplus x_3$ is independent of x_2, i.e. regardless of whatever value is assigned to x_2, as long as the values for x_1 and x_3 are fixed, the output of $f(x)$ is fixed. This kind of function is called *degenerate*. If every variable contributes to the output of a function $f(x)$, then $f(x)$

is called a *non-degenerate function* or a *complete function*. Properties of degeneration of Boolean functions have been studied in [18] and are not addressed in this paper. In this section we study another kind of degeneration. In order to distinguish this new concept from the known one, we call it the algebraic degeneration of Boolean functions. The *algebraic degeneration* of a Boolean function is defined as: if there exists an $n \times k$ $(k < n)$ binary matrix D and a Boolean function $g(y) \in \mathcal{F}_k$ such that $f(x) \equiv g(xD)$, then $f(x)$ is called *algebraic degenerate* and $g(y)$ is called an *algebraically degenerated function associated with* D. Note that matrix D is not unique and hence the degenerated[1] function $g(y)$ is not unique. The maximum possible value of $n-k$ is called the *algebraic degeneration* of $f(x)$ and is denoted by $AD(f)$. Here matrix D is assumed to be of rank k, because otherwise there will exist another algebraic degenerated function of lesser variables. A Boolean function which cannot be algebraically degenerated to a function with less variables is called an *algebraic non-degenerate function*.

Algebraic degeneration is an important criterion for measuring the insecurity of cryptographic Boolean functions. For example an effective attack on nonlinear filtered generators was observed by Siegenthaler [21] when the nonlinear filtered function is algebraic degenerate.

It is obvious that an incomplete function, or equivalently a degenerate function, is algebraic degenerate as well. However a complete function could be algebraic degenerate. For example the exclusive-or of all variables is non-degenerate, and by a linear transformation it can be algebraically degenerated to a function of only one variable. In this sense the concept of algebraic degeration is weaker.

In order to study the algebraic degeneration and correlation immunity of Boolean functions we introduce the Walsh transform of Boolean functions. Let $f(x) \in \mathcal{F}_n$. Then the Walsh transform of $f(x)$ is expressed as

$$S_f(\omega) = \sum_x f(x)(-1)^{\langle \omega, x\rangle}, \tag{1}$$

where $\omega, x \in F_2^n$ and $\langle \omega, x\rangle = \omega_1 x_1 \oplus \omega_2 x_2 \oplus \cdots \oplus \omega_n x_n$ is the inner product of vectors ω and x. Accordingly, the inverse transform is expressed as

$$f(x) = 2^{-n} \sum_\omega S_f(\omega)(-1)^{\langle \omega, x\rangle}. \tag{2}$$

Note that the summations in (1) and in (2) are over the real number field, and the Walsh transform of a Boolean function then is a real function. It should be noted that the value of $\langle \omega, x\rangle$ could be treated as a real value when executing the operations.

It is easy to deduce that

Lemma 1 *Let $f(x) \in \mathcal{F}_n$, D be an $n \times n$ nonsingular matrix over F_2. Let $g(x) = f(xD)$. Then*

$$S_g(\omega) = S_f(\omega (D^{-1})^T), \tag{3}$$

[1] We sometimes omit the word "algebraic" but mean the same thing.

where $(D^{-1})^T$ *is the transpose of* D^{-1}.

Lemma 2 *Let* $f(x) \in \mathcal{F}_n$, $g(x) = 1 \oplus f(x)$. *Then*

$$S_g(\omega) = \begin{cases} 2^n - S_f(\omega) & \text{if } \omega = 0, \\ -S_f(\omega) & \text{if } \omega \neq 0. \end{cases} \tag{4}$$

Proof: Note that the value of the $1 \oplus f(x)$ is equivalent to the real value of $1 - f(x)$, and $\sum_x (-1)^{\langle \omega, x \rangle}$ is 2^n if $\omega = 0$ and 0 else. So we have

$$\begin{aligned} S_g(\omega) &= \textstyle\sum_x g(x)(-1)^{\langle \omega, x \rangle} \\ &= \textstyle\sum_x (1 \oplus f(x))(-1)^{\langle \omega, x \rangle} \\ &= \textstyle\sum_x (1 - f(x))(-1)^{\langle \omega, x \rangle} \\ &= \textstyle\sum_x (-1)^{\langle \omega, x \rangle} - \sum_x f(x)(-1)^{\langle \omega, x \rangle} \\ &= \begin{cases} 2^n - S_f(\omega) & \text{if } \omega = 0, \\ -S_f(\omega) & \text{if } \omega \neq 0. \end{cases} \end{aligned}$$

□

Algebraic degeneration of Boolean functions can be described by means of Walsh transforms. A useful result can be found in [11] which describes the algebraic degeneration of Boolean functions precisely.

Lemma 3[11] *Let* $f(x) \in \mathcal{F}_n$. *Denote by* $V = \prec \{\omega : S_f(\omega) \neq 0\} \succ$ *the vector space generated by the vectors on which the Walsh transform takes nonzero values, or the linear span of* $S(f) = \{\omega : S_f(\omega) \neq 0\}$. *Suppose* $dim(V) = k$, *and let* $h_1, ..., h_k$ *be a basis of* V. *Write* $H = [h_1^T, h_2^T, ..., h_k^T]$, *where* h_i^T *is the transposed vector of* h_i. *Then there must exist a Boolean function* $g(y) \in \mathcal{F}_k$ *such that*

$$f(x) = g(xH) = g(y). \tag{5}$$

It can also be shown [23] that the dimension of the vector space V is the least number k that f has an algebraic degenerated function in $\mathcal{F}_k$.

Corollary 1 *Let* $f(x) \in \mathcal{F}_n$, A *be an* $n \times n$ *nonsingular matrix, and let* $g(x) = f(xA)$. *Then* $AD(g) = AD(f)$.

Corollary 2 *Let* $f(x) \in \mathcal{F}_n$. *If* $deg(f) = n$ *then* f *is algebraic non-degenerate.*

3 Correlation immunity of Boolean functions

Let $f(x) \in \mathcal{F}_n$. The function $f(x) \in \mathcal{F}_n$ is called *correlation immune with respect to the subset* $T \subset \{1, 2, ..., n\}$ if the probability for f to take any value from $\{0, 1\}$ is not changed given that the value of $\{x_i, i \in T\}$ are fixed in advance while other variables are chosen independently at random. The function $f(x)$ is called *correlation immune*

(CI) of order t if for every T of cardinality at most t, f is CI with respect to T. It is noticed that $f(x)$ is CI of order t implies that it is CI of any order less than t as well. The largest possible value of t is called the *correlation immunity* of f. Let $z = \bigoplus_{i=1}^{n} c_i x_i$ be another (nonzero) variable, where $c_i \in \{0,1\}$. Then the function $f(x)$ is said to be *correlation immune in* z if the probability for f to take any value from $\{0, 1\}$ is not changed given that z is assigned any fixed value in advance.

Lemma 4 *Let $f(x) \in \mathcal{F}_n$. Then $f(x)$ is CI of order t if and only if for every $\gamma \in F_2^n$ with $W_H(\gamma) \le t$, $f(x)$ is CI in $z = \langle \gamma, x \rangle$.*

Proof: It is trivial to prove that $f(x)$ is CI with respect to $T \in \{1, 2, ..., n\}$, if and only if $f(x)$ is CI in $z = \langle \gamma, x \rangle$ for all γ: $\gamma_i = 1$ implies that $i \in T$. A generalisation of this observation is that $f(x)$ is CI with respect to all T of cardinality $\le t$, if and only if $f(x)$ is CI in every $z = \langle \gamma, x \rangle$ with $W_H(\gamma) \le t$. Therefore the conclusion of lemma 4 follows. □

It should be noted that $f(x)$ is CI in z_1 and z_2 individually does not imply that it is CI in $z_1 \oplus z_2$. For example, although $f(x_1, x_2, x_3) = x_3 \oplus x_1x_2 \oplus x_1x_3 \oplus x_2x_3$ is a 1-st order CI function, it is easy to verify that it is not CI in $x_1 \oplus x_2$.

