Correlation Immune Functions with Controllable Nonlinearity

Seongtaek Chee, Sangjin Lee, Kwangjo Kim, and Daeho Kim

CONTENTS

ABSTRACT

In this paper, we consider the relationship between nonlinearity and correlation immunity of Boolean functions. In particular, we discuss the nonlinearity of correlation immune functions suggested by P. Camion *et al*. For the analysis of such functions, we present a simple method of generating the same set of functions, which makes it possible to construct correlation immune functions with controllable correlation immunity and nonlinearity. Also, we find a bound for the correlation immunity of functions having maximal nonlinearity.

I. INTRODUCTION

Cryptographic Boolean functions play an important role in the design of nonlinear filter functions or nonlinear combiners in stream cipher as well as primitive logics in block ciphers.

In particular, the function whose output leaks no information about its input values is of great importance. Such functions called correlation immune functions were firstly introduced by T. Siegenthaler [1]. Since then the topic has been an active research area [2]-[6] and many stream ciphers have employed the correlation immune functions. P. Camion *et al.* [2] presented a method for constructing balanced correlation immune functions. J. Seberry *et al.* [4] discussed the nonlinearity and propagation characteristics of such functions.

The objective of this paper is to discuss the relationship between correlation immunity and nonlinearity of Boolean functions. In particular, we focus our attention on the functions generated by P. Camion *et al.*'s method. In order to achieve such a goal, we present a simple method of generating the same set of functions, which makes it possible to construct correlation immune functions with controllable correlation immunity and nonlinearity.

The rest of this paper is organized as follows. Section II introduces notations and definitions that are needed in this paper. In Section III, we derive an upper bound for the nonlinearity of correlation immune function. In Section IV, we describe our method for constructing correlation immune functions and discuss some properties of the generated functions that were already analyzed in [4], which is rather complicated compared to ours. We also suggest a condition for obtaining maximal nonlinearity of the functions. In Section V, we present a systematic method to obtain correlation immune functions with controllable nonlinearity and give an example. Section VI describes the relationship between the correlation immunity and nonlinearity of functions discussed in Section V. In particular, we discuss the range of the correlation immunity of functions that have maximal nonlinearity. The conclusions are addressed in Section VII.

II. PRELIMINARIES

Let $\mathcal{Z}_2^n$ be the n-dimensional vector space with the binary n-tuples of elements $x = (x_1, \cdots, x_n)$. For $a = (a_1, \cdots, a_n)$, $b = (b_1, \cdots, b_n)$ in $\mathcal{Z}_2^n$, $a \cdot b = a_1 b_1 \oplus \cdots \oplus a_n b_n$ is the inner product of two vectors.

A function f is said to be balanced if $\#\{x \mid f(x) = 0\} = \#\{x \mid f(x) = 1\}$. A function f on $\mathcal{Z}_2^n$ is k-th order correlation immune$(1 \le k \le n)$ if $f(x)$ is statistically independent of any subset of k input variables $x_{i_1}, \cdots, x_{i_k} (1 \le i_1 < \cdots < i_k \le n)$ and k is called the correlation immunity of f. The algebraic normal form of f is as follows:

$$f(x_1, \cdots, x_n) = a_0 \oplus a_1 x_1 \oplus \cdots \oplus a_n x_n \oplus a_{12} x_1 x_2 \oplus \cdots \oplus a_{n-1,n} x_{n-1} x_n$$

$$\oplus a_{123}x_1x_2x_3 \oplus \cdots$$
$$\oplus a_{n-2,n-1,n}x_{n-2}x_{n-1}x_n$$
$$\vdots$$
$$\oplus a_{12\cdots n}x_1x_2\cdots x_n.$$

The algebraic degree of a Boolean function, denoted by $deg(f)$, is defined as the maximum of the order of its product terms that have a nonzero coefficient in the algebraic normal form. A Boolean function with $deg(f) \le 1$, i.e., $f(x) = a_0 \oplus a_1x_1 \oplus \cdots \oplus a_nx_n$ is said to be affine. In particular, if $a_0 = 0$, it is said to be linear.

For two Boolean functions f and g, we define the distance between f and g by $d(f,g) = \#\{x \mid f(x) \neq g(x)\}$. The minimum distance between f and the set of all affine functions Λ, i.e., $\min_{\lambda\in\Lambda} d(f,\lambda)$ is called the nonlinearity of f and denoted by $\mathcal{N}_f$. In most cases, it will be more convenient to deal with $\hat{f}(x) = (-1)^{f(x)}$ which takes values in $\{-1,1\}$.

The definitions of balancedness, correlation immunity and nonlinearity can be derived from the notions of Walsh-Hadamard transforms.

Definition 1. Let f be a Boolean function in the vector space $\mathcal{Z}_2^n$. The Walsh-Hadamard transform of $\hat{f}$ is the real-valued function $\hat{\mathcal{F}}$ over the vector space $\mathcal{Z}_2^n$ defined as

$$\hat{\mathcal{F}}(w) = \sum_x \hat{f}(x)(-1)^{w\cdot x}.$$

Lemma 1. A Boolean function f is balanced if and only if $\hat{\mathcal{F}}(0) = 0$.

