# Cryptographic Boolean Functions via Group Hadamard Matrices

Jennifer Seberry

Xian-Mo Zhang

Yuliang Zheng

Department of Computer Science
The University of Wollongong
Wollongong, NSW 2522, AUSTRALIA

**Abstract**

For any integers $n$, $m$, $2n > m > n$ we construct a set of boolean functions on $V_m$, say $\{f_1(z), \ldots, f_n(z)\}$, which has the following important cryptographic properties:

(i) any nonzero linear combination of the functions is balanced;

(ii) the nonlinearity of any nonzero linear combination of the functions is at least $2^{m-1} - 2^{n-1}$;

(iii) any nonzero linear combination of the functions satisfies the strict avalanche criterion;

(iv) the algebraic degree of any nonzero linear combination of the functions is $m-n+1$;

(v) $F(z) = (f_1(z), \ldots, f_n(z))$ runs through each vector in $V_n$ precisely $2^{m-n}$ times while $z$ runs through $V_m$.

## 1 Basic Definitions

Let $V_n$ be the vector space of $n$ tuples of elements from $GF(2)$. Let $\alpha, \beta \in V_n$. Write $\alpha = (a_1, \ldots, a_n)$, $\beta = (b_1, \ldots, b_n)$, where $a_i, b_i \in GF(2)$. Write $\langle \alpha, \beta \rangle = \sum_{j=1}^{n} a_j b_j$. Also write $\alpha = (a_1, \ldots, a_n) < \beta = (b_1, \ldots, b_n)$ if there exists $k$, $1 \leqq k \leqq n$, such that $a_1 = b_1$, $\ldots$, $a_{k-1} = b_{k-1}$ and $a_k = 0$, $b_k = 1$. Hence we can order all vectors in $V_n$ by the relation $<$

$$\alpha_0 < \alpha_1 < \cdots < \alpha_{2^n-1},$$

where

$$\alpha_0 = (0, \ldots, 0, 0), \ \ldots, \ \alpha_{2^{n-1}-1} = (0, 1, \ldots, 1),$$

$$\alpha_{2^{n-1}} = (1, 0, \ldots, 0), \ \ldots, \ \alpha_{2^n-1} = (1, 1, \ldots, 1).$$

**Definition 1** Let $f(x)$ be a function from $V_n$ to $GF(2)$, where $x = (x_1, \ldots, x_n)$, (or simply, a function on $V_n$). The (1 -1)-sequence $\eta = ((-1)^{f(\alpha_0)}\ (-1)^{f(\alpha_1)}\ \ldots\ (-1)^{f(\alpha_{2^n-1})})$ is called the *sequence of* $f(x)$. Similarly, the (0, 1)-sequence $(f(\alpha_0)\ f(\alpha_1)\ \ldots\ f(\alpha_{2^n-1}))$ is called the *truth table* of $f(x)$. In particular, if the truth table of $f(x)$ has $2^{n-1}$ zeros (ones) $f(x)$ is said to be *0-1 balanced (*or simply, *balanced).*

**Definition 2** We call $h(x) = a_1x_1 + \cdots + a_nx_n + c$, $a_j, c \in GF(2)$, *an affine function.* In particular, we will call $h(x)$ *a linear function* if $c = 0$. The sequence of an affine function (a linear function) will be called an *affine sequence* (a linear sequence).

**Definition 3** Let $f$ and $g$ be functions on $V_n$ whose sequences are $\xi$ and $\eta$ respectively. The *Hamming distance* between $f$ and $g$, denoted by $d(f, g)$, is the number of components where $\xi$ and $\eta$ differ. Let $\varphi_1, \ldots, \varphi_{2^n}, \varphi_{2^n+1}, \ldots, \varphi_{2^{n+1}}$ be all affine functions on $V_n$. $N_f = \min_{i=1,\ldots,2^{n+1}} d(f, \varphi_i)$ is called the *nonlinearity* of $f(x)$.

The nonlinearity is a crucial criterion for a good cryptographic design. It prevents a cryptosystem from being attacked by solving a set of linear equations.