Let $f(x) \in \mathcal{F}_n$, $g(y) \in \mathcal{F}_k$, $D = (d_1^T, d_2^T, ..., d_k^T)$ be an $n \times k$ binary matrix with $rank(D) = k$, where $d_i \in F_2^n$. Let $f(x) = g(xD) = g(y)$. It is known that each y_i is the linear combination of x_j's with coefficients the components of d_i, i.e., $y_i = \langle x, d_i \rangle = x \cdot d_i^T$. Let $z = \bigoplus_{i=1}^{n} c_i x_i$ be another variable. Then it is obvious that $f(x)$ is CI in z if and only if $g(y)$ is CI in z. Denote by $\gamma = (c_1, c_2, ..., c_n)$. We have

Lemma 5 *If $rank[D; \gamma^T] = k+1$, where $[A; B]$ means the concatenation of matrices A and B, then for any Boolean function $g(y) \in \mathcal{F}_k$, $g(xD)$ is independent of $z = \langle \gamma, x \rangle$ and hence is CI in z.*

Proof: Let $y = (y_1, y_2, ..., y_k) = xD$. It is noticed that $rank[D; \gamma^T] = k+1$ if and only if variables $y_1, y_2, ..., y_k$ together with z are all independent, and consequently $g(xD)$ is independent of z. So we have

$$Prob(g(xD) = 1 | z = 1) = Prob(g(y) = 1 | z = 1) = Prob(g(y) = 1).$$

This means that $g(xD)$ is CI in z. □

The following lemma has been proved both in [22] and in [24] using different methods.

Lemma 6 *If G is a generating matrix of an $[n, k, d]$ linear code, then for any $g(y) \in \mathcal{F}_k$, the correlation immunity of $f(x) = g(xG^T)$ is at least $d - 1$.*

In order for the function f to have correlation immunity of order larger than $d-1$, by the definition of correlation immunity and lemma 4 and lemma 5, we need to make $g(y)$, or equivalently $f(x) = g(xG^T)$, to be CI in every $z = \langle x, \gamma \rangle$ with $W_H(\gamma) = d$.

It is obvious that $rank[G^T, \gamma^T] = k$ if and only if γ is a codeword of C_G, the linear code generated by G. By lemma 5 we know that for those γ with Hamming weight d which are not codewords of C_G, the function f is already CI in $z = \langle x, \gamma \rangle$. So we have

Lemma 7 *Let G be a generating matrix of an $[n, k, d]$ linear code, and $f(x) = g(xG^T)$. Then f is CI of order $\geq d$ if and only if for every $\alpha \in F_2^k$ with $W_H(\alpha G) = d$, $g(y)$ is CI in $z = \langle \alpha, y \rangle$.*

Proof: It can be proved by setting $\gamma = \alpha G$ and consequently we have $\langle \alpha, y \rangle = \langle x, \gamma \rangle$. By lemma 4 the conclusion follows. □

By generalising lemma 7 we have

Theorem 1 *Let G be a generating matrix of an $[n, k, d]$ linear code, and $f(x) = g(xG^T)$. Then a necessary and sufficient condition for the function f to be CI of order m is that for every $\alpha \in F_2^k$ with $d \leq W_H(\alpha G) \leq m$, $g(y)$ is CI in $z = \langle \alpha, y \rangle$.*

Corollary 3 *If the i-th row vector of G is a codeword with nonzero minimum Hamming weight d and the function $g(y)$ is not CI in y_i, then the correlation immunity of $f(x) = g(xG^T)$ is exactly $(d-1)$.*

Now we consider the inverse question for general CI functions. Given an m-th order CI function $f \in \mathcal{F}_n$, can it be written as $f(x) = g(xD)$, where $g \in \mathcal{F}_k$ is algebraic non-degenerate and D^T is a generating matrix of an $[n, k, d]$ linear code with $k \leq n$ and $d \geq 1$? The answer is yes according to lemma 8. Furthermore it can be shown that the code generated by D^T is unique.

Lemma 8 *Let $f(x) \in \mathcal{F}_n$. Then it can be written as $f(x) = g(xD)$, where $g \in \mathcal{F}_k$ is algebraic non-degenerate and D^T is a generating matrix of an $[n, k, d]$ linear code with $k \leq n$ and $d \geq 1$. Moreover, the linear code is unique given that $f(x)$ is fixed.*

Proof: From the discussion above, what we need to show is the uniqueness of the code. On the contrary we suppose $f(x) = g_1(xD_1) = g_2(xD_2)$, where $C_{D_1^T} \neq C_{D_2^T}$. Then there must exist a column α of D_1 which is linearly independent of the column vectors of D_2. Without loss of generality let α be the first column of D_1. Then by lemma 5 we know that $f(x)$ is independent of $\langle \alpha, x \rangle$, and equivalently $g_1(y)$ must be independent of y_1. This is in contradiction with the premise of the lemma. So the conclusion is true. □

By lemma 8 we know that theorem 1 gives a necessary and sufficient condition for a general Boolean function to be CI. Since theorem 1 applies to every CI function, it can be used to develop exhaustive constructions of CI functions.

4 Some known constructions and their non-exhaustiveness

One aim of this paper is to develop the construction of CI functions described in lemma 6. The limitations of the construction of lemma 6 is shown in the example in the appendix where we construct some CI functions which are beyond the capability of lemma 6. Besides the construction of CI functions described in lemma 6, there have been numerous methods in constructing CI functions (see for example [2, 3, 6, 19, 20, 23, 12]). Some of these constructions are for functions over finite fields or Galois rings. As we are only concerned with Boolean functions in this paper, we will consider the following constructions which were initially studied in [20] and [2]. Some other constructions are extensions or variations of them.

Lemma 9 ([20]) *Let $f_1(x), f_2(x) \in \mathcal{F}_n$ be two m-th order CI functions with $W_H(f_1) = W_H(f_2)$. Then*

$$f(x_1, ..., x_{n+1}) = x_{n+1} f_1(x) \oplus (1 \oplus x_{n+1}) f_2(x) \tag{6}$$

is an m-th order CI function with $W_H(f) = 2W_H(f_1)$.

Lemma 10 ([2]) *Let $f_1(x) \in \mathcal{F}_n$ be balanced. Write $\overline{x} = (x_1 \oplus 1, ..., x_n \oplus 1)$. Then*

1. *$f(x_1, ..., x_{n+1}) = f_1(x) \oplus x_{n+1}$ is a balanced $(k+1)$-th order CI function in $\mathcal{F}_{n+1}$ if and only if $f_1(x)$ is a k-th order CI function of $\mathcal{F}_n$.*

2. *$f(x_1, ..., x_{n+1}) = f_1(x) \oplus x_{n+1}(f_1(x) \oplus f_1(\overline{x}))$ is a balanced $(k+1)$-th order CI function in $\mathcal{F}_{n+1}$ if and only if $f_1(x)$ is a k-th order CI function of $\mathcal{F}_n$.*

The two constructions above are both based on known CI functions. In [2] a more direct construction is proposed which can be described as follows:

Lemma 11 ([2]) *Let n_1, n_2, n be positive integers with $n_1 + n_2 = n$, $r(y), \phi_i(y) \in \mathcal{F}_{n_2}$, $i = 1, ..., n_1$. Let*

$$f(x; y) = \bigoplus_{i=1}^{n_1} x_i \phi_i(y) \oplus r(y). \tag{7}$$

Then $f(x; y)$ is a balanced Boolean function in $\mathcal{F}_n$ with correlation immunity of order

$$k \geq inf\{W_H(\phi_1(y), ..., \phi_{n_1}(y)) : \ y \in \mathcal{F}_2^{n_2}\}.$$

The non-exhaustiveness of the constructions studied in [2] (lemma 10 and lemma 11 above) is obvious because they can only construct balanced CI functions. As for the non-exhaustiveness of the construction of lemma 9, it can easily be checked when the CI function $x_1 \oplus x_1x_2 \oplus x_1x_3 \oplus x_2x_3$ is written as $x_i f_1(\hat{x}_i) \oplus (1 \oplus x_i) f_2(\hat{x}_i)$, where $\hat{x}_i$ is a collection of x_j excluding x_i, $f_2(\hat{x}_i)$ is always not CI at all. So it is beyond the capability of the construction described in lemma 9. We should also note that when a CI function is written in this way, $W_H(f_1) = W_H(f_2)$ is always true which is just part of the premise of lemma 9.

