Hyper-Bent Functions

A. M. Youssef [1] and G. Gong [2]

Center for Applied Cryptographic Research
[1]Department of Combinatorics & Optimization
[2]Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, Ontario N2L 3G1, CANADA
{a2youssef,ggong}@cacr.math.uwaterloo.ca

Abstract. Bent functions have maximal minimum distance to the set of affine functions. In other words, they achieve the maximal minimum distance to all the coordinate functions of affine monomials. In this paper we introduce a new class of bent functions which we call hyper-bent functions. Functions within this class achieve the maximal minimum distance to all the coordinate functions of *all* bijective monomials. We provide an explicit construction for such functions. We also extend our results to vectorial hyper-bent functions.

Key words. Boolean functions, bent functions, hyper-bent functions, nonlinearity.

1 Introduction

Nonlinearity [19] is a crucial requirement for the Substitution boxes (S-boxes) in secure block ciphers. The success of linear cryptanalysis [16] depends on how well the S-boxes functions can be approximated by an affine function. Highly nonlinear functions provide good resistant towards linear cryptanalysis. On the other hand, even such functions can be attacked by higher order differential cryptanalysis [13] [25] if they have a low algebraic degree. From the view point of polynomials, Jakobsen and Knudsen [13] introduced the interpolation attack on block ciphers. This attack is useful for attacking ciphers using simple algebraic functions as S-boxes. In [13] Jakobsen extended this cryptanalysis method to attack block ciphers with probabilistic nonlinear relation of low degree. The complexity of both attacks depends on the degree of the polynomial approximation and/or on the number of terms in the polynomial approximation expression. Along the same line of research, Gong and Golomb [11] introduced a new criterion for the S-box design. By showing that many block ciphers can be viewed as a non linear feedback shift register with input, Gong and Golomb proposed that S-boxes should not be approximated by a bijective monomial. The reason is that, for $gcd(c, 2^n-1) = 1$, the trace functions $Tr(\zeta_j x^c)$ and $Tr(\lambda x)$, $x \in GF(2^n)$, are both m-sequences with the same linear span.
For boolean functions with even number of input variables, bent functions achieve

the maximal minimum distance to the set of affine functions. In other words, they achieve the maximal minimum distance to all the coordinate functions of affine monomials (I.e., functions in the form $Tr(\lambda x) + e)$). However, this doesn't guarantee that such bent functions cannot be approximated by the coordinate functions of bijective monomials (I.e., functions in the form $Tr(\lambda x^c) + e, gcd(c, 2^n - 1) = 1)$. For example, 120 bent functions out of the 896 bent functions with 4 input variables, have a minimum Hamming distance distance of 2 from the coordinate functions of the monomial x^7 and their complements.
A natural question is whether there exists a class of functions that have the same distance to all the coordinate functions of all bijective monomials. In this paper we give an affirmative answer to this question and provide an explicit construction method for such functions. Functions obtained by our construction also achieve the maximum algebraic normal form degree. We also extend our results to vectorial boolean functions.

We conclude this section by the notation and concepts which will be used throughout the paper. For the theory of shift register sequences and finite fields, the reader is referred to [9], [15].

- $\mathbb{K} = GF(2^{2n})$.
- $\mathbb{E} = GF(2^n)$.
- $\mathbb{F} = GF(2)$.
- α a primitive element of $\mathbb{K}$.
- $Tr_M^N(x)$, $M|N$, represents the trace function from $\mathbb{F}_{2^N}$ to $\mathbb{F}_{2^M}$, i.e., $Tr_M^N(x) = x + x^q + \cdots + x^{q^{l-1}}$ where $q = 2^M$ and $l = N/M$. If $M = 1$ and the context is clear, we write it as $Tr(x)$.
- $\underline{\mathbf{a}} = \{a_i\}$, a binary sequence with period $s|2^{2n} - 1$. Sometimes, we also use a vector of dimension s to represent a sequence with period s. I.e., we also write $\underline{\mathbf{a}} = (a_0, a_1, \cdots, a_{s-1})$.
- $Per(\underline{\mathbf{b}})$, the least period of a sequence $\underline{\mathbf{b}}$.
- $wt(s)$: the number of 1's in one period of the sequence s or the number of 1's in the set of images of the function $s(x) : GF(2^m) \to GF(2)$. This is the so-called *the Hamming weight* of s whether s is a periodic binary sequence or a function from $GF(2^m)$ to $GF(2)$.

2 Preliminaries

There exists a 1-1 correspondence among the set of binary sequences with period $m|2^N - 1$, the set of polynomial functions from $GF(2^N)$ to $GF(2)$ and the set of Boolean functions in N variables through the trace representation of sequences. However, these connections are scattered in the literature. In this section, we will put this 1-1 correspondence together.

B. 1-1 Correspondence Among Periodic Sequences, Polynomial Functions and Boolean Functions

Let

- $\mathcal{S}$ be the set of all binary sequences with period $r|2^N - 1$,
- $\mathcal{F}$, the set of all (polynomial) functions from $GF(2^N)$ to $GF(2)$, and
- $\mathcal{B}$ the set of all Boolean functions in N variables.

There is a 1-1 correspondence among these three sets:

$$\mathcal{S} \longleftrightarrow \mathcal{F} \longleftrightarrow \mathcal{B}$$

which we will explain as follows.