Lemma 2. [6] For a Boolean function f, f is k-th order correlation immune if and only if $\hat{\mathcal{F}}(w) = 0$ holds for any w with $1 \le wt(w) \le k$, where $wt(w)$ is the Hamming weight of w.

Lemma 3. Let f be a Boolean function of n variables. The nonlinearity of f is

$$\mathcal{N}_f = 2^{n-1} - \frac{1}{2}\max_w |\,\hat{\mathcal{F}}(w)\,|\,.$$

Theorem 1(Parseval's Theorem). Let f be a Boolean function of n variables, then

$$\sum_{w\in\mathcal{Z}_2^n} \hat{\mathcal{F}}^2(w) = 2^{2n}.$$

III. CORRELATION IMMUNITY AND NONLINEARITY

It is well-known that the correlation immunity of a function on $\mathcal{Z}_2^n$ and its algebraic degree d are constrained by the relation $k + d \le n$ [1]. Naturally, we can imagine that there may be a similar relationship between the correlation immunity and nonlinearity.

In this section, we derive an upper bound for the nonlinearity of the correlation immune functions.

Lemma 4. If $f : \mathcal{Z}_2^n \to \mathcal{Z}_2$, then

$$\mathcal{N}_f \le 2^{n-1} - \frac{2^{n-1}}{\sqrt{\eta}},$$

where $\eta = \#\{w \in \mathcal{Z}_2^n \mid \hat{\mathcal{F}}(w) \neq 0\}$.

Proof. By Theorem 1, we have

$$\eta \cdot \max_w |\,\hat{\mathcal{F}}(w)\,|^2 \ge \sum_w \hat{\mathcal{F}}^2(w) = 2^{2n}.$$

Hence, $\max_w |\hat{\mathcal{F}}(w)| \geq \frac{2^n}{\sqrt{\eta}}$. Therefore, by Lemma 3, we have

$$\begin{aligned}\mathcal{N}_f &= 2^{n-1} - \frac{1}{2}\max_w |\hat{\mathcal{F}}(w)| \\ &\leq 2^{n-1} - \frac{2^{n-1}}{\sqrt{\eta}}.\end{aligned}$$

□

Theorem 2. If $f : \mathcal{Z}_2^n \to \mathcal{Z}_2$ is a k-th order correlation immune function, then

$$\mathcal{N}_f \leq 2^{n-1} - \frac{2^{n-1}}{\sqrt{\zeta_k}},$$

where $\zeta_k = 2^n - \left\{\binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{k}\right\}$.

Proof. Since f is k-th order correlation immune, by Lemma 2, $\hat{\mathcal{F}}(w) = 0$ holds for any w with $1 \leq wt(w) \leq k$. Hence

$$\begin{aligned}\eta &= 2^n - \#\{w \in \mathcal{Z}_2^n \mid \hat{\mathcal{F}}(w) = 0\} \\ &\leq 2^n - \#\{w \in \mathcal{Z}_2^n \mid 1 \leq wt(w) \leq k\} \\ &= 2^n - \left\{\binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{k}\right\} \\ &= \zeta_k.\end{aligned}$$

Therefore, the assertion holds by Lemma 4. □

If f is balanced and k-th order correlation immune, then $\hat{\mathcal{F}}(0) = 0$, i.e., $\eta \leq \zeta_k - 1$. Hence we have the following:

Corollary 1. *For a balanced and k-th order correlation immune function $f : \mathcal{Z}_2^n \to \mathcal{Z}_2$,*

$$\mathcal{N}_f \leq 2^{n-1} - \frac{2^{n-1}}{\sqrt{\zeta_k - 1}}.$$

IV. DESIGN OF CORRELATION IMMUNE FUNCTIONS

Throughout this section, we set the following notation:

n	an integer with $n \geq 4$
k	an integer with $1 \leq k \leq n-3$
m	an integer with $1 \leq m < n-k$
ϕ	a function on $\mathcal{Z}_2^m$ into $\mathcal{Z}_2^{n-m}$ with $wt(\phi(y)) \geq k+1$ for all $y \in \mathcal{Z}_2^m$
t_x	$\#\ \phi^{-1}(x),\ x \in \mathcal{Z}_2^{n-m}$
t	$\max_x t_x$
A_y	$\phi(y)$.

Now we present a method of constructing correlation immune functions.

Theorem 3. Define a Boolean function $f : \mathcal{Z}_2^n \to \mathcal{Z}_2$ by

$$f(y, x) = A_y \cdot x, \tag{1}$$

where $y = (y_1, \cdots, y_m) \in \mathcal{Z}_2^m$, $x = (x_1, \cdots, x_{n-m}) \in \mathcal{Z}_2^{n-m}$. Then the followings hold:

(i) f is balanced.

(ii) f is k-th order correlation immune.

(iii) $\mathcal{N}_f = 2^{n-1} - t2^{n-m-1}$.

(iv) Let $A_y(i)$ be the i-th component of A_y. If $\oplus_y A_y(i) = 1$ for some $i (1 \leq i \leq n-m)$, then $deg(f) = m+1$.

Proof.