5 Exhaustive construction of CI functions

Theoretically by using lemma 6 and theorem 1 the complete set of CI functions can be constructed. By applying theorem 1 we are able to see when the correlation immunity is larger than or equal to the minimum distance of the code. In order to do this, we need to construct Boolean functions which are CI in some of their variables and/or their linear combinations. Let $\hat{x_i} = (x_1, ..., x_{i-1}, x_{i+1}, ..., x_n)$. Then we have

Lemma 12 *Let $f(x) = x_i f_1(\hat{x_i}) \oplus f_2(\hat{x_i})$. Then $f(x)$ is CI in x_i if and only if*

$$W_H(f_1 \oplus f_2) = W_H(f_2). \tag{8}$$

Proof: By writing $f(x) = x_i(f_1(\hat{x_i}) \oplus f_2(\hat{x_i})) \oplus (1 \oplus x_i) f_2(\hat{x_i})$ it can be seen that $f(x)$ is CI in x_i if and only if $W_H(f_1 \oplus f_2) = W_H(f_2) = \frac{1}{2} W_H(f)$. □

Lemma 13 *Let $f(x) \in \mathcal{F}_n$. Then $deg(f) < n$ if and only if $2 | W_H(f)$, i.e., the Hamming weight of $f(x)$ is an even number.*

In [20] it was shown that if $f(x) \in \mathcal{F}_n$ is CI (of order ≥ 1), then $deg(f) \leq n-1$. We further prove that

Lemma 14 *Let $f(x) \in \mathcal{F}_n$. If $deg(f) = n$ then $f(x)$ is not CI in any linear combination of its variables.*

Proof: Assume the contrary, $f(x)$ is CI in $\langle \alpha, x \rangle$, and without loss of generality the first coordinate of α is assumed to be not zero. Denote by δ_i the vector in F_2^n with i consecutive ones followed by zeros. Let $D = [\alpha^T, \delta_2^T, ..., \delta_n^T]$. Then $g(x) = f(xD^{-1})$ is CI in x_1 and hence can be written as $g(x) = x_1 g_1(\hat{x_1}) \oplus g_2(\hat{x_1})$. By lemma 12 we know that

$$\begin{aligned} W_H(g_1) &= W_H((g_1 \oplus g_2) \oplus g_2) \\ &= W_H(g_1 \oplus g_2) + W_H(g_2) - 2W_H((g_1 \oplus g_2) \cdot g_2) \\ &= 2W_H(g_2) - 2W_H((g_1 \oplus g_2) \cdot g_2) \end{aligned}$$

is an even number and by lemma 13 we have $deg(f) = deg(g) = deg(g_1) + 1 < (n-1) + 1 = n$. This is a contradiction. So the conclusion of lemma 14 follows. □

Let $f(x) = g(xG^T)$ be a Boolean function of $\mathcal{F}_n$, where g is algebraic non-degenerate, and G is a generating matrix of an $[n, k, d]$ linear code. It is easy to see that by a linear transform on the rows of G, we can always make the row vectors of G satisfy

$$W_H(g_1) \leq W_H(g_2) \leq \cdots \leq W_H(g_k),$$

and there does not exist another basis $\beta_1, \beta_2, ..., \beta_k$ of C_G with $W_H(\beta_1) \leq W_H(\beta_2) \leq \cdots \leq W_H(\beta_k)$ such that $W_H(\beta_i) < W_H(g_i)$ for some $1 \leq i \leq k$. Constructions can always be based on this assumption. Such a matrix will be called a *minimum weight generating matrix.*

It is noticed that under a permutation of the variables of a Boolean function, the correlation immunity of the function is an invariant. To simplify the problem we will treat two CI functions as equivalent if they are equivalent by a variable permutation. For the function $f(x) = g(xG^T)$ a permutation of x is equivalent to the same permutation of the column vectors of G. Complements of CI functions can be left out in the first steps and then added at last. So the exhaustive construction can be outlined as follows:

For all integers $k \in \{1, 2, ..., n\}$ perform the following steps:

1. Search the minimum weight generating matrices G_i, $i \in I$, of $[n, k]$ codes such that they are not column-equivalent, where I is a set of complete index.
2. List all nontrivial Boolean functions $g(y) \in \mathcal{F}_k$ such that $g(\mathbf{0}) = 0$.
3. Match each $g(y)$ with every G_i to see if $f_i(x) = g(xG_i^T)$ is CI of any order according to theorem 1.
4. For those $f_i(x)$ with a certain order of CI, permute their variables to get an equivalent class of CI functions.
5. Complement every CI function obtained above.

Theoretically the above step can exhaustively generate all the CI functions. However because of the large number of CI functions of n variables when n is sufficiently large, it is not surprising to see that the above steps are not practically efficient in terms of computational complexity (such as step 3). So more efficient constructions of particular CI functions are required.

6 Construction of high order CI functions

From the above, every CI function can be written as $g(xD)$, where g is an algebraic non-degenerate function and D^T is a minimum weight generating matrix of an $[n, k, d]$ linear code. In this section we will concentrate mainly on the construction of those functions whose correlation immunity is not less than d.

For any Boolean function $f(x) \in \mathcal{F}_n$, set

$$\Delta_f = \{\delta \in F_2^n, \quad f(x) \text{ is CI in } \langle \delta, x \rangle\}. \tag{9}$$

Then by theorem 1 we have

Theorem 2 *Let $g(y) \in \mathcal{F}_k$ and G be a generating matrix of an $[n, k, d]$ linear code. Set $f(x) = g(xG^T)$. Then the correlation immunity of $f(x)$ is*

$$min_{\alpha \notin \Delta_g} W_H(\alpha G) - 1. \tag{10}$$

Moreover we have

$$AD(f) = n - k + AD(g). \tag{11}$$

Proof: The former part (equation 10) comes directly from theorem 1. So we need only to prove the latter part. Assume $AD(g) = t$, i.e., there exists an algebraic non-degenerate function $g_1 \in \mathcal{F}_{k-t}$ and a $k \times (k-t)$ matrix D such that $g(y) = g_1(yD)$. So $f(x) = g_1(xG^T D)$, and $AD(f) \geq n - (k-t) = n - k + AD(g)$.

On the other hand, since $rank(G) = k$, we can assume, without loss of generality, that the first k columns of G are linearly independent and we write $G = [G_1; G_2]$. Then $g(y) = f(yG_1^{-1}, 0, \cdots, 0)$. This means that if f can be algebraically degenerated to a function of r variables then g can be algebraically degenerated to a function of no more than r variables, i.e., $k - AD(g) \leq n - AD(f)$ or $AD(f) \leq n - k + AD(g)$.

In light of the above discussion, the conclusion follows. $\square$

In order to determine Δ_f for a general Boolean function $f(x) \in \mathcal{F}_n$ we have

Theorem 3 *Let $f(x) \in \mathcal{F}_n$ and $\delta \in F_2^n$. Then $\delta \in \Delta_f$ if and only if*

$$S_f(\delta) = 0. \tag{12}$$

Proof: $\delta \in \Delta_f \iff f(x)$ is CI in $\langle \delta, x \rangle \iff Prob(f(x) = 1 | \langle \delta, x \rangle = 0) = Prob(f(x) = 1 | \langle \delta, x \rangle = 1) \iff \sum_{\langle \delta, x \rangle = 0} f(x) - \sum_{\langle \delta, x \rangle = 1} f(x) = 0 \iff$
$S_f(\delta) = \sum_x f(x)(-1)^{\langle \delta, x \rangle} = \sum_{\langle \delta, x \rangle = 0} f(x) - \sum_{\langle \delta, x \rangle = 1} f(x) = 0.$ $\square$

By theorem 3, (10) can be rewritten as

$$\min_{\alpha:\ S_g(\alpha) \neq 0} \quad W_H(\alpha G) - 1. \tag{10'}$$

It is seen that using the techniques of Walsh transforms the correlation immunity of $f(x) = g(xG^T)$ can easily be determined by (10').

Note that $g(y)$ can always be chosen as algebraic non-degenerate which enables us to construct CI functions with least possible algebraic degeneration. When we use theorem 2 to construct CI functions, it is noticed that an $[n, k, d]$ linear code normally has several code words of Hamming weight d. So in general it is hard to find a Boolean function which can match a generating matrix of this linear code to generate CI functions of order $\geq d$. However it is easy to find Boolean functions which are CI in part of their variables and their linear combinations as shown in the following.

Corollary 4 *Let $g(y) \in \mathcal{F}_k$ be CI in its first t variables and their nonzero linear combinations. Let G be a generating matrix of an $[n-t, k-t, d]$ linear code. Then the correlation immunity of function $f(x) = g(x\hat{G}^T)$ is at least $d-1$, where*

$$\hat{G} = \begin{bmatrix} D & 0 \\ 0 & G \end{bmatrix},$$

and D is an arbitrary nonsingular binary matrix of order $t \times t$.

We note that when corollary 4 is used to construct CI functions, the size of D is normally small as the cases demonstrated in the example of the appendix. For special cases we have

Corollary 5 *If G is a generating matrix of an $[n,k,d]$ linear code and the row vectors of G include all the code words of Hamming weight d, then for any algebraic non-degenerate Boolean function $g(y)$ of k variables with correlation immunity of order t, $f(x) = g(xG^T)$ is a CI function of order $t+1$.*

7 Construction of CI functions with associated cryptographic properties

In practice a CI function is required to satisfy other cryptographic properties as well. Cryptographic properties of Boolean functions which have commonly been studied include the following:

- **Balance:** Let $f(x) \in \mathcal{F}_n$. The balance of $f(x)$ is defined as

$$\begin{aligned} Bal(f) &= 1 - |W_H(f) - 2^{n-1}|/2^{n-1} \\ &= \begin{cases} W_H(f)/2^{n-1} & \text{if } W_H(f) \le 2^{n-1}, \\ (2^n - W_H(f))/2^{n-1} & \text{if } W_H(f) > 2^{n-1}. \end{cases} \end{aligned}$$

  When $Bal(f) = 1$, $f(x)$ is called *balanced* and when $Bal(f) = 0$, $f(x)$ is called extremely unbalanced as $f(x)$ is a constant in this case.