B1. 1-1 Correspondence Between $\mathcal{S}$ and $\mathcal{F}$

Without loss of generality, assume that $f(0) = 0$. Any non-zero function $f(x) \in \mathcal{F}$ can be represented as

$$f(x) = \sum_{i=1}^{s} Tr_1^{m_{t_i}}(\beta_i x^{t_i}), \beta_i \in GF(2^{m_{t_i}})^*, \tag{1}$$

where t_i is a coset leader of a cyclotomic coset modulo $2^N - 1$, and $m_{t_i}|N$ is the size of the cyclotomic coset containing t_i. For any sequence $\underline{\mathbf{a}} = \{a_i\} \in \mathcal{S}$, there exists $f(x) \in \mathcal{F}$ such that

$$a_i = f(\alpha^i), i = 0, 1, \cdots,$$

where α is a primitive element of $\mathbb{K}$. $f(x)$ is called *the trace representation* of $\underline{\mathbf{a}}$. ($\underline{\mathbf{a}}$ is also referred to as an s-term sequence.) If $f(x)$ is any function from $\mathbb{K}$ to $\mathbb{F}$, by evaluating $f(\alpha^i)$, we get a sequence over $\mathbb{F}$ with period dividing $2^N - 1$. Thus

$$\delta : \underline{\mathbf{a}} \leftrightarrow f(x) \tag{2}$$

is a one-to-one correspondence between $\mathcal{F}$ and $\mathcal{S}$ through the trace representation in (1). We say that $f(x)$ is the *trace representation* of $\underline{\mathbf{a}}$ and $\underline{\mathbf{a}}$ is the *evaluation* of $f(x)$ at α. In this paper, we also use the notation $\underline{\mathbf{a}} \leftrightarrow f(x)$ to represent the fact that $f(x)$ is the trace representation of $\underline{\mathbf{a}}$. The set consisting of the exponents that appear in the trace terms of $f(x)$ is said to be the *null spectrum set* of $f(x)$ or $\underline{\mathbf{a}}$.

If $s = 1$, i.e.,

$$a_i = Tr_1^N(\beta\alpha^i), i = 0, 1, \cdots, \beta \in \mathbb{K}^*,$$

then $\underline{\mathbf{a}}$ is an m-sequence over $\mathbb{F}$ of period $2^N - 1$ of degree N. (For a detailed treatment of the trace representation of sequences, see [17].)

B2. 1-1 Correspondence between $\mathcal{F}$ and $\mathcal{B}$

Let $\{\alpha_1, \cdots \alpha_N\}$ be a basis of $\mathbb{K}/\mathbb{F}$ and let α be a primitive element of K. For $x \in \mathbb{K}$ we can represent x as

$$x = x_0\alpha_0 + x_1\alpha_1 + \cdots + x_{N-1}\alpha_{N-1}, x_i \in \mathbb{F}.$$

Thus we have

$$f(x) = f(\sum_{i=0}^{N-1} x_i\alpha_i) = g(x_0, \cdots, x_{N-1}),$$

i.e.,

$$\sigma : f(x) \to g(x_0, \cdots, x_{N-1}) \tag{3}$$

is a bijective map from $\mathcal{F}$ to $\mathcal{B}$. On the other hand, from the Lagrange interpolation [15], for a given Boolean function $g(x_0, \cdots, x_{N-1}) \in \mathcal{B}$, we can determine its polynomial representation $f(x)$ as follows

$$f(x) = \sum_{i=0}^{2^N-1} d_i x^i, \tag{4}$$

where

$$d_i = \sum_{x \in \mathbb{K}^*} (g(x_0, \cdots, x_{N-1}) - g(0, \cdots, 0))x^{-i},$$

where $x = \sum_{i=0}^{N-1} x_i\alpha_i$. Thus (4) gives a bijective map from $\mathcal{B}$ to $\mathcal{F}$ which is the inverse of (3). The correspondence among $\mathcal{S}$, $\mathcal{F}$ and $\mathcal{B}$ is shown in Figure 1.

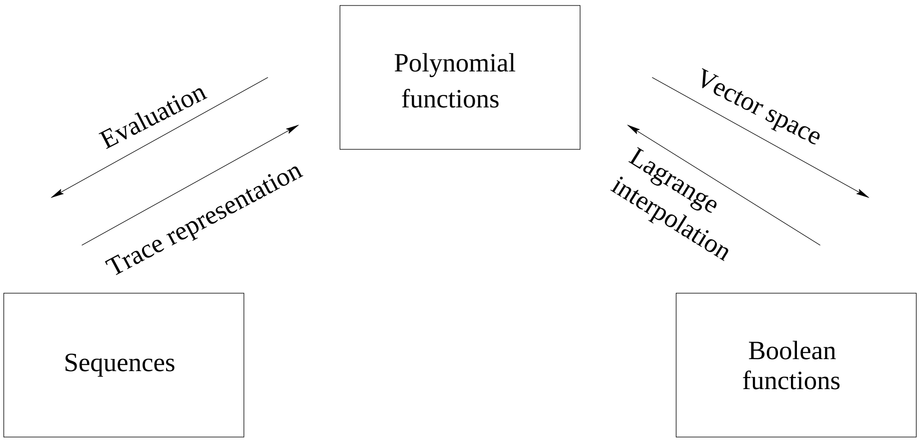

Fig. 1. Correspondence among $\mathcal{S}$, $\mathcal{F}$ and $\mathcal{B}$

From the above diagram, we have a 1-1 correspondence between $\mathcal{S}$, the set of all binary sequences with period dividing $2^N - 1$, and $\mathcal{B}$, the set of Boolean functions in N variables.

