JOURNAL OF THE
CHUNGCHEONG MATHEMATICAL SOCIETY
Volume **17**, No.1, April 2004

# AUTO-CORRELATIONS AND BOUNDS ON THE NONLINEARITY OF VECTOR BOOLEAN FUNCTIONS

Wansoon Kim* and Junseok Park**

Abstract. The nonlinearity of a Boolean function $f$ on $GF(2)^n$ is the minimum hamming distance between $f$ and all affine functions on $GF(2)^n$ and it measures the ability of a cryptographic system using the functions to resist against being expressed as a set of linear equations. Finding out the exact value of the nonlinearity of given Boolean functions is not an easy problem therefore one wants to estimate the nonlinearity using extra information on given functions, or wants to find a lower bound or an upper bound on the nonlinearity. In this paper we extend the notion of auto-correlations of Boolean functions to vector Boolean functions and obtain upper bounds and a lower bound on the nonlinearity of vector Boolean functions in the context of their auto-correlations. Also we can describe avalanche characteristics of vector Boolean functions by examining the extended notion of auto-correlations.

## 1. Introduction

The nonlinearity of a Boolean function $f$ on $GF(2)^n$ is the minimum hamming distance between $f$ and all affine functions on $GF(2)^n$ and it measures the ability of a cryptographic system using the functions to resist against being expressed as a set of linear equations. Finding out the exact value of the nonlinearity of given Boolean functions is not an easy problem therefore one wants to estimate the nonlinearity using extra information on given functions, or wants to find a lower bound or an upper bound on the nonlinearity when the exact value is

Received by the editors on March 5, 2004.
2000 *Mathematics Subject Classifications* : Primary 94A60.
Key words and phrases: nonlinearity, global avalanche characteristics.

not easily obtainable. It is well known that the nonlinearity of vector Boolean functions $F$ on n-dimensional vector space $GF(2)^n$ to $GF(2)^m$ is bounded above by $2^{n-1} - 2^{\frac{n}{2}-1}$. Zhang and Zheng [7] obtained upper bounds and lower bounds on the nonlinearity of Boolean functions using the notion of auto-correlation. In this paper we extend the notion of auto-correlations of Boolean functions to vector Boolean functions and obtain upper bounds and a lower bound on the nonlinearity of vector Boolean functions in the context of their auto-correlations. This result generalizes Zhang and Zheng's results [7]. Also we can describe avalanche characteristics of vector Boolean functions by examining the extended notion of auto-correlations.

## 2. Basic definitions and properties

In this section, we introduce notations, definitions and well known properties for cryptographic Boolean functions which will be used in this paper. Let $GF(2)^n$ be an n-dimensional vector space over the Galois field $GF(2)$. Put $GF(2)^{n*} = GF(2)^n - \{0\}$. A function $f$ from $GF(2)^n$ to $GF(2)$ is called a Boolean function on $GF(2)^n$. Let $B_n$ denote the set of all Boolean functions on $GF(2)^n$. Let $f \in B_n$ be a Boolean function. The truth table of $f$ is a $(0,1)$-sequence defined by $(f(a_0), f(a_1), \cdots, f(a_{2^n-1}))$ where $a_0 = (0,0,\cdots,0), a_1 = (0,0,\cdots,1), \cdots, a_{2^n-1} = (1,1,\cdots,1)$ . The sequence of $f$ is a $(1,-1)$-sequence defined by $((-1)^{f(a_0)}, (-1)^{f(a_1)}, \cdots, (-1)^{f(a_{2^n-1})})$ where each exponent is regarded as being real-valued. Let $a = (a_1, \cdots, a_n)$ and $b = (b_1, \cdots, b_n)$ be two vectors (or sequences), the scalar product of $a$ and $b$ denoted by $< a, b >$ is defined as the sum of the component-wise multiplications. In particular, when $a$ and $b$ are $(0,1)$-sequences, $< a, b >= a_1b_1 \oplus \cdots \oplus a_nb_n$ , where the addition and multiplications are over $GF(2)$, and when $a$ and $b$ are $(1,-1)$-sequences, $< a, b >= a_1b_1 + \cdots + a_nb_n$ where the addition and multiplications are over the reals. A function $f \in B_n$ that takes the form of $f(x) = a_1x_1 + \cdots + a_nx_n$