- **Algebraic degree:** The algebraic degree or simply degree of a Boolean function is defined as the largest number of variables in one product term of its polynomial expression and denoted by $deg(f)$.

- **Nonlinearity:** The nonlinearity of a Boolean function $f(x) \in \mathcal{F}_n$, denoted by N_f, is the minimum distance of f from all affine functions in $\mathcal{L}_n$.

- **Propagation criterion:** A Boolean function $f(x) \in \mathcal{F}_n$ is said to satisfy the propagation criterion with respect to a non-zero vector α if $f(x) \oplus f(x \oplus \alpha)$ is balanced.

  A Boolean function $f(x)$ is said to satisfy the propagation criterion of order k if it satisfies the propagation criterion with respect to all α with $1 \le W_H(\alpha) \le k$, and denoted by $PC(f) = k$.

  Note: *Strict Avalanche Criterion* (SAC) is equivalent to the propagation criterion of order 1 ($PC(f) = 1$) and *perfect nonlinearity* defined in [15] is equivalent to the propagation criterion of order n ($PC(f) = n$).

- **Linear structure:** A boolean function $f(x) \in \mathcal{F}_n$ is said to have a linear structure $\alpha \in F_2^n$ if $f(x) \oplus f(x \oplus \alpha) \equiv c$, where c is a constant of $\{0,1\}$. In particular α is called an *invariant linear structure* if $c = 0$ and a *complement linear structure* if $c = 1$.

- **Algebraic degeneration:** As described earlier in this paper.

From the discussions above we know that every CI function can be written as $f(x) = g(xG^T)$, where g is an algebraic non-degenerate Boolean function of k variables and G is a generating matrix of an $[n, k, d]$ linear code. We will show that some cryptographic properties of g can be inherited by the CI function f.

7.1 CI functions with good balance

From the view point of cryptographic applications, we aim to construct CI functions with as good a balance as possible. The balance of CI function given in the form $f(x) = g(xG^T)$ can easily be controlled by choosing g to be of a good balance.

Lemma 15 *Let $f(x) = g(xD)$, where g is an algebraic non-degenerate Boolean function of k variables and D^T is a generating matrix of an $[n, k, d]$ linear code. Then*

$$Bal(g) = Bal(f).$$

Particularly, $f(x)$ is balanced if and only if $g(y)$ is such.

Proof. Denote by $KerD = \{x : \ xD = 0\}$. For any $y \in F_2^k$, since $rank(D) = k$, there must exist an $x \in F_2^n$ such that $y = xD$. So $x + KerD$ is the set of all solutions of equation $xD = y$. This means that when there exists an y such that $g(y) = 1$, there will exist 2^{n-k} $\overline{x}_i$ such that $\overline{x}_i D = y$ and $f(\overline{x}_i) = 1$. So we have that $W_H(f) = 2^{n-k} \cdot W_H(g)$. By the definition we have the conclusion. □

7.2 CI functions with high algebraic degree

Algebraic degree is one criterion to measure the nonlinearity of Boolean functions. In practical applications, a CI function is required to have as high algebraic degree as possible. Otherwise there may be a risk in decreasing its security when the low order approximation technique [16] is applied. It can be shown that the degree of f is the same as that of g.

Lemma 16 : *Let $f(x) \in \mathcal{F}_n$ and A be an $n \times n$ nonsingular binary matrix. Then $deg(f(xA)) = deg(f(x))$.*

Proof: Denote by $f_1(x) = f(xA)$. It is obvious that the expansion of $f(xA)$ does not generate a term with degree $> deg(f(x))$, so we have $deg(f_1(x)) \leq deg(f(x))$. On the other hand, from the non-singularity of A we have $f(x) = f_1(xA^{-1})$ and hence $deg(f(x)) \leq deg(f_1(x))$. Therefore, $deg(f_1(x)) = deg(f(x))$. □

Theorem 4 *Let D be an $n \times k$ $(k \leq n)$ binary matrix and let $f(x) = g(xD)$, where $g \in \mathcal{F}_k$. Then $deg(f) = deg(g)$ holds for any g if and only if $rank(D) = k$.*

Proof. By row-transformation, matrix D can be written as

$$D = A \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} P$$

where A is an $n \times n$ nonsingular matrix, I_r is an $r \times r$ $(r \leq k)$ identity matrix and P is a $k \times k$ permutation matrix. Then

$$f(x) = g(xD) = g(xA \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} P)$$

Denote by $f_1(x) = f(xA^{-1})$, $g_1(y) = g(yP)$, where $x \in F_2^n$ and $y \in GF^k(2)$. Then

$$\begin{aligned} f_1(x) &= f(xA^{-1}) = g(xA^{-1}D) = g(x \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} P) \\ &= g_1(x \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}) = g_1(x_1, \ldots, x_r, 0, \ldots, 0). \end{aligned}$$

From the equation above we see that

$$deg(f_1) = deg(g_1(x_1, \ldots, x_r, 0, \ldots, 0)) = deg(g_1(y))$$

holds for any $g_1(y) \in \mathcal{F}_k$ if and only if $r = k$, i.e., if and only if rank(D)=k. Notice that by lemma 16, $deg(g_1) = deg(g)$ and $deg(f_1) = deg(f)$. So we have $deg(f) = deg(g)$ holds for any $g(y) \in \mathcal{F}_k$ if and only if $rank(D)$=k. □

From theorem 4 we see that the maximum algebraic degree of the function written as $f(x) = g(xD)$ is k. In this case by corollary 3 and lemma 14, the correlation immunity of $f(x)$ is exactly $d - 1$, where D^T is the generating matrix of an $[n, k, d]$ linear code. This is consistent with Siegenthaler's inequality [20]. The discussion above also shows that we can construct CI functions which meet the equality (maximum correlation immunity/algebraic degree) of Siegenthaler's inequality.

7.3 CI functions with high nonlinearity

Nonlinearity of Boolean functions is a measurement of the distance of Boolean functions to the nearest affine one [15]. If the nonlinearity of a Boolean function is very low, then it can be approximated by an affine Boolean function with high correlation with the affine function [7] and hence is cryptographically insecure. By using the Walsh spectral techniques it is easy to deduce that

Lemma 17

$$N_f = min\{W_H(f),\ 2^n - W_H(f),\ 2^{n-1} - \max_{\omega \neq 0} |S_f(\omega)|\}. \tag{13}$$

Lemma 18 *Let* $f(x) = g(xG^T)$, *where* g *is an algebraic non-degenerate Boolean function of* k *variables and* G *is a generating matrix of an* $[n,k,d]$ *linear code. Then*

$$N_f \leq 2^{n-k} N_g.$$

Proof: By the definition of nonlinearity there exists an affine function $l(y)$ of k variables such that $W_H(g(y) \oplus l(y)) = N_g$. Hence we have $W_H(g(xG^T) \oplus l(xG^T)) = 2^{n-k} N_g$ and again by the definition we have $N_f \leq 2^{n-k} N_g$. □

Furthermore we can prove

Theorem 5 *Let* D *be an* $n \times k$ $(k \leq n)$ *binary matrix. Then* $rank(D) = k$ *if and only if for any Boolean function* $g(y) \in \mathcal{F}_k$ *and* $f(x) = g(xD)$ *we have*

$$N_f = 2^{n-k} N_g. \tag{14}$$

In order to prove theorem 5, the following lemmas will be used.

Lemma 19 *Let* V *be a vector subspace of* F_2^n. *Then*

$$\sum_{x \in V} (-1)^{\langle \omega,\, x \rangle} = \begin{cases} \#(V) & \text{if } \omega \in V^{\perp}, \\ 0 & \text{otherwise}, \end{cases} \tag{15}$$

where $\#(A)$ *denotes the cardinality of the set* A *and* $V^{\perp} = \{y : \langle x,\, y \rangle = 0$ *for every* $x \in V\}$ *is the orthogonal space of* V.

Lemma 20 *Let* $f_1(x) = f_2(xA)$, *where* A *is an* $n \times n$ *nonsingular matrix. Then* $N_{f_1} = N_{f_2}$.

Lemma 21 *Let* $D = \begin{bmatrix} D_1 \\ 0 \end{bmatrix}$ *be an* $n \times k$ *binary matrix, where* D_1 *is a* $k \times k$ *non-singular matrix. Let* $f(x) = g(xD)$. *Then* $N_f = 2^{n-k} N_g$.