(i) Since $wt(A_y) \geq k+1$, we have $A_y \neq 0$. Thus $\sum_x (-1)^{A_y \cdot x} = 0$. Therefore, we have

$$\hat{\mathcal{F}}(0)=\sum_{y,x}(-1)^{f(y,x)}=\sum_{y,x}(-1)^{A_y\cdot x}$$
$$=\sum_{y}\sum_{x}(-1)^{A_y\cdot x}=0.$$

By Lemma 1, f is balanced.

(ii) For any $(b,a)\in\mathcal{Z}_2^n$ with $1\le wt(b,a)\le k$, note that

$$\hat{\mathcal{F}}(b,a)=\sum_{y,x}(-1)^{f(y,x)}(-1)^{(b,a)\cdot(y,x)}$$
$$=\sum_{y,x}(-1)^{A_y\cdot x}(-1)^{b\cdot y\oplus a\cdot x}$$
$$=\sum_{y}(-1)^{b\cdot y}\sum_{x}(-1)^{(A_y\oplus a)\cdot x}. \quad (2)$$

Since $0\le wt(a)\le k$ and $wt(A_y)\ge k+1$, we have $a\oplus A_y\neq 0$. Thus $\sum_x(-1)^{(A_y\oplus a)\cdot x}=0$. Therefore, by (2), we have $\hat{\mathcal{F}}(b,a)=0$. By Lemma 2, f is k-th order correlation immune.

(iii) By (2), we know that

$$\hat{\mathcal{F}}(b,a)=\sum_{y}(-1)^{b\cdot y}\sum_{x}(-1)^{(A_y\oplus a)\cdot x}$$
$$=2^{n-m}\sum_{\{y|A_y=a\}}(-1)^{b\cdot y}. \quad (3)$$

Hence we have

$$\max_{b,a}|\hat{\mathcal{F}}(b,a)|=\max_{b=0,a}|\hat{\mathcal{F}}(b,a)|=t\cdot 2^{n-m}.$$

By Lemma 3, $\mathcal{N}_f=2^{n-1}-t\cdot 2^{n-m-1}$.

(iv) We note that

$$f(y,x)=(y_1\oplus 1)(y_2\oplus 1)\cdots(y_m\oplus 1)A_0\cdot x$$
$$\oplus(y_1\oplus 1)(y_2\oplus 1)\cdots y_mA_1\cdot x$$
$$\vdots$$
$$\oplus y_1y_2\cdots y_mA_{2^m-1}\cdot x.$$

If $\oplus_y A_y(i)=1$, then in the above expression, the term $y_1y_2\cdots y_mx$ is not cancelled. Hence $deg(f)=m+1$. □

Corollary 2. For any $r:\mathcal{Z}_2^m\to\mathcal{Z}$, if we define $f(y,x)=A_y\cdot x\oplus r(y)$, then properties (i), (ii) and (iv) of Theorem 3 hold. And the nonlinearity of f is bounded by

$$\mathcal{N}_f\le 2^{n-1}-2^{n-m-1}, \quad (4)$$

where equality holds if and only if ϕ is injective, i.e., t=1.

Proof. We only prove (4), since the proofs of rests are trivial. By (3), we have

$$\hat{\mathcal{F}}(b,a)=2^{n-m}\sum_{\{y|A_y=a\}}(-1)^{b\cdot y\oplus r(y)}. \quad (5)$$

Then $\max_{b,a}|\hat{\mathcal{F}}(b,a)|\ge 2^{n-m}$ and $\mathcal{N}_f\le 2^{n-1}-2^{n-m-1}$.

Suppose ϕ is injective, then clearly we have $\mathcal{N}_f=2^{n-1}-2^{n-m-1}$ by (5). Conversely, assume that ϕ is not injective and $\max_{b,a}|\hat{\mathcal{F}}(b,a)|=2^{n-m}$. Then by Parseval's Theorem, we have

$$2^{2n}=\sum_{b,a}\hat{\mathcal{F}}^2(b,a)$$
$$=\sum_{b}\sum_{a\in im(\phi)}\hat{\mathcal{F}}^2(b,a)$$
$$<2^m2^m2^{2n-2m}=2^{2n}.$$

This is a contradiction. Therefore, equality in (4) holds if and only if ϕ is injective. □

Similar results are studied in [4]. The main advantage of our method is that it is simple enough to analyze the relationship

between correlation immunity and nonlinearity. And, we derive the exact nonlinearity, while only a lower bound was given in [4]. This fact drives us to study under what conditions we can make the nonlinearity maximal.

For convenience, we denote by $\mathcal{C}_n^m(k)$ the set of Boolean functions generated by Theorem 3^l.

The following lemma is useful to find conditions for maximal nonlinearity of a function in $\mathcal{C}_n^m(k)$.

Lemma 5. For given positive integers n and $k(n \geq 4, 1 \leq k \leq n-3)$, and any positive integer t, let l_t be the smallest l such that

$$t\left\{\binom{l}{k+1}+\binom{l}{k+2}+\cdots+\binom{l}{l}\right\} \geq 2^{n-l} .$$

Then we have $2^{l_1} \leq t \cdot 2^{l_t}$, i.e., $\min\{t \cdot 2^{l_t} | t = 1,2,\cdots\} = 2^{l_1}$.