3 Extended Transform Domain Analysis for Boolean Functions

The Hadamard transform of $f : \mathbb{E} \to \mathbb{F}$ is defined by [1]

$$\hat{f}(\lambda) = \sum_{x \in \mathbb{E}} (-1)^{f(x)+Tr(\lambda x)}, \lambda \in \mathbb{E}. \tag{5}$$

The Hadamard transform spectrum of f exhibits the nonlinearity of f. More precisely, the nonlinearity of f is given by

$$NL(f) = 2^{n-1} - \frac{1}{2} \max_{\lambda \in \mathbb{E}} |\hat{f}(\lambda)|.$$

I.e., the absolute value of $\hat{f}(\lambda)$ reflects the difference between agreements and disagreements of $f(x)$ and the linear function $Tr(\lambda x)$. Only Bent functions [23] have a constant spectrum of their Hadamard transform. Gong and Golomb [11] showed that many block ciphers can be viewed as a non linear feedback shift register with input. In the analysis of shift register sequences [9], all m-sequences are equivalent under the decimation operation on elements in a sequence. The same idea can be used to approximate boolean functions, i.e., we can use monomial functions instead of linear functions to approximate Boolean functions. In other words, for $gcd(c, n-1) = 1$, the trace functions $Tr(\zeta_j x^c)$ and $Tr(\lambda x)$ have the same linear span. From the view point of m-sequences, both of the sequences $\{Tr(\zeta \alpha^{ic})\}_{i \geq 0}$ and $\{Tr(\lambda \alpha^i)\}_{i \geq 0}$ are m-sequences of period $2^n - 1$. The former can be obtained from the later by decimation c. Gong and Golomb [11] introduced the concept of extended Hadamard transform (*EHT*) for a function from $\mathbb{E}$ to $\mathbb{F}$. The extended Hadamard transform is defined as follows.

Definition 1. *Let f(x) be a function from $\mathbb{E}$ to $\mathbb{F}$. Let*

$$\hat{f}(\lambda, c) = \sum_{x \in \mathbb{K}} (-1)^{f(x)+Tr(\lambda x^c)} \tag{6}$$

where $\lambda \in \mathbb{E}$ and c is a coset leader modulo $2^n - 1$ co-prime to $2^n - 1$. Then we call $\hat{f}(\lambda, c)$ an extended Hadamard transform of the function f.

Notice that the Hadamard transform of f, defined by (5), is $\hat{f}(\lambda, 1)$. The numerical results in [11] show that, for all the coordinate functions $f_i, i = 1, \cdots, 32$ of the DES s-boxes, the distribution of $\hat{f}_i(\lambda, c)$ in λ is invariant for all c.

Thus a new generalized nonlinearity measure can be defined as

$$NLG(f) = 2^{n-1} - \frac{1}{2} \max_{\substack{\lambda \in \mathbb{E}, \\ c : gcd(c, 2^n - 1) = 1}} |\hat{f}(\lambda, c)|.$$

This leads to a new criterion for the design of Boolean functions used in conventional cryptosystems. The *EHT* of Boolean functions should not have any large component.

Throughout the rest of the paper, we consider functions with even number of input variables.

4 Construction for Hyper-Bent Functions

In this section we introduce a new class of functions which have a constant *EHT* spectrum. A binary function $f : \mathbb{K} \to \mathbb{F}$ is said to be hyper-bent if and only if the *EHT* of f,

$$\widehat{f}(\lambda, c) = \sum_{x \in \mathbb{K}} (-1)^{Tr(\lambda x^c) + f(x)} = \pm 2^n$$

for all $\lambda \in \mathbb{K}$ and for all c, $gcd(c, 2^{2n} - 1) = 1$. Clearly a hyper-bent function must be bent.

Let $\underline{\mathbf{b}} = \{b_j\}_{j \geq 0}$ be a binary sequence with period $2^n + 1$. In the following, first we will give the criterion such that $g(x) \leftrightarrow \underline{\mathbf{b}}, g(0) = 0$, is hyper-bent and count the number of such functions. Then we will show that all such hyper-bent functions obtained from our construction achieve maximal algebraic degree.

Theorem 1. *With the above notation, then $g(x)$ is hyper-bent if and only if $wt(\underline{\mathbf{b}}) = 2^{n-1}$, i.e., $\underline{\mathbf{b}}$ is balanced (Note that a sequence is said to be balanced if the disparity between the number of 1's and the number of 0's in one period is not to exceed 1).*

In order to prove Theorem 1 we need the following Lemmas. Let $r = 2^n - 1, d = 2^n + 1$. Write $u_i = Tr(\zeta \alpha^{ci}), i = 0, 1, \cdots$. Thus $\underline{\mathbf{u}} = \{u_i\}$ is an m-sequence of period $2^{2n} - 1$. Let $\underline{\mathbf{v}} = \underline{\mathbf{u}} + \underline{\mathbf{b}}$. Then $\underline{\mathbf{v}}$ can be written into a (d, r)-interleaved sequence [10], i.e., $\underline{\mathbf{v}}$ can be arranged into the following array

$$\begin{bmatrix} v_0 & v_1 & \cdots & v_{d-1} \\ v_d & v_{d+1} & \cdots & v_{2d-1} \\ \vdots & \vdots & \vdots & \vdots \\ v_{d(r-1)} & v_{d(r-1)+1} & \cdots & v_{rd-1} \end{bmatrix} = (\underline{\mathbf{v}}_0, \underline{\mathbf{v}}_1, \cdots \underline{\mathbf{v}}_{d-1}),$$

where $\underline{\mathbf{v}}_j$'s are columns of the matrix.