Proof: For any vector $\alpha \in F_2^n$ we will write $\underline{\alpha}_1 = (\alpha_1, \cdots, \alpha_k)$. It is easy to see that

$$\begin{array}{l} KerD = \{(0, ..., 0, x_{k+1}, ..., x_n) : \ x_i \in F_2\}, \\ (KerD)^{\perp} = \{(x_1, ..., x_k, 0, ..., 0) : \ x_i \in F_2\}. \end{array}$$

Noticing that $F_2^n = (KerD)^{\perp} \oplus KerD$, we have

$$\begin{aligned} S_f(\omega) &= \textstyle\sum_x f(x)(-1)^{\langle \omega,\, x \rangle} \\ &= \textstyle\sum_x g(xD)(-1)^{\langle \omega,\, x \rangle} \\ &= \textstyle\sum_{x \in (KerD)^{\perp}} \sum_{y \in KerD} g((x \oplus y)D)(-1)^{\langle \omega,\, (x \oplus y) \rangle} \\ &= \textstyle\sum_{x \in (KerD)^{\perp}} g(xD)(-1)^{\langle \omega,\, x \rangle} \sum_{y \in KerD} (-1)^{\langle \omega,\, y \rangle}. \end{aligned}$$

By lemma 19 we know that $S_f(\omega) = 0$ if $\omega \notin (KerD)^\perp$. If $\omega \in (KerD)^\perp$ we have

$$\begin{aligned} S_f(\omega) &= 2^{n-k}\sum_{x\in(KerD)^\perp} g(xD)(-1)^{\langle \omega,\, x\rangle} \\ &= 2^{n-k}\sum_{x\in(KerD)^\perp} g(\underline{x}_1 D_1)(-1)^{\langle \underline{\omega}_1,\, \underline{x}_1\rangle} \quad \text{(by lemma 1)} \\ &= 2^{n-k} S_g(\underline{\omega}_1 (D_1^{-1})^T). \end{aligned}$$

This means that

$$max_{\omega\neq 0}\ |S_f(\omega)| = \max_{\omega_1 \neq 0}\ 2^{n-k}|S_g(\underline{\omega}_1)|.$$

Notice that $W_H(f) = 2^{n-k} W_H(g)$. By lemma 17 we have $N_f = 2^{n-k} N_g$. □

Proof of theorem 5: Necessity: Since $rank(D) = k$, there must exist a nonsingular $n \times n$ matrix R such that $RD = D' = \begin{bmatrix} D_1 \\ 0 \end{bmatrix}$. Write

$$f_1(x) = f(xR) = g(xRD) = g(xD').$$

Then by lemma 21 we have $N_{f_1} = 2^{n-k} N_g$. But by lemma 20 we have $N_f = N_{f_1}$. So the conclusion follows.

Sufficiency: On the contrary we assume that $rank(D) < k$. Then the columns of $D = [d_1^T, \cdots, d_k^T]$ are linearly dependent, i.e., for some i-th column of D, there must exist $a_j \in F_2$ such that

$$d_i = a_1 d_1 \oplus \cdots \oplus a_{i-1} d_{i-1} \oplus a_{i+1} d_{i+1} \oplus \cdots \oplus a_k d_k.$$

If d_i is an all-zero vector, then for any $j \neq i$, set $g(y) = y_i y_j$ to be a quadratic function which has nonzero nonlinearity, $f(x) = g(xD) = (xd_i^T)(xd_j^T) = 0$ has zero nonlinearity. If d_i is a nonzero vector, then set $g(y) = y_i(a_1 y_1 \oplus \cdots \oplus a_{i-1} y_{i-1} \oplus a_{i+1} y_{i+1} \oplus \cdots \oplus a_k y_k)$ to be a quadratic function which has nonzero nonlinearity. Then

$$\begin{aligned} f(x) &= g(xD) \\ &= (xd_i^T)(a_1 x d_1^T \oplus \cdots \oplus a_{i-1} x d_{i-1}^T \oplus a_{i+1} x d_{i+1}^T \oplus \cdots \oplus a_k x d_k^T) \\ &= (xd_i^T)(xd_i^T) \\ &= xd_i^T \end{aligned}$$

is a linear function which has zero nonlinearity. This is a contradiction with (14) and hence the conclusion of theorem 5 is true. □

From theorem 5 we know that, if a CI function is constructed in the form $f(x) = g(xD)$, where D is an $n \times k$ matrix with $rank(D) = k$, then $f(x)$ has maximum possible nonlinearity if and only if $g(x)$ has the maximum possible nonlinearity as well. There have been alternative methods for constructing Boolean functions with high nonlinearity (refer to [4, 5, 19, 27]). With Boolean functions having high order nonlinearity, CI functions having high nonlinearity can be constructed according to theorem 5.

7.4 CI functions with propagation criterion

Unlike other properties, the propagation property is not inheritable from g to f for the expression $f(x) = g(xD)$, i.e., g satisfies propagation criterion does not guarantee that f does. For example, let

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Although $g(y) = y_1y_2 \oplus y_2y_3 \oplus y_3y_4 \oplus y_4y_5 \oplus y_1y_5$ satisfies the propagation criterion of order 4, $f(x) = g(xD) = x_1x_2 \oplus x_2x_3 \oplus x_3x_4 \oplus x_4x_5 \oplus x_1x_5 \oplus x_6$ does not satisfy the propagation criterion of order 1. In order to study the way that the propagation property of f relates to that of g more precisely, for $f(x) \in \mathcal{F}_n$, we denote by $NP(f) = \{\alpha \in F_2^n,\ f(x) \oplus f(x \oplus \alpha) \text{ is not balanced}\}$.

Theorem 6 *Let $f(x) = g(xD)$, where $g(y) \in \mathcal{F}_k$ and D is an $n \times k$ binary matrix with $rank(D) = k$. Then the propagation criterion order of $f(x)$ is*

$$PC(f) = \min_{\alpha D \in NP(g)} W_H(\alpha) - 1.$$

Proof: We first prove that $\alpha \in NP(f)$ if and only if $\alpha D \in NP(g)$. It is easy to verify (refer the proof of lemma 15) that when $rank(D) = k$, xD forms k uniform random variables provided that x is a collection of n uniform random variables. So $g(xD) \oplus g(xD \oplus \beta)$ is unbalanced if and only if $\beta \in NP(g)$. So $\alpha \in NP(f)$ $\Longleftrightarrow$ $f(x) \oplus f(x \oplus \alpha)$ is unbalanced $\Longleftrightarrow$ $g(xD) \oplus g(xD \oplus \alpha D)$ is unbalanced $\Longleftrightarrow$ $\alpha D \in NP(g)$. By the definition that

$$PC(f) = \min_{\alpha \in NP(f)} W_H(\alpha) - 1$$

the conclusion follows. □

Particularly, when g satisfies the propagation criterion of the maximum order k, i.e., g is a bent function (or g is perfect nonlinear and k is even in this case), we have

Corollary 6 *Let $g \in \mathcal{F}_k$ be such that g satisfies the propagation criterion of order k, i.e., g is perfect nonlinear, and let D be an $n \times k$ matrix with $rank(D) = k$. Then $f(x) = g(xD)$ satisfies the propagation criterion of order k.*

Proof: Note that $g(y) \in \mathcal{F}_k$ satisfies the propagation criterion of order k if and only if $NP(g) = \{0\}$. Since $rank(D) = k$, it is obvious that $\alpha D \in NP(g)$ or equivalently

$\alpha D = 0$ only if $W_H(\alpha) \geq k+1$. We can also find an α with $W_H(\alpha) = k+1$ such that $\alpha D = 0$. So by theorem 6 the conclusion of corollary 6 is true. $\square$

In the case of corollary 6, function $f(x)$ has the same propagation criterion order as that of $g(y)$. Is it possible that $f(x)$ has a higher propagation criterion order than that of $g(y)$? The answer is yes as demonstrated by the following example. It can be verified that $g(x_1, ..., x_5) = x_1x_2 \oplus x_3x_4 \oplus x_5$ satisfies propagation criterion of order 0. Let

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Then $f(x_1, ..., x_6) = g((x_1, ..., x_6)A) = x_1 \oplus x_1x_2 \oplus x_2x_3 \oplus x_4 \oplus x_1x_4 \oplus x_3x_4 \oplus x_1x_5 \oplus x_4x_5 \oplus x_6 \oplus x_1x_6 \oplus x_4x_6 \oplus x_5x_6$ satisfies the propagation criterion of order 3. We can easily find more such examples. However, as the propagation criterion characteristics of different functions are very different, and the choice of the matrices can be variant, we do not have a systematic way for constructing CI functions in the form $f(x) = g(xD)$ such that the propagation criterion order of f is higher than that of g. We leave this as an open problem.

7.5 Linear structure characteristics of CI functions

It is known that the more linear structures a Boolean function has, the closer the function is related to an affine function. In the extreme case when every vector is a linear structure of a Boolean function, it must be an affine one. From a cryptographic point of view, a Boolean function is required to have as few linear structures as possible. However, when a Boolean function can be written as $f(x) = g(xD)$, it definitely has linear structures if $k < n$. The relationship between the linear structures of f and that of g can be described as follows.