To prove Lemma 5, we need some lemmas. Lemma 6 is well-known and we omit its proof.

Lemma 6. For positive integers n and k, the following holds.

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}.$$

Lemma 7. For positive integers n and k, the following holds.

$$2\left\{\binom{n}{k+1}+\binom{n}{k+2}+\cdots+\binom{n}{n}\right\} \leq \binom{n+1}{k+1}+\binom{n+1}{k+2}+\cdots+\binom{n+1}{n+1}.$$

Proof. By Lemma 6, we have

$$\begin{aligned}
&\binom{n+1}{k+1}+\binom{n+1}{k+2}+\cdots+\binom{n+1}{n+1} \\
&=\left\{\binom{n}{k+1}+\binom{n}{k}\right\}+\left\{\binom{n}{k+2}+\binom{n}{k+1}\right\}+\cdots \\
&\quad+\left\{\binom{n}{n}+\binom{n}{n-1}\right\}+1 \\
&=\left\{\binom{n}{k+1}+\binom{n}{k+2}+\cdots+\binom{n}{n}\right\} \\
&\quad+\left\{\binom{n}{k}+\binom{n}{k+1}+\cdots+\binom{n}{n-1}\right\}+1 \\
&=2\left\{\binom{n}{k+1}+\binom{n}{k+2}+\cdots+\binom{n}{n}\right\}+\binom{n}{k}. \quad \square
\end{aligned}$$

Lemma 8. If $l \geq l_t$, then

$$t\left\{\binom{l}{k+1}+\binom{l}{k+2}+\cdots+\binom{l}{l}\right\} \geq 2^{n-l}$$

where l_t is the value defined in Lemma 5.

Proof. By the definition of l_t,

$$t\left\{\binom{l_t}{k+1}+\binom{l_t}{k+2}+\cdots+\binom{l_t}{l_t}\right\} \geq 2^{n-l_t}.$$

Since $l \geq l_t$, we have

$$\begin{aligned}
&t\left\{\binom{l}{k+1}+\binom{l}{k+2}+\cdots+\binom{l}{l_t}+\cdots+\binom{l}{l}\right\} \\
&\geq t\left\{\binom{l_t}{k+1}+\binom{l_t}{k+2}+\cdots+\binom{l_t}{l_t}\right\} \\
&\geq 2^{n-l_t} \geq 2^{n-l}. \quad \square
\end{aligned}$$

Now, we are ready to prove Lemma 5.

Proof of Lemma 5. If $t = 1$, then $t \cdot 2^{l_t} = 2^{l_1}$. So let's show that $t \cdot 2^{l_t} \geq 2^{l_1}$ when $t \geq 2$. If $t \geq 2$, then there is p $(p \geq 1)$ such that $2^p \leq t < 2^{p+1}$. If $l_1 - p \leq l_t$, then

$$t \cdot 2^{l_t} \geq t \cdot 2^{l_1-p} \geq 2^p 2^{l_1-p} = 2^{l_1}.$$

Hence if $l_1 - p \leq l_t$, the proof is completed. It remains to show that the case $l_t < l_1 - p$

can not happen. Suppose $l_t < l_1 - p$, i.e., $l_t \le l_1 - p - 1$. Then, by Lemma 8,

$$t\left\{\binom{l_1-p-1}{k+1}+\binom{l_1-p-1}{k+2}+\cdots+\binom{l_1-p-1}{l_1-p-1}\right\} \ge 2^{n-(l_1-p-1)}. \quad (6)$$

Also, by Lemma 7,

$$\begin{aligned} &t\left\{\binom{l_1-p-1}{k+1}+\binom{l_1-p-1}{k+2}+\cdots+\binom{l_1-p-1}{l_1-p-1}\right\} \\ &\le 2^{p+1}\left\{\binom{l_1-p-1}{k+1}+\binom{l_1-p-1}{k+2}+\cdots+\binom{l_1-p-1}{l_1-p-1}\right\} \\ &\le 2^{p}\left\{\binom{l_1-p}{k+1}+\binom{l_1-p}{k+2}+\cdots+\binom{l_1-p}{l_1-p}\right\}. \quad (7) \end{aligned}$$

By applying Lemma 7 to (7), we obtain

$$\begin{aligned} &t\left\{\binom{l_1-p-1}{k+1}+\binom{l_1-p-1}{k+2}+\cdots+\binom{l_1-p-1}{l_1-p-1}\right\} \\ &\le 2\left\{\binom{l_1-1}{k+1}+\binom{l_1-1}{k+2}+\cdots+\binom{l_1-1}{l_1-1}\right\}. \end{aligned}$$

Hence by (6),

$$2\left\{\binom{l_1-1}{k+1}+\binom{l_1-1}{k+2}+\cdots+\binom{l_1-1}{l_1-1}\right\} \ge 2^{n-(l_1-p-1)}.$$

Since $p \ge 1$, we have

$$\begin{aligned} \binom{l_1-1}{k+1}+\binom{l_1-1}{k+2}+\cdots+\binom{l_1-1}{l_1-1} &\ge 2^{n-l_1+p} \\ &\ge 2^{n-(l_1-1)}. \end{aligned}$$

By the definition of l_1, $l_1 \le l_1 - 1$. But this is a contradiction. □

The following theorem is one of the major results in this paper.