where $a_j \in GF(2), j = 1, 2, \cdots n$ is called an affine function. The Hamming weight $W(x)$ of $x \in GF(2)^n$ is the number of ones in $x$. The Hamming distance between two functions $f$ and $g$ is defined by $\#\ \{x|f(x) \neq g(x)\}$. We denote it by $wt(f+g)$. The minimal distance between $f$ and any affine function from $GF(2)^n$ into $GF(2)$ is the nonlinearity of $f$, that is,

$$N(f) = \min_{\phi \in \Gamma} wt(f + \phi)$$

where $\Gamma$ is the set of all affine functions over $GF(2)^n$. The nonlinearity of Boolean functions measures the ability of a cryptographic system using the functions to resist against being expressed as a set of linear equations. It is known that the nonlinearity of arbitrary Boolean function is bounded above by $N(f) \leq 2^{n-1} - 2^{\frac{n}{2}-1}$. A function with this maximal nonlinearity is called a bent function and exists if and only if $n$ is even. The Walsh–Hadamard transformation of a Boolean function $f$ is defined as $W_f(a) = \sum_{x \in GF(2)^n} (-1)^{f(x)+<a,x>}$, for $a \in GF(2)^n$ . Since $W_f(a) = wt(f(x)+ < a, x >) - wt(f(x)+ < a, x > +1)$, we have

$$N(f) = 2^{n-1} - \frac{1}{2} \max_{a \in GF(2)^n} |W_f(a)|.$$

Since a bent function has the maximal nonlinearity $2^{n-1} - 2^{\frac{n}{2}-1}$, equivalently, a bent function is defined as a Boolean function with $W_f(a) = \pm 2^{\frac{n}{2}}$ for all $a \in GF(2)^n$.

Cryptographic applications, such as the design of strong substitution boxes, require that when input coordinates of a Boolean function are selected independently, at random, the output of the function must behave as a uniformly distributed random variable. This yields to the definition of balancedness. A Boolean function $f \in B_n$ is balanced if

$$\#\{x \in GF(2)^n | f(x) = 0\} = \#\{x \in GF(2)^n | f(x) = 1\}.$$

Both linear functions and affine functions are balanced functions.

We say a Boolean function $f$ satisfies the propagation criterion (PC) with respect to a vector $a \in GF(2)^n$ if and only if $\#\{x \in GF(2) | f(x+a) = f(x)\} = 2^{n-1}$ or equivalently $f(x+a)+f(x)$ is balanced. A Boolean function is said to satisfy k-th order propagation characteristic if it is balanced for all $a \in GF(2)^n$ with $1 \leq wt(a) \leq k$. For a Boolean function $f$, if $f(x+a)+f(x)$ is a constant for $a \in GF(2)^n$, a is called a linear structure of $f$. The following results can be found in [6].

LEMMA 2.1. *Let $B_n$ be a Boolean function on $GF(2)^n$. Then the following statements are equivalent.*

*(1) $f$ is bent.*

*(2) $< \xi, l >= \pm 2^{\frac{1}{2}}$ for any affine sequence $l$ of length $2^n$ , where $\xi$ is the sequence of $f$.*

*(3) $f(x)+f(x+a)$ is balanced for any nonzero $a \in GF(2)^n$.*

LEMMA 2.2. *Let $f$ be a bent function. Then the following holds.*

*(1) $f$ satisfies PC of degree $k$ for all $1 \leq k \leq n$.*

*(2) $f$ has maximum nonlinearity.*

*(3) $f$ has no linear structure.*

*(4) $f$ is not balanced.*

*(5) $f$ satisfies SAC.*

Given a Boolean function $f$ on $GF(2)^n$ and a vector $a \in GF(2)^n$, we denote by $\xi(a)$ the sequence of $f(x+a)$. The auto-correlation of $f$ with a shift $a$ is defined by $\Delta_f(a) = < \xi(0), \xi(a) >$. To further simplify our discussions, $\Delta_f(a)$ will be written as $\Delta(a)$ if the function under consideration is clear. Obviously, $\Delta(a) = 0$ if and only if $f(x)+f(x+a)$ is balanced, and $|\Delta(a)| = 2^n$ if and only if $f(x)+f(x+a)$ is a constant, i.e., $a$ is a linear structure of $f$. The following lemmas on upper bounds and a low bound on nonlinearity of Boolean functions [7] will be used in Section 3.