Theorem 7 *Let $f(x) = g(xD)$, where D is an $n \times k$ ($k \leq n$) matrix with $rank(D) = k$. Then α is an invariant (a complement) linear structure of f if and only if αD is an invariant (a complement) linear structure of g.*

Proof. The sufficiency is obvious. So we only need to present the proof of the necessity. Assume the contrary, i.e., there exists a vector $\alpha \in F_2^n$ such that $f(x) \oplus f(x \oplus \alpha) \equiv c$ and $g(y) \oplus g(y \oplus \alpha D) \not\equiv c$. Let $g(y') \oplus g(y' \oplus \alpha D) \neq c$. Since $rank(D) = k$, there must exist an $x' \in F_2^n$ such that $y' = x'D$. So we have

$$f(x') \oplus f(x' \oplus \alpha) = g(x'D) \oplus g((x' \oplus \alpha)D) = g(y') \oplus g(y' \oplus \alpha D) \neq c.$$

This is a contradiction of the assumption. So the conclusion is true. $\square$

Corollary 7 *Let $f(x) = g(xD)$, where D is an $n \times k$ $(k \leq n)$ matrix with $rank(D) = k$. Denote by V_f and V_g the set of linear structures of f and g respectively. Then $dim(V_f) = (n-k) + dim(V_g)$, where $dim(.)$ means the dimension of a vector space.*

It can be seen from corollary 7 that even if g has no nonzero linear structures, f may have because the all-zero vector is an invariant linear structure (trivial) of every function. It also implies that a Boolean function may have many invariant linear structures but no complement ones.

We have shown above that if a function is algebraic degenerate, it must have nonzero invariant linear structures. Is this also a sufficient condition for a Boolean function to be algebraic degenerate? The following gives a positive answer.

Theorem 8 *Let $f(x) \in \mathcal{F}_n$, $V_I(f)$ be the linear space of all the invariant linear structures of $f(x)$ and $dim(V_I(f)) = k$. Then there must exist a nonsingular matrix A over F_2 such that*

$$g(x_1, ..., x_n) = f((x_1, ..., x_n)A) = g_1(x_{k+1}, ..., x_n),$$

where $g_1(x_{k+1}, ..., x_n)$ has no nonzero invariant linear structures. Moreover, $g_1(x_{k+1}, ..., x_n)$ has a complementary linear structure, or equivalently it can be written as $g_1(x_{k+1}, ..., x_n) = x_{k+1} \oplus g_2(x_{k+2}, ..., x_n)$, if and only if f has a complementary linear structure.

Proof: Let A be an $n \times n$ binary matrix such that the first k rows of A, $\alpha_1, ..., \alpha_k$, form a basis of $V_I(f)$. Let $e_i \in F_2^n$ be the vector with the i-th coordinate being one and zero elsewhere. Set $g(x) = f(xA)$. It is easy to check that $e_1, ..., e_k$ form a basis of $V_I(g)$. This means that $g(x)$ is independent of $x_1, ..., x_k$ and hence can be written as $g(x) = g_1(x_{k+1}, ..., x_n)$. Also note that α is a complementary linear structure of $f(x)$ if and only if αA^{-1} is a complementary linear structure of $g(x)$. So the conclusion follows. □

Note that this result is similar to the one in [10]. However here we precisely describe the value of k which is the dimension of $V_I(f)$. The proof here is also simpler.

From theorem 8 we have

Corollary 8 *Let $f(x) \in \mathcal{F}_n$, $V_I(f)$ be the linear space of all the invariant linear structures of $f(x)$. Then $AD(f) = dim(V_I(f))$. Particularly, $f(x)$ is algebraically non-degenerate if and only if it has no nonzero invariant linear structures.*

Corollary 8 gives a relationship between the algebraic degeneration and linear structure characteristics of Boolean functions. We further know that an algebraic non-degenerate function can have at most one complementary linear structure.

Lemma 22 *Let $f(x) \in \mathcal{F}_n$, where α is a complementary linear structure of $f(x)$. Then there exists an $n \times n$ nonsingular matrix D such that $g(x) = f(xD) = x_1 \oplus g_1(x_2, ..., x_n)$, where g_1 has no linear structures. In this case, $f(x)$ is balanced.*

Proof: Let $D = \begin{bmatrix} \alpha \\ D_1 \end{bmatrix}$ be a nonsingular matrix. Then e_1 is a complementary linear structure of $g(x)$ and by theorem 8 $g(x)$ can be written as $x_1 \oplus g_1(x_2, ..., x_n)$. It is easy to verify that $\beta = (0, b_2, ..., b_n)$ is an invariant linear structure of $f(x)$ if and only if $\beta_1 = (b_2, ..., b_n)$ is an invariant linear structure of g_1, and $\beta = (1, b_2, ..., b_n)$ is an invariant linear structure of $f(x)$ if and only if $\beta_1 = (b_2, ..., b_n)$ is a complementary linear structure of g_1. Since $f(x)$ has no invariant linear structures, g_1 must have no linear structures. □

Considering the CI functions without linear structures, from the discussion above it is known that they are algebraic non-degenerate functions which do not have a complementary linear structure. From lemma 22 it is known that those unbalanced CI functions which are algebraic non-degenerate satisfy the requirement, i.e., they do not have linear structures. In the next section we give constructions of algebraic non-degenerate CI functions which can be formulated by the constructions for CI functions having no linear structures.

7.6 Construction of algebraic non-degenerate CI functions

Note that the construction of CI functions discussed above is based on the expression $f(x) = g(xD)$. When D is a square nonsingular matrix, this method is no longer effective. So we need other methods to construct algebraic non-degenerate CI functions.

It is seen that for any $i \in \{1, ..., n\}$ and for any Boolean function $f(x) \in \mathcal{F}_n$, it can be written as $f(x) = x_i f_1(\hat{x}_i) \oplus (1 \oplus x_i) f_2(\hat{x}_i)$, and by lemma 12 we know that $f(x)$ is CI in x_i implies that $W_H(f_1) = W_H(f_2)$. We adopt the result of lemma 9 for the construction of non-degenerate CI functions here.

In order for the method of lemma 9 to be able to construct algebraic non-degenerate CI functions, we need to know when f is algebraic non-degenerate. Denote by $\overline{\omega} = (\omega, \omega_{n+1})$ and $\overline{x} = (x, x_{n+1})$. Then for the functions of (6) we have

$$\begin{aligned} S_f(\overline{\omega}) &= \sum_{\overline{x}} f(\overline{x})(-1)^{\langle \overline{\omega}, \overline{x} \rangle} \\ &= \sum_{x_{n+1}=1} \sum_{x} f_1(x)(-1)^{\langle \omega, x \rangle \oplus \omega_{n+1}} + \sum_{x_{n+1}=0} \sum_{x} f_2(x)(-1)^{\langle \omega, x \rangle} \\ &= (-1)^{\omega_{n+1}} S_{f_1}(\omega) + S_{f_2}(\omega). \end{aligned} \tag{16}$$

It is easy to check that when the dimension of the linear span of $\{\omega : S_{f_1}(\omega) + S_{f_2}(\omega) \neq 0\}$ is n, the dimension of the linear span of $\{\overline{\omega} : S_f(\overline{\omega}) \neq 0\}$ is $n+1$ and hence f is algebraic non-degenerate. So we have

Theorem 9 *Let $f_1(x), f_2(x) \in \mathcal{F}_n$ be two m-th order CI functions with $W_H(f_1) = W_H(f_2)$. If $\prec \omega : S_{f_1}(\omega) + S_{f_2}(\omega) \neq 0 \succ$ forms the whole vector space F_2^n, then $f(x_1, \ldots, x_{n+1}) = x_{n+1} f_1(x) \oplus (1 \oplus x_{n+1}) f_2(x)$ is an algebraic non-degenerate m-th order CI function of $n+1$ variables.*

Theorem 9 gives a sufficient condition for function f defined by (6) to be algebraic non-degenerate. When the condition of theorem 9 can be satisfied is still not clear. It is anticipated that when one or both of f_1 and f_2 are algebraic non-degenerate, f is likely to be so. It is noticed that in the example of the appendix, 96 algebraic non-degenerate CI functions are listed, among them half have Hamming weight 6 and another half have Hamming weight 10. By checking every pair of them with the same Hamming weight we found that among $2 \times \binom{96}{2} = 9120$ pairs, there are 7680 pairs which can form an algebraic non-degenerate CI function of five variables according to (6) while another 1440 pairs cannot.