Theorem 4. For $f \in \mathcal{C}_n^m(k)$, the maximal nonlinearity of f is $\mathcal{N}_f = 2^{n-1} - 2^{l_1-1}$ and it can be obtained if $m = n - l_1$, $t = 1$, where l_1 is the value defined in Lemma 5.

Proof. By the definition of A_y, t and m satisfy the following inequality:

$$t\left\{\binom{n-m}{k+1}+\binom{n-m}{k+2}+\cdots+\binom{n-m}{n-m}\right\} \ge 2^m. \quad (8)$$

In (8), if we substitute $n-m$ with l, then

$$t\left\{\binom{l}{k+1}+\binom{l}{k+2}+\cdots+\binom{l}{l}\right\} \ge 2^{n-l} \quad (9)$$

and $\mathcal{N}_f = 2^{n-1} - t \cdot 2^{l-1}$ by Theorem 3-(iii). Hence for each t $(t = 1, 2, \cdots)$, the maximum nonlinearity is obtained if l is the smallest value satisfying (9). That is, $\max_m \mathcal{N}_f = \max_l \mathcal{N}_f = 2^{n-1} - t \cdot 2^{l_t-1}$. Therefore, by Lemma 5,

$$\begin{aligned} &\max_{m,t} \mathcal{N}_f \\ &= \max_{l,t} \mathcal{N}_f = 2^{n-1} - \min_t t \cdot 2^{l_t-1} \\ &= 2^{n-1} - 2^{l_1-1}. \end{aligned}$$

□

V. CONSTRUCTION OF CORRELATION IMMUNE FUNCTIONS WITH CONTROLLABLE NONLINEARITY

In this section, by using Theorem 4, we suggest a method for constructing correlation immune functions with controllable nonlinearity .

Method for constructing k-th order correlation immune functions with nonlinearity $\mathcal{N}_f = 2^{n-1} - 2^{l_1-1}$

Input. n ($n \geq 4$; the number of input variables of Boolean function),

k ($1 \leq k \leq n-3$; correlation immunity)

Step 1. For $k+1 \leq l \leq n$, find the smallest l satisfying

$$\binom{l}{k+1} + \binom{l}{k+2} + \cdots + \binom{l}{l} \geq 2^{n-l} \quad (10)$$

and call this value l_1.

Step 2. Choose 2^{n-l_1} vectors $A_0, A_1, \cdots, A_{2^{n-l_1}-1}$ in $\mathcal{Z}_2^{l_1}$ with weight greater than or equal to $k+1$.

Step 3. Define $f : \mathcal{Z}_2^n \to \mathcal{Z}_2$ by

$$f(y_1, \cdots, y_{n-l_1}, x_1, \cdots, x_{l_1}) = f(y, x) = A_y \cdot x.$$

We now discuss the above method step by step. First, for the input, since a function with algebraic degree 0 is constant, it is not balanced. And a function with algebraic degree 1, i.e., an affine function, is of nonlinearity 0. Moreover, the sum of correlation immunity k and algebraic degree d for a function f is less than or equal to the number of input variables n and in particular if f is balanced, $k + d \leq n-1$ [1]. Hence, for a balanced, nonlinear and correlation immune function, we have

$$k + d \leq n-1, \; k \geq 1, \; d \geq 2.$$

That is, $1 \leq k \leq n-1-d \leq n-3$, where the smallest n is 4.

The second, in Step 1, by the above, $k+2 \leq n-1$. Hence

$$\binom{n-1}{k+1} + \binom{n-1}{k+2} + \cdots + \binom{n-1}{n-1} \geq \binom{n-1}{k+1} + \binom{n-1}{k+2} \geq 2.$$

Hence we can find $l_1 (l_1 \leq n-1)$ satisfying (10) in Step 1.

The third, in Step 2, the number of vectors in $\mathcal{Z}_2^{l_1}$ with weight greater than or equal to $k+1$ is

$$\binom{l_1}{k+1} + \binom{l_1}{k+2} + \cdots + \binom{l_1}{l_1}$$

and it is greater than or equal to 2^{n-l_1} by Step 1. Hence we can choose 2^{n-l_1} vectors with weight greater than or equal to $k+1$.

The function defined in Step 3 fulfills the requirements of Theorem 4. So the balanced k-th order correlation immune function in Step 3 has nonlinearity $\mathcal{N}_f = 2^{n-1} - 2^{l_1-1}$.

Example 1. We construct a balanced and nonlinear 1st order correlation immune function $f : \mathcal{Z}_2^7 \to \mathcal{Z}_2$.

Input: $n = 7, k = 1$

Step 1: Since $\binom{3}{2} + \binom{3}{3} \not\geq 2^{7-3}$ and $\binom{4}{2} + \binom{4}{3} + \binom{4}{4} \geq 2^{7-4}$, we have $l_1 = 4$.