**Definition 4** Let $f(x)$ be a function on $V_n$. If $f(x) + f(x + \alpha)$ is 0-1 balanced for every $\alpha \in V_n$ with $W(\alpha) = 1$, where $W(\alpha)$ denotes the number of nonzero components (*the Hamming weight)* of $\alpha$, we say that $f(x)$ satisfies the *strict avalanche criterion the* $(SAC)$.

The strict avalanche criterion was originally defined in [16], [17], and was generalized in two different directions [2], [5], [8], [9], [10], [14]. The 0-1 balance, the nonlinearity and the avalanche criterion are important criteria for cryptographic functions [1], [5], [7], [10].

**Definition 5** A (1, -1)-matrix of order $n$ will be called a *Hadamard matrix* if $HH^T = nI_n$.

If $n$ is the order of an Hadamard matrix then $n$ is 1, 2 or divisible by 4 [15]. A special kind of Hadamard matrix defined below will be relevant:

**Definition 6** A *Sylvester-Hadamard matrix (*or *Walsh-Hadamard matrix)* of order $2^n$, denoted by $H_n$, is generated by the recursive relation

$$H_n = \begin{bmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -H_{n-1} \end{bmatrix}, \quad n = 1, 2, \ldots, \quad H_0 = 1.$$

**Notation 1** *For a vector* $\delta = (i_1, \ldots, i_p) \in V_p$, *we define a function on* $V_p$*:*

$$D_\delta(y_1, \ldots, y_p) = D_{i_1,\ldots,i_p}(y_1, \ldots, y_p) = (y_1 + \bar{i_1}) \cdots (y_p + \bar{i_p})$$

*where* $\bar{i} = 1 + i$.

**Notation 2** *Define a matrix of order $s+t$, denoted by $Q(s,t)$, whose entries come from $GF(2)$, such that*

$$Q(s,t) = \begin{bmatrix} I_s & 0_{s\times t} \\ D & I_t \end{bmatrix},$$

*where $I_i$ is the identity matrix of order $i$, $0_{s\times t}$ is the zero-matrix of order $s \times t$,*

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

Obviously $Q(s,t)$ is a nonsingular matrix.

## 2 The Properties of Balance, Nonlinearity and SAC

In this section we review a number of results on balance, nonlinearity and the SAC. These results will be employed in the later part of the paper.

**Lemma 1**

$$D_{i_1,\ldots,i_p}(y_1,\ldots,y_p) = \begin{cases} 1 & \text{if } (y_1,\ldots,y_p) = (i_1,\ldots,i_p), \\ 0 & \text{if } (y_1,\ldots,y_p) \neq (i_1,\ldots,i_p). \end{cases}$$

*Proof.* The verification is straightforward. □

**Lemma 2** *Let $\xi_{i_1,\ldots,i_p}$ be the sequence of a function $f_{i_1,\ldots,i_p}(x_1,\ldots,x_q)$ on $V_q$. Set $\xi = (\xi_{0,\ldots,0,0}, \xi_{0,\ldots,0,1}, \ldots, \xi_{1,\ldots,1,1})$. Then $\xi$ is the sequence of the function*

$$f(y_1,\ldots,y_p,x_1,\ldots,x_q) = \sum_{(i_1,\ldots,i_p)\in V_p} D_{i_1,\ldots,i_p}(y_1,\ldots,y_p) f_{i_1,\ldots,i_p}(x_1,\ldots,x_q), \tag{1}$$

*that is a function on $V_{q+p}$.*

(See Lemma 1 of [11].)

**Lemma 3** *$f(y_1,\ldots,y_p,x_1,\ldots,x_q)$, defined as in ( 1) is the zero function on $V_{q+p}$ if and only if each $f_{i_1,\ldots,i_p}(x_1,\ldots,x_q)$ is the zero function on $V_q$.*

*Proof.* $f(y_1,\ldots,y_p,x_1,\ldots,x_q)$ is the zero function on $V_{q+p}$ if and only if $f(i_1,\ldots,i_p,x_1,\ldots,x_q)$ is the zero function on $V_q$ for any fixed $(i_1,\ldots,i_p) \in V_p$. From Lemma 1, $f(i_1,\ldots,i_p,x_1,\ldots,x_q) = f_{i_1,\ldots,i_p}(x_1,\ldots,x_q)$. □

¿From the proof of Lemma 3, any function can be uniquely presented by (1).