Lemma 1. *$\underline{\mathbf{v}}_j = \underline{\mathbf{u}}_j + \underline{\mathbf{b}}_j$ where $\underline{\mathbf{u}}_j$'s are columns of the matrix*

$$\begin{bmatrix} u_0 & u_1 & \cdots & u_{d-1} \\ u_d & u_{d+1} & \cdots & u_{2d-1} \\ \vdots & \vdots & \vdots & \vdots \\ u_{d(r-1)} & u_{d(r-1)+1} & \cdots & u_{rd-1} \end{bmatrix},$$

and $\underline{\mathbf{b}}_j = (b_j, b_j, \cdots, b_j)$ is a constant sequence.

Proof. The result follows by noting that $Per(\underline{\mathbf{u}}) = dr$ and $Per(\underline{\mathbf{b}}) = d$.

□

Lemma 2. *With the notation in Lemma 1, we have*

$$wt(\underline{\mathbf{v}}) = \sum_{j=0}^{d-1} wt(\underline{\mathbf{u}}_j + \underline{\mathbf{b}}_j).$$

Proof. Note that

$$wt(\underline{\mathbf{v}}) = \sum_{j=0}^{d-1} wt(\underline{\mathbf{v}}_j).$$

Applying Lemma 1, the result follows immediately.

□

Lemma 3. *Let* $wt(\underline{\mathbf{b}}) = t$ *and*

$$\widehat{g}(\lambda, c) = \sum_{x \in \mathbb{K}} (-1)^{Tr(\lambda x^c)+g(x)}.$$

Then $\widehat{g}(\lambda, c) \in \{2t, 2(t-2^n)\}, \forall \lambda \in \mathbb{K}^*$ *and* $\widehat{g}(0, c) = (2^n - 1)(d - 2t) + 1$.

Proof.

$$\begin{aligned}\widehat{g}(0, c) &= \sum_{x \in K} (-1)^{g(x)} = 1 + \sum_{x \in K^*} (-1)^{g(x)} = 1 + r \sum_{k=0}^{d-1} (-1)^{b_k} \\ &= 1 + r(d - 2wt(\underline{\mathbf{b}})).\end{aligned}$$

For $\lambda \neq 0$,

$$\begin{aligned}\widehat{g}(\lambda, c) &= \textstyle\sum_{x \in \mathbb{K}} (-1)^{Tr(\lambda x^c)+g(x)} = 1 + \sum_{i=0}^{2^{2n}-1} (-1)^{u_i+b_i} \\ &= 1 - wt(\underline{\mathbf{v}}) + (q - 1 - wt(\underline{\mathbf{v}})) = 2^{2n} - 2wt(\underline{\mathbf{v}}).\end{aligned} \tag{7}$$

So, we only need to determine the value of the Hamming weight of $\underline{\mathbf{v}}$. From Lemma 2 we have

$$wt(\underline{\mathbf{v}}) = \sum_{j=0}^{d-1} wt(\underline{\mathbf{u}}_j + \underline{\mathbf{b}}_j).$$

Note that $\{u_j\}$ is a binary m-sequence of period $2^{2n}-1$. So one of the component sequences $\underline{\mathbf{u}}_j$ is a zero sequence and the rest of 2^n sequences are shifts of the binary m-sequence $\{Tr(\alpha^{icd})\}$ with period r. Without loss of generality, we can suppose that that $\underline{\mathbf{u}}_0 = (0, 0, \cdots 0)$ is the zero sequence. Therefore $wt(\underline{\mathbf{u}}_0) = 0$ and $wt(\underline{\mathbf{u}}_j) = 2^{n-1}, 1 \leq j < 2^n$. We have the following two cases
Case 1. $b_0 = 0$.

$$\begin{aligned}wt(\underline{\mathbf{v}}) &= \sum_{j>0, b_j=0} wt(u_j) + \sum_{j>0, b_j=1} wt(u_j + 1) \\ &= (d - t - 1)2^{n-1} + t(2^{n-1} - 1) \\ &= d2^{n-1} - t - 2^{n-1}.\end{aligned}$$

Thus in this case, we have

$$\widehat{g}(\lambda, c) = 1 + (2^{2n} - 1) - 2(d2^{n-1} - t - 2^{n-1}) = 2^{2n} - d2^n + 2t + 2^n = 2t.$$

Case 2. $b_0 = 1$

$$
\begin{aligned}
wt(\underline{\mathbf{v}}) &= 2^n - 1 + \sum_{j>0, b_j=0} wt(u_j) + \sum_{j>0, b_j=1} wt(u_j + 1) \\
&= 2^n - 1 + (d-t)2^{n-1} + (2^{n-1} - 1)(t-1)) \\
&= 2^{n-1} + d2^{n-1} - t.
\end{aligned}
$$

Substituting into (7) we get

$$\widehat{g}(\lambda, c) = 2^{2n} - 2(2^{n-1} + d2^{n-1} - t) = 2^{2n} - 2^n - d2^n + 2t = 2(t - 2^n).$$

Thus $\widehat{g}(\lambda, c) \in \{2t, 2(t - 2^n)\}, \forall \lambda \in \mathbb{K}^*$.