LEMMA 2.3. *For any Boolean function $f$ on $GF(2)^n$, the nonlinearity of $f$ satisfies*

$$N_f \leq 2^{n-1} - \frac{1}{2}\sqrt[4]{2^{2n} + \sum_{j=1}^{2^{n-1}} \Delta^2(a_j)}.$$

It is easy to verify that the bound does not exceed the well known bound $2^{n-1} - 2^{\frac{1}{2}n-1}$ . In addition, as the equality holds if $f$ is bent, the bound is tight.

LEMMA 2.4. *For any Boolean function $f$ on $GF(2)^n$, the nonlinearity of $f$ satisfies*

$$N_f \leq 2^{n-1} - \frac{1}{2}\sqrt{2^n + \Delta_{max}}$$

*where* $\Delta_{max} = max\{|\Delta(a)||a \in GF(2)^{n*}\}$.

LEMMA 2.5. *For any Boolean function $f$ on $GF(2)^n$, the nonlinearity of $f$ satisfies*

$$N_f \geq 2^{n-2} - \frac{1}{4}\Delta_{min}$$

*where* $\Delta_{min} = min\{|\Delta(a)||a \in GF(2)^{n*}\}$.

## 3. Auto-correlation and bounds on the nonlinearities of vector Boolean functions

Now we introduce vector Boolean functions and extend the notion of auto-correlation of Boolean functions to vector Boolean functions and derive upper bounds and a lower bound of nonlinearity of vector Boolean functions in terms of those notions.

A function $F : GF(2)^n \to GF(2)^m$ is called a vector Boolean function on $GF(2)^n$. When $n \leq m$, $F$ is said to be balanced if and only if $\{x \in GF(2)^n | F(x) = b\} = 2^{n-m}$ for any $b \in GF(2)^m$. Note that if a basis of $GF(2)^m$ over $GF(2)$ is specified, there are unique boolean functions $f_i$'s such that $F = (f_1, f_2, \cdots, f_m)$. We denote by $b \cdot F$ the Boolean

function $b_1f_1 + b_2f_2 + \cdots + b_mf_m$ for $b = (b_1, b_2, \cdots, b_m) \in GF(2)^m$. The nonlinearity of $F$, $N(F)$, is defined as

$$N(F) = \min_{b \in GF(2)^m} N(b \cdot F) = \min_{b \neq 0, \phi \in \Gamma} wt(b \cdot F + \phi)$$

where $\Gamma$ is the set of all affine functions over $GF(2)$. It is known that the nonlinearity of arbitrary vector boolean function is bounded above by $N(F) \leq 2^{n-1} - 2^{\frac{n}{2}-1}$. A function with this maximal nonlinearity is called a bent function and exists if and only if $n \geq 2m$ and $n$ is even. Equivalently, a bent function can be defined as a Boolean function with $W_{b \cdot F}(a) = \pm 2^{\frac{n}{2}}$ for all $a \in GF(2)^n$ and $b \in GF(2)^m$. A bent function has cryptographically ideal nonlinearity, but it is not balanced and is only defined over vector spaces with even dimension. Also $F$ is bent if and only if $b \cdot F$ is bent for any $b \in GF(2)^{n*}$. The following Lemma follows immediately from the definition of bent function [1, 2] and Lemma 2.14 in [4].

LEMMA 3.1. *Let $F$ be a bent function. Then for any vector $b$ in $GF(2)^{m*}$ we have the followings:*

*(1) $b \cdot F$ satisfies PC of degree $k$ for all $1 \leq k \leq n$.*

*(2) $b \cdot F$ satisfies SAC.*

*(3) $b \cdot F$ has maximum nonlinearity.*

*(4) $b \cdot F$ has no linear structure.*

*(5) $b \cdot F$ is not balanced.*

We define the auto-correlation $\Delta_F(a)$ of $F$ with a shift $a$ as follows.

DEFINITION 3.1. Let $F$ be a vector Boolean function on $GF(2)^n$ to $GF(2)^m$. For any vector $a \in GF(2)^n$ the auto-correlation of $F$ with a shift a is defined as

$$\Delta_F(a) = (\frac{1}{2^m - 1} \sum_{b \neq 0} \Delta^2_{b \cdot F}(a))^{\frac{1}{2}}.$$

By definition if $\Delta_F(a) = 0$ , $b \cdot F$ satisfies Propagation Characteristic for all $b \in GF(2)^{n*}$ and $a$. The converse is also true. Now we want to derive upper bounds and a lower bound on linearity of vector Boolean functions.