In practice it is suggested to use the definition to check whether the constructed CI function according to lemma 9 is algebraically non-degenerate. Notice in the proof of theorem 9 that for every $\overline{\omega} = (\omega, \omega_{n+1})$, $(-1)^{\omega_{n+1}} S_{f_1}(\omega) + S_{f_2}(\omega) = 0$ if and only if $S_{f_1}(\omega) + (-1)^{\omega_{n+1}} S_{f_2}(\omega) = 0$. So we have

Corollary 9 *Let $f_1(x), f_2(x) \in \mathcal{F}_n$. Then $x_{n+1} f_1(x) \oplus (1 \oplus x_{n+1}) f_2(x)$ is algebraic non-degenerate if and only if $(1 \oplus x_{n+1}) f_1(x) \oplus x_{n+1} f_2(x)$ is algebraic non-degenerate.*

Let $f(x) \in \mathcal{F}_n$. Now we consider the function $F(\overline{x}) = F(x_1, \ldots, x_{n+1}) = x_{n+1} \oplus f(x)$. It is easy to check that $AD(F) \leq AD(f) + 1$. So $F(\overline{x})$ is algebraic degenerate if $f(x)$ is such. When $f(x)$ is algebraic non-degenerate, the algebraic degeneration of $F(\overline{x})$ is at most one. It is interesting to know when $F(\overline{x})$ is algebraic non-degenerate as well. We have

Theorem 10 *Let $f(x) \in \mathcal{F}_n$ be an algebraic non-degenerate function and $F(\overline{x}) = x_{n+1} \oplus f(x)$. Then $F(\overline{x})$ is algebraic non-degenerate if and only if $f(x)$ has no complement linear structures.*

Proof: Necessity: Assume that $f(x)$ has a complement linear structure α, then $(\alpha, 1)$ is an invariant linear structure of $F(\overline{x})$. By theorem 8, $F(\overline{x})$ is algebraic degenerate.

Sufficiency: If $x_{n+1} \oplus f(x)$ is algebraic, then by corollary 8, $x_{n+1} \oplus f(x)$ must have an invariant linear structure $(a_1, ..., a_{n+1})$. It can easily be verified in this case that $(a_1, ..., a_n)$ is an invariant linear structure of $f(x)$ if $a_{n+1} = 0$ and is a complementary linear structure of $f(x)$ if $a_{n+1} = 1$. □

By theorem 10 and lemma 10 we know that, if $f(x)$ is a balanced algebraic non-degenerate m-th order CI function and has no complement linear structures, then $x_{n+1} \oplus f(x)$ is a balanced algebraic non-degenerate $(m+1)$-th order CI function of $n+1$ variables. Note that this construction cannot be preceded further as $x_{n+1} \oplus f(x)$ has at least one complement linear structure. As an example of this construction, we found that the function

$$f(x_1, \ldots, x_5) = x_1 \oplus x_5 \oplus x_2x_3 \oplus x_3x_4 \oplus x_3x_5 \oplus x_1x_2x_3 \oplus x_1x_2x_4$$

$$\oplus x_1x_2x_5 \oplus x_1x_3x_4 \oplus x_1x_3x_5 \oplus x_2x_3x_5$$

is balanced, algebraic non-degenerate, and 1-st order CI, and has no complement linear structures. Then by theorem 10 and lemma 10 we can construct a Boolean function $x_6 \oplus f(x)$ which is balanced, algebraic non-degenerate, 2-nd order CI, and having only one complementary linear structure (000001).

8 Conclusion

In this paper we have revealed the inherent structure of CI functions and described constructions for such functions. We particularly used the universal form $f(x) = g(xG^T)$, where g is an algebraic non-degenerate function of k variables and G is a generating matrix of an $[n, k]$ linear code. It is also shown that most other cryptographic properties of g, such as *balance, nonlinearity, etc.,* can be inherited by the CI function f. We have studied the constructions of CI functions satisfying at least one more cryptographic property. Based on the study it can naturally be extended for the constructions of CI functions having additional cryptographic properties. Preliminary constructions for algebraic non-degenerate CI functions are also given.

References

[1] R.J.Anderson, Searching for the optimum correlation attacks, *Proc. of K.U.Leuven workshop on Cryptographic Algorithms*, Leuven, Belgium, 1994, pp.56-62.

[2] P.Camion, C. Carlet, P. Charpin, and N. Sendrier, On correlation-immune functions, *Advances in Cryptology, Proc. of CRYPTO'91*, Springer-Verlag 1992, pp.86-100.

[3] P. Camion, A. Canteaut, Construction of t-resilient functions over a finite alphabet, *Advances in Cryptology, Proc. Eurocrypt'96*, Springer-Verlag 1996, pp.283-293.

[4] C. Carlet, Partially-bent functions, *Designs Codes and Cryptography*, Vol.3, 1993, 135-145.

[5] C. Carlet, Generalized Partial Spreads, *IEEE Trans. on Information Theory*, Vol.41, No.5, 1995, pp.1482-1487.

[6] C.Carlet, More correlation-immune and resilient functions over Galois fields and Galois rings, *Advances in Cryptology, Proc. of Eurocrypt'97*, Springer-Verlag 1997, pp.422-433.

[7] C.Ding, G.Xiao, and W.Shan, *The Stability Theory of Stream Ciphers*, Springer-Verlag, 1991.

[8] J.Dj.Golic, On the security of shift register based keystream generators, *Fast Software Encryption (Cambridge'93)*, Springer-Verlag 1994, pp.90-100.

[9] J.Dj.Golic, Correlation properties of a general binary combiner with memory, *Journal of Cryptology,* Vol.9, No.2, 1996, pp.111-126.

[10] X.Lai, Additive and linear structures of cryptographic functions, *Fast Software Encryption*, Springer-Verlag 1995, pp.75-85.

[11] R.L.Lechner, Harmonic Analysis of Switching Functions, in *Recent Developments in switching Theory,* Edited by A.Mukhopadhyay, Academic Press, 1971.

[12] M. Liu, P. Lu, and G. L. Mullen, Correlation-Immune Functions over Finite Fields, *IEEE Trans. on Information Theory*, Vol.IT-44, No.3, May 1998, pp.1273-1275.

[13] S. Lloyd, Counting binary functions with certain cryptographic properties, *Journal of Cryptology,* Vol.5, 1992, pp.107-131.

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

[15] W.Meier and O.Staffelbach, Nonlinearity criteria for cryptographic functions, *Advances in Cryptology, Proc. of Eurocrypt'89*, Springer-Verlag 1990, pp.549-562.

[16] W. Millan, Low Order Approximation of Cipher Functions, in *Cryptography: Policy and Algorithms*, Springer-Verlag, 1996, pp. 144-155.

[17] C. Mitchell, Enumerating Boolean functions of cryptographic significance, *Journal of Cryptology,* Vol.2, No.3, 1990, pp.155-170.

[18] L.J.O'Connor, *An analysis of product ciphers based on the properties of Boolean functions,* Ph.D. thesis, Queensland University of Technology, 1992.

[19] J.Seberry, X.M.Zhang, and Y.Zheng, On constructions and nonlinearity of correlation immune functions (extended abstract), *Advances in Cryptology, Proc. of Eurocrypt'93*, Springer-Verlag 1993, pp.181-197.

[20] T.Siegenthaler, Correlation-immunity of nonlinear combining functions for cryptographic applications, *IEEE Trans. on Infor. Theory,* Vol. IT-30, No.5, 1984, pp.776-780.

[21] T. Siegenthaler, Cryptanalysts' representation of nonlinearly filtered m-sequences, *Advances in Cryptology, Proc. of Eurocrypt'85*, Springer-Verlag 1986, pp.103-110.

[22] T. Siegenthaler, *Methoden fur den Entwurf von Stream Cipher Systemen,* Diss. ETH Nr. 8185, 1986

[23] C.K.Wu, *Boolean functions in Cryptology,* Ph.D. Thesis, Xidian University, 1993.

[24] C.K.Wu, X.M.Wang, and E.Dawson, Construction of correlation immune functions based on the theory of error-correcting codes, *Proc. ISITA96,* Canada, September 1996, pp.167-170.

[25] G.Z.Xiao and J.L.Massey, A spectral characterization of correlation-immune combining functions, *IEEE Trans. Inform. Theory,* Vol. IT-34, 1988, pp.569-571.

[26] X.M.Zhang and Y.Zheng, On nonlinear resilient functions (extended abstract), *Advances in Cryptology, Proc. of Eurocrypt'95*, Springer-Verlag 1995, pp.274-288.

[27] X.M.Zhang and Y.Zheng, Auto-correlations and new bounds on the nonlinearity of boolean functions, *Advances in Cryptology, Proc. of Eurocrypt'96*, Springer-Verlag 1996, pp.294-306.

Appendix:
An example of exhaustive construction

It is not surprising that to accomplish an exhaustive construction of CI functions of n variables is not practical when n is fairly large, even if the method described in section 5 is used. However, as an interesting practice we show here a small example of how all the CI functions are constructed.