$\square$

Proof of Theorem 1. If $g(x)$ is hyper-bent, then $|\widehat{g}(\lambda, c)| = 2^n$. From Lemma 3, we have $\widehat{g}(\lambda, c) = 2t$ or $2(t - 2^n)$. Thus we have $2t = 2^n \Leftrightarrow t = 2^{n-1}$ or $2(2^n - t) = 2^n \Leftrightarrow t = 2^{n-1}$. Thus if $g(x)$ is hyper-bent then $wt(\underline{\mathbf{b}}) = 2^{n-1}$. Conversely, if $wt(\underline{\mathbf{b}}) = 2^{n-1} = t$, then according to Lemma 3 we have

$$\widehat{g}(\lambda, c) = 2t = 2^n,$$

or

$$\widehat{g}(\lambda, c) = 2(t - 2^n) = 2(2^{n-1} - 2^n) = -2^n,$$

which implies that

$$|\widehat{g}(\lambda, c)| = 2^n, \forall \lambda \in \mathbb{K}^*.$$

and

$$\widehat{g}(0, c) = (2^n - 1)(2^n + (d - 2t)) + 1 = (2^n - 1)(2^n + 1 - 2^n) + 1 = 2^n.$$

Thus $g(x)$ is bent. $\square$

In the following theorem, we will count how many hyper-bent functions can be obtained using this construction and show these functions achieve maximal algebraic degree.

Note. Let $f(x) = \sum_{i=0}^{2^m-1} c_i x^i$ be a function from $GF(2^m)$ to $GF(2)$. The algebraic degree of $f(x)$ is defined by

$$AD_f = max\{AD_{x^i} | i : c_i \neq 0\} \text{ where } AD_{x^i} = wt(i).$$

So the algebraic degree of $f(x)$ is equal to the algebraic degree of a boolean form of $f(x)$.

Theorem 2. *Let*

$$S(d) = \{\underline{\mathbf{b}} = \{b_i\}_{i \geq 0} | b_i \in \mathbb{F} \text{ and } Per(\underline{\mathbf{b}}) = d\},$$

and

$$\Gamma(d) = \{h(x) : \mathbb{K} \to \mathbb{F} | h(x) \leftrightarrow \underline{\mathbf{b}} \in S(d)\}.$$

Then there are $\binom{2^n+1}{2^{n-1}}$ *hyper-bent functions in* $\Gamma(d)$ *and each of such functions has algebraic degree* n*, which is the maximal algebraic degree that bent functions in* $2n$ *variables can achieve.*

In order to prove Theorem 2 we need the following lemma.

Lemma 4. *Let* $0 < a < 2^n + 1$, $a \equiv 1$ *mod* 2. *Then*

$$wt(a(2^n - 1)) = n.$$

Proof. If $a = 2^n$, then $wt(2^n(2^n - 1)) = wt(2^n - 1) = n$. If $0 < a < 2^n$, then we can write

$$a(2^n - 1) = a2^n - a = 2^n(a - 1) + (2^n - a).$$

Since $(2^n - a) < 2^n$ and $\geq 2^n 2^n(a - 1) \leq 2^{2n}$ then

$$wt(a(2^n - 1)) = wt(2^n(a - 1)) + wt(2^n - a). \tag{8}$$

Since $a \equiv 1$ mod 2, we can write

$$a = 1 + 2^{i_1} + 2^{i_2} + \cdots + 2^{i_k},$$

where $0 < i_1 < i_2 < \cdots < i_k < n$. Thus

$$a - 1 = 2^{i_1} + \cdots + 2^{i_k} \Rightarrow wt(2^n(a - 1)) = k. \tag{9}$$

We also have

$$2^n - a = 2^n - 1 - (a - 1) = 1 + 2 + \cdots + 2^n - 1 - (2^{i_1} + \cdots + 2^{i_k}).$$

Therefore

$$wt(2^n - a) = n - k. \tag{10}$$

Substituting (9) and (10) into (8), we obtain that $wt(a(2^n - 1)) = k + n - k = n$.