THEOREM 3.2. *For any vector Boolean function $F$ on $GF(2)^n$, the nonlinearity of $F$ satisfies*

$$N_F \leq 2^{n-1} - \frac{1}{2}\sqrt[4]{2^{2n} + \sum_{j=1}^{2^{n-1}} \Delta_F^2(a_j)}.$$

*Proof.* Firstly, for $b*$ in $GF(2)^{m*}$ we may assume the following equality holds.

$$\textstyle\sum_{j=1}^{2^{n-1}} \Delta_F^2(a_j) = max\{\sum_{j=1}^{2^{n-1}} \Delta_{b\cdot F}^2(a_j) | b \in GF(2)^{m*}\}.$$

The right hand-side of the inequality of Theorem 3.2 is

$$\begin{aligned}
&2^{n-1} - \tfrac{1}{2}\sqrt[4]{2^{2n} + \textstyle\sum_{j=1}^{2^{n-1}} \Delta_F^2(a_j)} \\
&= 2^{n-1} - \tfrac{1}{2}\sqrt[4]{2^{2n} + \textstyle\sum_{j=1}^{2^{n-1}} \frac{1}{2^m-1} \sum_{j=1}^{2^{n-1}} \Delta_{b\cdot F}^2(a_j)} \\
&\geq 2^{n-1} - \tfrac{1}{2}\sqrt[4]{2^{2n} + \tfrac{1}{2^m-1} \textstyle\sum_{b\neq 0} \sum_{j=1}^{2^{n-1}} \Delta_{b*\cdot F}^2(a_j)} \\
&= 2^{n-1} - \tfrac{1}{2}\sqrt[4]{2^{2n} + \textstyle\sum_{j=1}^{2^{n-1}} \Delta_{b*\cdot F}^2(a_j)} \qquad \text{(By definition of } b* \text{ )} \\
&\geq N_{b*\cdot F} \qquad \text{(By Lemma 2.3.)} \\
&\geq N_F. \qquad \text{(By definition of } N(F)) \quad \square
\end{aligned}$$

THEOREM 3.3. *For any vector Boolean function $F$ on $GF(2)^n$, the nonlinearity of $F$ satisfies*

$$N_F \leq 2^{n-1} - \frac{1}{2}\sqrt{2^n + \Delta_F^{max}}$$

*where* $\Delta_F^{max} = max\{\Delta_F(a) | a \in GF(2)^n*\}$.

*Proof.* For $a* \in GF(2)^{n*}$ and $b* \in GF(2)^{m*}$ we may assume the following equality holds.

$\Delta^2_{b_*\cdot F}(a*) = max\{\Delta^2_{b\cdot F}(a*)|b* \in GF(2)^{m*}\}$.

The right hand-side of the inequality of Theorem 3.3 is

$$\begin{aligned}
&2^{n-1} - \tfrac{1}{2}\sqrt{2^n + \Delta_F^{max}} \\
&= 2^{n-1} - \tfrac{1}{2}\sqrt{2^n + \Delta_F(a*)} && \text{(By definition of } a*) \\
&= 2^{n-1} - \tfrac{1}{2}\sqrt{2^n + (\tfrac{1}{2^m-1}\textstyle\sum_{b\neq 0}\Delta^2_{b\cdot F}(a*))^{\frac{1}{2}}} \\
&= 2^{n-1} - \tfrac{1}{2}\sqrt{2^n + (\tfrac{1}{2^m-1}\textstyle\sum_{b\neq 0}\Delta^2_{b_*\cdot F}(a*))^{\frac{1}{2}}} && \text{(By definition of } b*) \\
&= 2^{n-1} - \tfrac{1}{2}\sqrt{2^n + \Delta_{b_*\cdot F}(a*)} \\
&= 2^{n-1} - \tfrac{1}{2}\sqrt{2^n + \Delta^{max}_{b_*\cdot F}} && \text{(By definition of } \Delta^{max}_{b_*\cdot F}) \\
&\geq N_{b_*\cdot F} && \text{(By Lemma 3.3.)} \\
&\geq N_F. && \text{(By definition of } N(F)) \qquad \square
\end{aligned}$$

THEOREM 3.4. *For any vector Boolean function $F$ on $GF(2)^n$, the nonlinearity of $F$ satisfies*

$$N_F \geq 2^{n-2} - \frac{1}{4}\Delta_F^{min}$$

*where* $\Delta_F^{min} = max\{\Delta^{min}_{b\cdot F}|b \in GF(2)^m*\}$.