We consider the correlation immunity of Boolean functions of $n = 4$ variables. All CI functions will be presented by means of representatives, i.e., their complements and/or variable-permutation equivalences. First of all we know that

$$f(x_1, x_2, x_3, x_4) = c_1x_1 \oplus c_2x_2 \oplus c_3x_3 \oplus c_4x_4$$

is CI of order $W_H(\gamma) - 1$ if $\gamma = (c_1, c_2, c_3, c_4) \neq \mathbf{0}$, or 4 if $\gamma = \mathbf{0}$. Then we consider functions in the form $g(xG^T)$, where g is an algebraic non-degenerate Boolean function of 2 variables and G is a generating matrix of [4, 2] code. It is easy to see that g is algebraic non-degenerate if and only if $deg(g) = 2$, and by lemma 14 such a function is not CI in any linear combination of its variables. All possible representatives of such functions are as follows:

$$\begin{aligned}
&y_1y_2,\\
&y_1y_2 \oplus y_1,\\
&y_1y_2 \oplus y_2,\\
&y_1y_2 \oplus y_1 \oplus y_2.
\end{aligned}$$

In order for the constructed function to be CI of order at least one, the only possible codes useful are [4, 2, 2] codes. Recall that a permutation on the column vectors of matrix G is equivalent to the same permutation performed on the variables of the constructed CI functions. So under column permutation equivalence we have three different linear codes with matrices

$$\begin{bmatrix} 1100 \\ 1010 \end{bmatrix}, \begin{bmatrix} 1100 \\ 0011 \end{bmatrix}, \begin{bmatrix} 1100 \\ 1011 \end{bmatrix}.$$

By corollary 3 we know that all the constructed functions (with 12 representatives) are exactly 1-st order correlation immune. All these functions also have the properties that *algebraic degree = 2, nonlinearity = 4, number of invariant linear structures = 4, number of complement linear structures = 0.*

Now we consider algebraic non-degenerate functions of 3 variables and the family of [4, 3] linear codes. It is known that there are totally $2^{2^3} = 256$ Boolean functions of 3 variables. Among them half are of degree 3 which are algebraic non-degenerate according to corollary 2 (they are useless in constructing CI functions according to corollary 3 because every [4, 3] linear code has a code word with Hamming weight one), and $2^{3+1} = 16$ are affine ones. So only 112 functions are of degree 2 with half are complements of the other. It can be checked that those algebraic degenerate functions can always be written as y_1y_2, $y_1y_2 \oplus y_1$, $y_1y_2 \oplus y_2$ and $y_1y_2 \oplus y_1 \oplus y_2$ and their complements. When y_1 and y_2 are as follows (order is ignored):

$$\begin{cases} y_1 = x_1 \oplus x_2 \\ y_2 = x_3 \end{cases}, \begin{cases} y_1 = x_1 \oplus x_3 \\ y_2 = x_2 \end{cases}, \begin{cases} y_1 = x_2 \oplus x_3 \\ y_2 = x_1 \end{cases}, \begin{cases} y_1 = x_1 \oplus x_2 \\ y_2 = x_2 \oplus x_3 \end{cases},$$

they form 16 algebraic degenerate functions of degree 2. When $y_1 = 1$ while y_2 is any Boolean function of two variables from x_1, x_2, x_3 with degree 2, y_1y_2 has 12 different forms. All together we have 28 algebraic degenerate functions of degree 2 and with constant term 0. So there are 28 algebraic non-degenerate Boolean functions of degree 2 which have constant term 0, namely

$$\begin{gathered} x_1x_2 \oplus \{x_3,\ x_1 \oplus x_3,\ x_2 \oplus x_3,\ x_1 \oplus x_2 \oplus x_3\}, \\ x_1x_3 \oplus \{x_2,\ x_1 \oplus x_2,\ x_2 \oplus x_3,\ x_1 \oplus x_2 \oplus x_3\}, \\ x_2x_3 \oplus \{x_1,\ x_1 \oplus x_2,\ x_1 \oplus x_3,\ x_1 \oplus x_2 \oplus x_3\}, \\ x_1x_2 \oplus x_1x_3 \oplus \{x_2,\ x_3,\ x_1 \oplus x_2,\ x_1 \oplus x_3\}, \\ x_1x_2 \oplus x_2x_3 \oplus \{x_1,\ x_3,\ x_1 \oplus x_2,\ x_2 \oplus x_3\}, \\ x_1x_3 \oplus x_2x_3 \oplus \{x_1,\ x_2,\ x_1 \oplus x_3,\ x_2 \oplus x_3\}, \\ x_1x_2 \oplus x_1x_3 \oplus x_2x_3 \oplus \{0,\ x_1 \oplus x_2,\ x_1 \oplus x_3,\ x_2 \oplus x_3\} \end{gathered} .$$

It is easy to check that no function above is CI. So by theorem 1, in order for the function $g(xG^T)$ to be CI, there are at most 2 linearly independent code words with Hamming weight one. Therefore only the following minimum weight generating matrices of [4, 3] linear codes need to be considered (without being column permutation equivalent):

$$G_1 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \ G_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \ \text{and} \ G_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

Matching the 28 functions above with G_1, we can construct 28 first order CI functions. These functions are actually constructed based on lemma 6 and have been discussed in [24]. By theorem 1 if $g(y)$ is CI in y_1 then $g(xG_1^T)$ is CI of order ≥ 1. Among the above algebraic non-degenerate functions only the following ones are CI in x_1:

$$x_1x_2 \oplus \{x_3,\ x_1 \oplus x_3,\ x_2 \oplus x_3,\ x_1 \oplus x_2 \oplus x_3\}$$
$$x_1x_3 \oplus \{x_2,\ x_1 \oplus x_2,\ x_2 \oplus x_3,\ x_1 \oplus x_2 \oplus x_3\}$$
$$x_1x_2 \oplus x_1x_3 \oplus \{x_2,\ x_3,\ x_1 \oplus x_2,\ x_1 \oplus x_3\}.$$

Matching them with G_2 we can generate 12 1-st order CI functions of 4 variables. By variable permutations more CI functions can be generated. Note that all these functions are not constructible by the methods in [24].

It can also be checked that functions

$$x_1x_2 \oplus \{x_3,\ x_1 \oplus x_3,\ x_2 \oplus x_3,\ x_1 \oplus x_2 \oplus x_3\}$$

are also CI in x_2 as well. Matching with G_3 we can get 4 more 1-st order CI functions of 4 variables which are not constructible by the methods in [24] either. In addition, all of the above constructed functions also have the properties that *algebraic degree = 2, nonlinearity = 4, number of invariant linear structures = 2, number of complement linear structures = 2.*

By computing search we found that there are 192 functions in $\mathcal{F}_4$ which are algebraic non-degenerate and with 1-st order correlation immunity. They also have the properties that *algebraic degree = 3, nonlinearity = 4, number of invariant linear structures = 1, number of complement linear structures = 0, and propagation criterion order = 0.* 96 of them are listed below by truth table expression and the other 96 are just the complements of those in the list.

0001011010011000	0001011010100100	0001011011000010	0001100101101000
0001100110100100	0001100111000010	0001101001100100	0001101010010100
0001101011000001	0001110001100010	0001110010010010	0001110010100001
0010010101101000	0010010110011000	0010010111000010	0010011001011000
0010011010010100	0010011011000001	0010100101011000	0010100101100100
0010100111000001	0010110001010010	0010110001100001	0010110010010001
0011010001001010	0011010010000110	0011010010001001	0011100001000110
0011100001001001	0011100010000101	0011110111011010	0011110111100110
0011110111101001	0011111011010110	0011111011011001	0011111011100101
0100001101101000	0100001110011000	0100001110100100	0100011000111000
0100011010010010	0100011010100001	0100100100111000	0100100101100010
0100100110100001	0100101000110100	0100101001100001	0100101010010001
0101001000101100	0101001010000110	0101001010001001	0101100000100110
0101100000101001	0101100010000011	0101101110111100	0101101111100110
0101101111101001	0101111010110110	0101111010111001	0101111011100011
0110000100101100	0110000101001010	0110000110001001	0110001000011100
0110001001001001	0110001010000101	0110010000011010	0110010000101001
0110010010000011	0110011110111100	0110011111011010	0110011111101001

0110100000011001	0110100000100101	0110100001000011	0110101101111100
0110101111011001	0110101111100101	0110110101111010	0110110110111001
0110110111100011	0110111001111001	0110111010110101	0110111011010011
0111011010011110	0111011010101101	0111011011001011	0111100101101110
0111100110101101	0111100111001011	0111101001101101	0111101010011101
0111101011000111	0111110001101011	0111110010011011	0111110010100111

All the CI functions of 4 variables can be obtained by a variable permutation and/or the complementation of the above constructed functions.

(Received 1/4/99)