$\square$

Proof of Theorem 2. According to Theorem 1, $h(x)$ is hyper-bent if and only if $wt(\underline{\mathbf{b}}) = 2^{n-1}$. I.e., there are 2^{n-1} 1's in $\{b_0, b_1, \cdots, b_{d-1}\})$ where $Per(\underline{\mathbf{b}})|d$. In the following, we first show $Per(\underline{\mathbf{b}}) = d$ provided $wt(\underline{\mathbf{b}}) = 2^{n-1}$ and $Per(\underline{\mathbf{b}})|d$. I.e., if $wt(\underline{\mathbf{b}}) = 2^{n-1}$ and $Per(\underline{\mathbf{b}})|d$, then $\underline{\mathbf{b}} \in S(d)$. Let $d = kPer(\underline{\mathbf{b}})$. Then $wt(\underline{\mathbf{b}})k = 2^{n-1} \Longrightarrow k|2^c$. Since $gcd(d, 2) = 1$ and $k|d$, this forces $k = 1$. Therefore $Per(\underline{\mathbf{b}}) = d$. Hence $\underline{\mathbf{b}} \in S(d)$. Note that the number of sequences with weight 2^{n-1} in $S(d)$ is $\binom{2^n + 1}{2^{n-1}}$ which is just the number of bent functions in $\Gamma(d)$.

It can also be shown that for each $h(x) \in \Gamma(d)$, there exists $f : \mathbb{K} \to \mathbb{F}$ such that $h(x) = f(x^r)$ and the evaluation of $f(x)$ has period dr. Therefore, the exponent of any trace term in $h(x)$ has r as a divisor, i.e., it can be expressed as rs where $1 \leq s \leq 2^n$ and s is a coset leader. Applying Lemma 4, we have $wt(rs) = n$. Thus all exponents in $h(x)$ have weight n. Therefore $h(x)$ has algebraic degree n which is maximal since bent functions in $2n$ variables has algebraic degree $\leq n$ [21].

Remark 1. It is easy to show that the complement of the functions obtained from our constructions are also hyper-bent. Thus the total number of hyper-bent functions (with $2n$ input bits) obtained from our construction is $2\binom{2^n + 1}{2^{n-1}}$.

Remark 2. After this paper was accepted for publications, Claude Carlet pointed out that this class of hyper-bent functions corresponds to the class of bent functions (family PS/ap) introduced by Dillon in his dissertation. In [6] Dillon constructed this class using the partial spread difference sets.

Remark 3. The special case where $g(x)$ in Theorem 1 is given by $Tr(\alpha x^{2^n-1})$ and $\alpha \in \mathbb{K}$ such that $\underline{\mathbf{b}} \leftrightarrow Tr(\alpha x^{2^n-1})$ is balanced, is equivalent to the difference set construction in [7]. In [7] Dillon showed that the difference sets constructed using $Tr(\alpha x^{2^n-1})$ is inequivalent to the Maiorana-McFarland construction [18]. Using a similar approach, it is easy to show that the construction in Theorem 1 is inequivalent to the Maiorana-McFarland construction.

5 Construction of Balanced Functions with Large *NLG*

By randomly complementing 2^{n-1} zeros of a hyper-bent functions we obtain a balanced functions with $NLG \geq 2^{2n-1} - 2^n$. Using this construction procedure, for $2n = 8$, we were able to obtain balanced functions with $NLG = 116$. (Note that the best known nonlinearity for balanced functions with 8 input variables is 116 [8], [22]).

Example 1. The $(17, 15)$ interleaved sequence corresponding the a balanced function $f : GF(2^8) \rightarrow GF(2)$ with $NLG = 116$ is shown below. The ones in the complemented positions is surrounded by brackets.

$$\begin{bmatrix}
1&1&0&0&1&1&0&1&1&0&0&0&1&0&0&1&0\\
1&1&(1)&0&1&1&0&1&1&0&0&0&1&0&0&1&0\\
1&1&(1)&0&1&1&0&1&1&0&0&0&1&(1)&0&1&0\\
1&1&0&0&1&1&0&1&1&0&0&0&1&0&0&1&0\\
1&1&0&0&1&1&0&1&1&0&0&0&1&0&0&1&0\\
1&1&0&0&1&1&0&1&1&0&0&0&1&0&0&1&0\\
1&1&0&0&1&1&0&1&1&0&0&0&1&0&0&1&0\\
1&1&(1)&0&1&1&0&1&1&0&0&0&1&0&0&1&0\\
1&1&(1)&0&1&1&0&1&1&0&0&0&1&0&0&1&0\\
1&1&0&0&1&1&0&1&1&0&0&0&1&0&0&1&0\\
1&1&0&0&1&1&0&1&1&0&0&0&1&0&0&1&0\\
1&1&(1)&0&1&1&0&1&1&0&0&0&1&0&0&1&0\\
1&1&(1)&0&1&1&0&1&1&0&0&0&1&0&0&1&0\\
1&1&(1)&0&1&1&0&1&1&0&0&0&1&0&0&1&0\\
1&1&0&0&1&1&0&1&1&0&0&0&1&0&0&1&0
\end{bmatrix}$$

6 Construction for Vectorial Hyper-Bent Functions

Let $\{\eta_0, \eta_1, \cdots, \eta_{m-1}\}$ be a basis of $GF(2^m)$ over $GF(2)$. Then we can write $h : GF(2^{2n}) \rightarrow GF(2^m)$ as

$$h(x) = \sum_{j=0}^{m-1} h_j(x)\eta_j.$$

We call $h(x)$ *VHB* if and only if every nonzero linear combination of its output coordinates is a hyper-bent function, i.e., $h(x)$ is a *VHB* function if and only for any non-zero m-tuple $(c_0, c_1, \cdots, c_{m-1}) \in \mathbb{F}^m$, the function $\sum_{i=0}^{m-1} c_i h_i(x)$ from $\mathbb{K}$ to $\mathbb{F}$ is hyper-bent. Clearly, *VHB* functions are a sub-class of perfect nonlinear functions [21], [24]. It is known that such functions exist only for $m \le n$ [21]. In this section we will apply the results in Section 4 to present a new construction for vectorial hyper-bent (*VHB*) functions with the maximum possible number of outputs and maximal algebraic degree.