*Proof.* It follows immediately from the definition of $\Delta_F^{min}$. $\square$

All theorems above are independent of the dimension of codomain of $F$.

The overall avalanche characteristic of a function $f$ can be measured by examining $|\Delta(a)|$ for all nonzero vectors $a$. We can say that a function has a good GAC(Global Avalanche Characteristic) if for most nonzero $a$, $|\Delta(a)|$ is zero or very close to zero. This observation leads us to the following definition [5, 6, 7]. Let $F : GF(2)^n \to GF(2)^m$ be a vector Boolean function. We define the sum-of-square indicator $\sigma_F$ for

the global avalanche characteristics of $F$ by

$$\sigma_F = \sum \Delta_F^2(a) = \sum_a \frac{1}{2^m - 1} \sum_{b \neq 0} \Delta_{b \cdot F}^2(a)$$

and the absolute indicator $\Delta_F$ for the global avalanche characteristic of $F$ by

$$\Delta_F = max\{\Delta_F(a) | a \in GF(2)^{n*}\}.$$

The smaller $\sigma_F$ and $\Delta_F$ the better the GAC of a function $F$. Also in general the larger the nonlinearity the smaller(i.e. the better) the GAC of a function $F$.

PROPOSITION 3.5. *Let $F : GF(2)^n \to GF(2)^m$ be a vector Boolean function on $GF(2)^n$. Then we have*

*(1) $2^{2n} \leq \sigma_f \leq 2^{3n}$*

*(2) $\sigma_F = 2^{2n}$ if and only if $F$ is a bent function.*

*(3) $\sigma_F = 2^{3n}$ if and only if $F$ is an affine function.*

*Proof.* It follows immediately from definition of $\sigma_F$ and Theorem 3.2 in [4]. □

By definition $\Delta_F$ is the maximum among $\Delta_F(a), a \neq 0$ and for any $b$ in $GF(2)^m * \Delta_{b \cdot F}(a) = \pm 2^n$ if and only if $a$ is a linear structure of $b \cdot F$. Thus the following result is straightforward.

PROPOSITION 3.6. *Let $F : GF(2)^n \to GF(2)^m$ be a vector Boolean function on $GF(2)^n$. Then we have $0 \leq \Delta_F \leq 2^n$. Moreover, $\Delta_F = 0$ if and only if $F$ is a bent function.*

## REFERENCES

1. J. Cheon, *Nonlinear Vector Resilient Function*, Lecture Notes in Computer Science, Springer-Verlag, 1999.
2. J. Cheon and J. Silverman, *An Algebraic Approach to Boolean Functions*, preprint.
3. J. Daemen and V. Rijmen, *The Design of Rijndael*, Springer, 1991.

4. W. Kim, Y. Kim and M. S. Rhee, *Global avalanche criterion for the S-boxes of DES*, Journal of Korea Soc. Math. Educ. Ser. B: Pure Appl. Math. **8** (2001), pp. 163–174.
5. J. Seberry, X. Zhang and Y. Zheng, *The relationship between propagation characteristics and nonlinearity of cryptographic functions*, Journal of Universal Computer Science **1** (1995), pp. 136–150.
6. X. Zhang and Y. Zheng, *The criterion for global avalanche characteristics of cryptographic function*, Journal of Universal Computer Science **1** (1995), pp. 316–333.
7. X. Zhang and Y. Zheng, *Auto-correlations and new bounds on the nonlinearity of Boolean functions*, EUROCRYPT '96 Proceedings, LNCS 1070, Springer-Verlag, Berlin, Heidelberg, 1996.
8. X. Zhang and Y. Zheng, *Characterizing the structures of cryptographic functions satisfying the propagation criterion for almost all vectors*, Design Codes and Cryptography **7** (1996), pp. 111–134.

*
Department of Mathematics
Hoseo University
Asan 336-795, Korea
*E-mail*: kimws@office.hoseo.ac.kr

**
Department of Mathematics
Hoseo University
Asan 336-795, Korea
*E-mail*: junspk@office.hoseo.ac.kr