Step 2: Choose 8 vectors $A_i \in \mathcal{Z}_2^4$ with $wt(A_i) \geq 2$, say

$$A_0 = (1,1,0,0) \quad A_1 = (1,0,1,0)$$

$$A_2=(1,0,0,1) \quad A_3=(0,1,1,0)$$
$$A_4=(0,1,0,1) \quad A_5=(0,0,1,1)$$
$$A_6=(1,1,1,0) \quad A_7=(1,1,0,1).$$

Step 3: Define $f:\mathcal{Z}_2^7 \to \mathcal{Z}_2$ as follows.

$$\begin{aligned} f(y,x)=A_y\cdot x &= (y_1\oplus 1)(y_2\oplus 1)(y_3\oplus 1)(x_1\oplus x_2)\\ &\oplus(y_1\oplus 1)(y_2\oplus 1)y_3(x_1\oplus x_3)\\ &\oplus(y_1\oplus 1)y_2(y_3\oplus 1)(x_1\oplus x_4)\\ &\oplus(y_1\oplus 1)y_2y_3(x_2\oplus x_3)\\ &\oplus y_1(y_2\oplus 1)(y_3\oplus 1)(x_2\oplus x_4)\\ &\oplus y_1(y_2\oplus 1)y_3(x_3\oplus x_4)\\ &\oplus y_1y_2(y_3\oplus 1)(x_1\oplus x_2\oplus x_3)\\ &\oplus y_1y_2y_3(x_1\oplus x_2\oplus x_4). \end{aligned}$$

Then f is balanced, 1st-order correlation immune with $\mathcal{N}_f=2^6-2^3=56$. And, since $\oplus_y A_y(1)=1$, $\oplus_y A_y(2)=1$, $\oplus_y A_y(3)=0$, and $\oplus_y A_y(4)=0$ by Theorem 3-(iv), $\deg(f)=n-l_1+1=4$.

VI. NONLINEARITY AND CORRELATION IMMUNITY OF $\mathcal{C}_n^m(k)$

We now discuss the relationship between correlation immunity and nonlinearity of functions in $\mathcal{C}_n^m(k)$.

Lemma 9. Let n, k, l_1 be given as Lemma 5. Then $l_1 \ge \left\lfloor \frac{n}{2} \right\rfloor + 1$.

Proof. By the definition of l_1,

$$\binom{l_1}{k+1}+\binom{l_1}{k+2}+\cdots+\binom{l_1}{l_1} \ge 2^{n-l_1}.$$

Thus

$$2^{l_1}-2^{n-l_1} \ge \binom{l_1}{0}+\binom{l_1}{1}+\cdots+\binom{l_1}{k}. \quad (11)$$

Since the right hand side of (11) is positive, $l_1 > n-l_1$. In all, $l_1 \ge \left\lfloor \frac{n}{2} \right\rfloor + 1$. □

Lemma 10. The necessary and sufficient condition for two integers n and x to satisfy the following equation is $x \le \left\lfloor \frac{n+1}{2} \right\rfloor$.

$$\binom{n}{x+1}+\binom{n}{x+1}+\cdots+\binom{n}{n} \ge 2^{n-1}. \quad (12)$$

Proof. If n is even, then

$$\binom{n}{\frac{n}{2}}+\binom{n}{\frac{n}{2}+1}+\cdots+\binom{n}{n} > 2^{n-1},$$
$$\binom{n}{\frac{n}{2}+1}+\binom{n}{\frac{n}{2}+2}+\cdots+\binom{n}{n} < 2^{n-1}.$$

So, (12) holds if $x \le \left\lfloor \frac{n+1}{2} \right\rfloor$. And, if n is odd, then

$$\binom{n}{\frac{n+1}{2}}+\binom{n}{\frac{n+1}{2}+1}+\cdots+\binom{n}{n} = 2^{n-1}.$$

So, (12) holds if $x \le \left\lfloor \frac{n+1}{2} \right\rfloor$. □

Theorem 5. For a Boolean function f in $\mathcal{C}_n^m(k)$, we have $\mathcal{N}_f \le 2^{n-1}-2^{\lfloor \frac{n}{2} \rfloor}$, and the equality holds if and only if k satisfies the following:

$$\begin{cases} \binom{\frac{n}{2}+1}{k+1}+\binom{\frac{n}{2}+1}{k+2}+\cdots+\binom{\frac{n}{2}+1}{\frac{n}{2}+1} \ge 2^{\frac{n}{2}-1} \\ \qquad\qquad \text{if } n \text{ is even}, \\ k \le \left\lfloor \frac{n}{4} \right\rfloor \qquad \text{if } n \text{ is odd}. \end{cases}$$

Proof. By Lemma 9, since $l_1 \geq \left\lfloor \frac{n}{2} \right\rfloor + 1$, we have

$$\mathcal{N}_f = 2^{n-1} - 2^{l_1-1} \leq 2^{n-1} - 2^{\lfloor \frac{n}{2} \rfloor},$$

and the equality holds if $l_1 = \left\lfloor \frac{n}{2} \right\rfloor + 1$. Hence by the definition of l_1, equality holds if the followings are satisfied:

$$\binom{\lfloor \frac{n}{2} \rfloor + 1}{k+1} + \binom{\lfloor \frac{n}{2} \rfloor + 1}{k+2} + \cdots + \binom{\lfloor \frac{n}{2} \rfloor + 1}{\lfloor \frac{n}{2} \rfloor + 1} \geq 2^{n-(\lfloor \frac{n}{2} \rfloor + 1)}, \quad (13)$$

$$\binom{\lfloor \frac{n}{2} \rfloor}{k+1} + \binom{\lfloor \frac{n}{2} \rfloor}{k+2} + \cdots + \binom{\lfloor \frac{n}{2} \rfloor}{\lfloor \frac{n}{2} \rfloor} < 2^{n-\lfloor \frac{n}{2} \rfloor}. \quad (14)$$