Keep r and d as defined before. Let $\gamma = \alpha^d$ and $\mathbb{E} = GF(2^n)$. Then γ is a primitive element of $\mathbb{E}$. Let $\{\beta_0, \cdots, \beta_{n-1}\}$ be the dual basis of $\{1, \gamma, \cdots, \gamma^{n-1}\}$ of $\mathbb{E}$ over $\mathbb{F}$ and $\pi : \mathbb{E} \to \mathbb{E}$ be a permutation. We define

$$b_{i,j}^{\pi} = \begin{cases} Tr(\beta_j \pi(\gamma^i)), & 0 \le i \le 2^n - 2, \\ Tr(\beta_j \pi(0), & i = 2^n - 1, \\ 0, & i = 2^n, \end{cases} \tag{11}$$

Let $\underline{\mathbf{b}}_j^{\pi} = \{b_{i,j}^{\pi}\}_{i \ge 0}$ where $b_{i,j}^{\pi} = b_{s,j}^{\pi}$ for $i = kd + s$ with $0 \le s < d$. From the construction above, it is easy to see that $Per(\underline{\mathbf{b}}_j^{\pi}) = d$.

Theorem 3. *With the notation above, let Π be the set consisting of all permutations of $\mathbb{E}$ and*

$$P = \{h(x) : \mathbb{K} \to \mathbb{E} | h(x) = \sum_{j=0}^{n-1} h_j(x)\gamma^j, h_j(x) \leftrightarrow \underline{\mathbf{b}}_j^{\pi}, \pi \in \Pi\},$$

Then any function $h \in P$ is a VHB *function with maximum algebraic degree. Moreover, we have*

$$|P| = 2^n!.$$

Proof. We can write $\pi(x) = \sum_{j=0}^{n-1} Tr(\beta_j \pi(x))\gamma^j$. Since π is a permutation of $\mathbb{E}$, then $wt(Tr(\beta_j \pi(x)) = 2^{n-1}$ for $0 \le j < n$. From (11), $\underline{\mathbf{b}}_j^{\pi}$ is obtained from the evaluation of $Tr(\beta_j \pi(x))$ by lengthening it by one zero bit. Thus $wt(\underline{\mathbf{b}}_j^{\pi}) = 2^{n-1}$ for each $j : 0 \le j < n$. For any nonzero m-tuple $(c_0, c_1, \cdots, c_{m-1}) \in \mathbb{F}^m$, the evaluation of $D(x) = \sum_{i=0}^{m-1} c_i h_i(x)$ can be obtained from the evaluation of $C(x) = \sum_{i=0}^{m-1} c_i Tr(\beta_i \pi(x))$ by lengthening it by one zero bit. Note that

$$C(x) = Tr\left(\sum_{i=0}^{m-1} c_i \beta_i \pi(x)\right) = Tr(\theta\pi(x))$$

where $\theta = \sum_{i=0}^{m-1} c_i \beta_i \ne 0, \theta \in \mathbb{E}$. Thus $wt(C(x)) = wt(Tr(\theta\pi(x))) = 2^{n-1}$. Therefore the evaluation of $D(x)$ has weight 2^{n-1}. According to Theorem 4, $D(x)$ is hyper-bent. Thus $h(x)$ is a *VHB* function. The proof regarding the algebraic degree is identical to that in the proof of Theorem 2. Since the number of permutations of E is $2^n!$, then $|P| = 2^n!$.

□

Remark 4. Note that we can insert one zero bit in equation (11) at any place other than $i = 2^n$ and we still have $wt(\underline{\mathbf{b}}_j) = 2^{n-1}$. Hence we have $(2^n+1)!/2$ such constructions, which corresponds to the number of ordered partitions of $2^n + 1$ elements of which 2 elements are the same and the rest of the $2^n - 1$ elements are all different. It is easy to show that the complement of the functions obtained from our constructions are also *VHB* functions. Thus the total number of *VHB* functions (with $2n$ input bits) obtained from our construction is

$$(2^n + 1)!.$$

Remark 5. Note that the notion of hyper-bent functions investigated in this paper is different from the one used in [5]. In fact, the class of functions considered in [5] are those Boolean functions on $(GF(2))^m$ (m even) such that, for a given even integer k ($2 \leq k \leq m-2$), any of the Boolean functions on $(GF(2))^{m-k}$ obtained by fixing k coordinates of the variable is bent.

7 Conclusions

Boolean functions used in block cipher design should have a large Hamming distance to functions with simple algebraic description. In this paper, we presented a method to construct bent functions which achieve the maximal minimum distance to the set of all bijective monomials. Functions obtained from our construction achieve the maximum algebraic degree. These functions can be modified to achieve the balance property while maintaining large distance to bijective monomials. We also presented a method to construct vectorial bent functions for which every non zero linear combination of its coordinate functions satisfy the above property. These functions also achieve both the largest degree and the largest number of output bits. It should also be noted that while Rijndael (the NIST's Selection for the AES [14]) S-boxes are constructed by the monomial $x^{-1} = x^{254}$ over $GF(2^8)$, an affine transformation over $GF(2)$ is applied to the output of these S-boxes and hence the equivalent S-boxes will not have a simple algebraic description when looked at as a polynomial over $GF(2^8)$.

Acknowledgment

After this paper was accepted for publications, we realized that the term "hyper-bent" has been used by Claude Carlet to refer to a different class of Boolean functions. We would like to thank Claude Carlet for encouraging us to keep the "hyper-bent" term and for attracting our attention to references [6] and [7] .

References

1. R.E. Blahut, *Theory and Practice of Error Control Codes*, Addison-Wesley Publishing Company, 1983.

2. C. Carlet, A construction of bent functions, *Proc. of Third Conference of Finite Fields and Applications*, Glasgow, London Mathematical Society, Lecture Series 233, Cambridge University Press, 1996, pp. 47-58.
3. C. Carlet, Two new classes of bent functions, *Advances in Cryptology-EuroCrypt'85*, Lecture Notes in Computer Science, No. 765, Springer-Verlag, 1994, pp. 77-101.
4. C. Carlet , P. Charpin and V. Zinoviev, *Codes, bent functions and permutations suitable for DES-like cryptosystems*, Designs, Codes and Cryptography. vol.15, no.2; Nov. 1998; pp.125-156.
5. C. Carlet, Hyper-bent functions, *PRAGOCRYPT'96*, Czech Technical University Publishing House, Prague, pp. 145-155, 1996.
6. J. F. Dillon, *Elementary Hadamard Difference sets*, Ph.D. Dissertation, University of Maryland, 1974.
7. J. F. Dillon, *Elementary Hadamard Difference sets*, in Proc. Sixth S-E Conf. Comb. Graph Theory and Comp., 237-249, F. Hoffman et al. (Eds), Winnipeg Utilitas Math (1975) .
8. H. Dobbertin, Construction of bent functions and balanced Boolean functions with high nonlinearity, *Proceedings of Fast Software Encryption, Second International Workshop*, Springer-Verlag, 1995, pp. 61-74.
9. S.W. Golomb,*Shift Register Sequences*, Aegean Park Press. Laguna Hills, California. 1982.
10. G. Gong, Theory and applications of q-ary interleaved sequences, *IEEE Trans. on Inform. Theory*, vol. 41, No. 2, 1995, pp. 400-411.
11. G. Gong and S. W. Golomb, *Transform Domain Analysis of DES*, IEEE transactions on Information Theory. Vol. 45. no. 6. pp. 2065-2073. September, 1999.
12. T. Jakobsen and L. Knudsen, *The Interpolation Attack on Block Ciphers, LNCS 1267*, Fast Software Encryption. pp. 28-40. 1997.
13. T. Jakobsen, *Cryptanalysis of Block Ciphers with Probabilistic Non-linear Relations of Low Degree*, Proceedings of Crypto'99. LNCS 1462. pp. 213-222. 1999.
14. J. Daemen and V. Rijmen, *AES Proposal: Rijndael*, `http://csrc.nist.gov/encryption/aes/rijndael/`
15. R. Lidl and H. Niederreiter, *Finite Fields*, Encyclopedia of Mathematics and its Applications, Volume 20, Addison-Wesley, 1983.
16. M. Matsui, *Linear Cryptanalysis method for DES cipher* Advances in Cryptology, Proceedings of Eurocrypt'93, LNCS 765, pp. 386-397, Springer-Verlag, 1994.
17. R. J. McEliece, *Finite Fields For Computer Scientists and Engineers* , Kluwer Academic Publishers, Dordrecht, 1987.
18. R.L. McFarland, *A family of Noncyclic Difference Sets* , Journal of Comb. Th. (Series A) 15, pp. 1-10, 1973.
19. W. Meier and O. Staffelbach, *Nonlinearity criteria for cryptographic functions*, Proceedings of EUROCRYPT '89, Springer-Verlag, Berlin, Germany, 1990 pp. 549-62.
20. K. Nyberg, S-boxes and round functions with controllable linearity and differential uniformity, *Proceedings of Fast Software Encryption, Second International Workshop*, Springer-Verlag, Berlin, Germany, 1995, pp.111-130.
21. K. Nyberg, Perfect nonlinear S-boxes, *Proceedings of EUROCRYPT '91*, Springer-Verlag, Berlin, Germany, 1991, pp.378-86.
22. P. Sarkar and S. Maitra, Nonlinearity Bounds and Constructions of Resilient Boolean Functions, *Proceedings of CRYPTO '2000*, Springer-Verlag, Berlin, Germany, LNCS 1880, pp. 515-532.

23. O.S. Rothaus , On bent functions, *J. Combinatorial Theory*, vol. 20(A), 1976, pp.300-305.
24. T. Satoh , T. Iwata and K. Kurosawa, On cryptographically secure vectorial Boolean functions, *Proceedings of ASIACRYPT'99*, Springer-Verlag, Berlin, Germany, 1999, pp. 20-28.
25. T. Iwata and K. Kurosawa, Probabilistic higher order differential cryptanalysis and higher order bent functions. *Proceedings of ASIACRYPT'99*, Springer-Verlag, Berlin, Germany, 1999, pp. 62-74.