Since (14) holds for any k, equality holds if and only if (13) holds. If n is odd, since the right hand side of (13) is $2^{\lfloor \frac{n}{2} \rfloor}$, by Lemma 10, (13) holds if and only if

$$k+1 \leq \left\lfloor \frac{\lfloor \frac{n}{2} + 1 + 1 \rfloor}{2} \right\rfloor = \left\lfloor \frac{n-1}{4} \right\rfloor + 1 = \left\lfloor \frac{n}{4} \right\rfloor + 1$$

i.e., $k \leq \left\lfloor \frac{n}{4} \right\rfloor$. □

By Theorem 5, we can construct $\left\lfloor \frac{n}{4} \right\rfloor$-th order correlation immune functions with nonlinearity $2^{n-1} - 2^{\lfloor \frac{n}{2} \rfloor}$ if n is odd. If n is even, we have the following:

Corollary 3. If n is even and $\mathcal{N}_f = 2^{n-1} - 2^{\lfloor \frac{n}{2} \rfloor}$, the approximative upper bound of the correlation immunity k in Theorem 5 is

$$k \leq \left\lfloor \frac{n}{4} + 0.335\sqrt{\frac{n}{2} + 1} \right\rfloor. \quad (15)$$

The proof of Corollary 3 is left to the appendix.

In fact, for $n = 4, 6, \cdots, 98, 100$, if a function has nonlinearity $\mathcal{N}_f = 2^{n-1} - 2^{\frac{n}{2}}$, which is the maximum among functions in $\mathcal{C}_n^m(k)$, the range of correlation immunity given in Corollary 3 is actually correct except only for the case $n = 38$. In that case, (15) gives $k \leq 10$, where, in fact, $k \leq 11$.

It is natural to question what is the range of k if the nonlinearity of function in $\mathcal{C}_n^m(k)$ is fixed. The following corollary, which can be verified similarly to Corollary 3, solves the problem.

Corollary 4. If the function in $\mathcal{C}_n^m(k)$ has nonlinearity $\mathcal{N}_f = 2^{n-1} - 2^{l_1-1}$, then the approximate range of the correlation immunity k is as follows:

$$\left\lfloor \frac{l_1 - 1}{2} + \frac{1}{2} + \frac{\sqrt{l_1 - 1}}{2} z_{2^{n-2(l_1-1)}} \right\rfloor \leq k \leq \left\lfloor \frac{l_1 - 1}{2} + \frac{1}{2}\sqrt{l_1} z_{2^{n-2l_1}} \right\rfloor,$$

where $P(Z \geq z_\alpha) = \alpha$ and $Z \sim N(0,1)$.

For the case where a function in $\mathcal{C}_n^m(k)$ has maximum correlation immunity, we have the following:

Corollary 5. Let $k = n - 3$. Then the maximum nonlinearity of Boolean function $f \in \mathcal{C}_n^m(k)$ is $\mathcal{N}_f = 2^{n-1} - 2^{n-2} = 2^{n-2}$.

Figure 1 represents the relationship between correlation immunity and nonlinearity of functions in $\mathcal{C}_n^m(k)$ for $n = 22$. Since the nonlinearity is too big to investigate its behavior precisely as k increases, we present the relationship between the correlation immunity and $l_1 - 1$, which determines the nonlinearity in Fig. 2 as a solid line. In Fig. 2, the dotted curve, which is derived

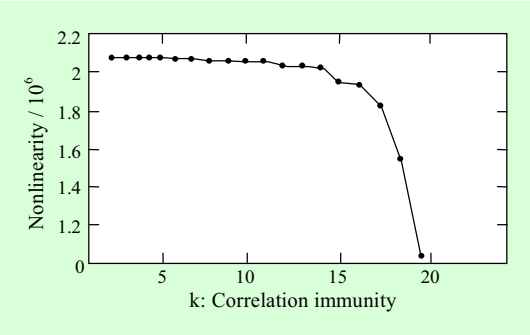

Fig. 13. Relationship between nonlinearity and correlation immunity.

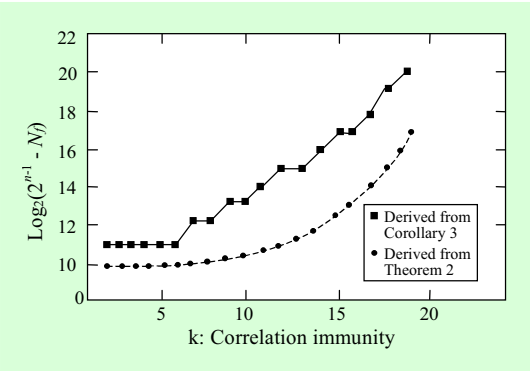

Fig. 14. Relationship between $l_1 - 1 = \log_2(2^{n-1} - \mathcal{N}_f)$ and correlation immunity.

from Theorem 2 is concerned with 'general' correlation immune functions. We notice that the nonlinearity does not decrease until the correlation immunity k reaches about $\lfloor n/4 \rfloor$.

From the above two corollaries, we can construct Boolean functions with controllable nonlinearity and correlation immunity. Even though the range in the Corollaries 3 and 4 are approximative, the error range was verified as small enough by our simulations.

VII. CONCLUSIONS

The main results of this paper are associated with studying the relationship between correlation immunity and nonlinearity of functions in [2], [4]. Such results were possible by making the generating method simple and clear enough to discuss several properties. This paper may provide us with a new avenue towards studying the relationship between correlation immunity and nonlinearity.

APPENDIX

Proof of Corollary 3. If a random variable X has the binomial distribution with parameters n and $\frac{1}{2}$, i.e., $X \sim b\left(n, \frac{1}{2}\right)$, then

$$P(X \geq x) = \sum_{k=x}^{n} \binom{n}{k} \left(\frac{1}{2}\right)^n.$$

Hence the condition for Theorem 5 holds if

$$P(X \geq k+1) \geq \frac{1}{4},$$

where $X \sim b\left(\frac{n}{2}+1, \frac{1}{2}\right)$. Then the following holds if $X \sim b\left(\frac{n}{2}+1, \frac{1}{2}\right)$ and $\frac{n}{2}+1$ is large enough.[1]

$$P(X \leq k) \simeq P\left(Z \leq \frac{k - \frac{\frac{n}{2}+1}{2} + \frac{1}{2}}{\frac{1}{2}\sqrt{\frac{n}{2}+1}}\right),$$

[1]In general, we assume that $\left(\frac{n}{2}+1\right)\frac{1}{2} > 5$, i.e., $\frac{n}{2}+1 > 10$.

where $Z \sim N(0,1)$. Since $P(Z \le 0.67) = 0.7486 \simeq \frac{3}{4}$, we have

$$\frac{k - \frac{\frac{n}{2}+1}{2} + \frac{1}{2}}{\frac{1}{2}\sqrt{\frac{n}{2}+1}} \le 0.67.$$

Since $k \le \frac{n}{4} + 0.335\sqrt{\frac{n}{2}+1}$, we have

$$k \le \left\lfloor \frac{n}{4} + 0.335\sqrt{\frac{n}{2}+1} \right\rfloor.$$

ACKNOWLEDGMENT

The authors are grateful to anonymous referees for their comments. We also give special thanks to Dr. Soo Hak Sung, Sung Mo Park and Dr. Jung Hee Cheon for helpful discussions.

REFERENCES

[1] T. Siegenthaler, "Correlation immunity of nonlinear combining functions for cryptographic applications," *IEEE Trans. on Inf. Th.*, vol. 30, pp. 776–780, 1984.

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

[3] W. Meier and O. Staffelbach, "Nonlinearity criteria for cryptographic functions," *Advances in Cryptology - EUROCRYPT'89*, Springer-Verlag, pp. 549–562, 1990.

[4] J. Seberry, X. M. Zhang, and Y. Zheng, "On constructions and nonlinearity of correlation immune functions," *Advances in Cryptology - EUROCRYPT'93*, Springer-Verlag, pp. 181–199, 1994.

[5] Y. Xian, "Correlation-immunity of Boolean functions," *Electronics Letters*, vol. 23, pp. 1335–1336, 1987.

[6] G. Xiao and J. Massey, "A spectral characterization of correlation-immune combining functions," *IEEE Trans. on Inf. Th.*, vol. 34, pp. 569–571, 1988.

Seongtaek Chee received B.S. and M.S. degrees in mathematics from Sogang University in 1985 and 1987, respectively. Since 1989 he has been on the research staff at ETRI, where he is currently a senior member of Coding Technology Department. His research interests are cryptographic functions.

Sangjin Lee received Ph. D. in mathematics from Korea University, Seoul, Korea in 1994. He joined ETRI in 1989, where he is currently working as a senior member of Coding Technology Department. His research interests include finite filed theory, cryptography, and cryptanalysis.

Kwangjo Kim received the B.S. and M.S. degrees in electronic engineering from Yonsei University in 1980 and 1983 respectively. He also received the Ph.D degree in electrical and information engineering from Yokohama National University, Japan in 1991. Since 1979 he has been with ETRI and is currently working at Department of Coding Technology. He served as a program co-chair of Asiacrypt'96 conference which was held in Kyongju, Nov. 3∼7, 1996. His current interests include

information security, cryptology and microwave communications. He is now a member of IEEE, IACR and IEICE and a director of KIISC.

Daeho Kim received M.S. degree in electrical engineering from Hanyang University in 1984. He joined ETRI in 1977, where he is currently working as Director of Coding Technology Department. His research interests include coding theory, information security, and cryptology. He is a regular member of the Korean Communication Professional Engineer Association.