**Lemma 4** *$D_\delta(y+\beta) = D_{\delta+\beta}(y)$ where $y,\ \delta \in V_p$.*

*Proof.* Since $D_\delta(y+\beta) = 1$ if and only if $y+\beta = \delta$. $D_{\delta+\beta}(y) = 1$ if and only if $y = \delta+\beta$. This proves the lemma. □

**Lemma 5** *Write* $H_n = \begin{bmatrix} \ell_0 \\ \ell_1 \\ \vdots \\ \ell_{2^n-1} \end{bmatrix}$ *where $\ell_i$ is a row of $H_n$. Then each $\ell_i$ is the sequence of the linear function $h_i(x) = \langle \alpha_i, x \rangle$ where $\alpha_i$, $0 \leqq i \leqq 2^n - 1$, is a vector in $V_n$, $x \in V_n$.*

(See Lemma 2 of [11].)

¿From Lemma 5, the rows of $H_n$ comprise all the sequences of linear functions on $V_n$ and hence the rows of $\pm H_n$ comprise all the sequences of affine functions on $V_n$.

**Lemma 6** *Let $f$ and $g$ be functions on $V_n$ whose sequences are $\eta_f$ and $\eta_g$ respectively. Then $d(f,g) = 2^{n-1} - \frac{1}{2}\langle \eta_f, \eta_g \rangle$.*

(See Lemma 3 of [11].)

**Lemma 7** *For any function $f$ on $V_n$, $N_f \leqq 2^{n-1} - 2^{\frac{1}{2}n-1}$.*

(See Lemma 4 of [11].)

**Lemma 8** *Let $f(x)$ be a function on $V_n$, $A$ be a nonsingular matrix of order $n$, with entries from $GF(2)$. Set $f(xA) = \psi(x)$. Then*

*(i) $f$ is balanced if and only if $\psi$ is balanced,*

*(ii) $N_f = N_\psi$.*

*Proof.* (i) Note that $\psi(x_0) = 0$ if and only if $f(x_0 A) = 0$.

(ii) Let $h(x)$ be an affine function on $V_n$. Set $h_A(x) = h(xA)$. $\psi(x_0) \neq h_A(x_0)$ if and only if $f(x_0 A) \neq h(x_0 A)$. Thus $d(f,h) = d(\psi, h_A)$. Note that while $h$ runs through all affine functions on $V_n$, $h_A$ runs through all affine functions on $V_n$ since $A$ is nonsingular. □

**Theorem 1** *Let $f(x)$ be a function on $V_n$, $A$ be a nonsingular matrix of order $n$, with entries from $GF(2)$. Set $f(xA) = \psi(x)$. Let $\gamma_i$ denote the $i$th row of $A$. If $f(x) + f(x+\gamma_i)$ is balanced for $i = 1, \ldots, n$ then $\psi(x)$ satisfies the SAC.*

*Proof.* Let $\delta_i$ denote the vector $V_n$, whose the $i$th entry is 1 and others 0. Note that $IA = A$. Thus $\delta_i A = \gamma_i$, $i = 1, \ldots, n$. Note that $\psi(x) + \psi(x + \delta_i) = f(xA) + f((x + \delta_i)A) = f(u) + f(u + \gamma_i)$, where $u = xA$. Since $A$ is nonsingular $uA^{-1} = x$ will go through $V_n$ while $u$ runs through $V_n$. Thus $\psi(x) + \psi(x + \delta_i)$ is balanced, $i = 1, \ldots, n$, that is to say, $\psi(x)$ satisfies the $SAC$. □

**Lemma 9** *Let $g(y_1, \ldots, y_s)$ be a function on $V_s$. Set $f(y_1, \ldots, y_s, x_1, \ldots, x_t) = g(y_1, \ldots, y_s)$, a function on $V_{s+t}$.*

*(i) If $g$ is balanced then $f$ is balanced,*

*(ii) $N_f \geqq 2^t N_g$.*

*Proof.* (i) $g(y_1, \ldots, y_s)$ takes the value 0 and the value 1 both $2^{s-1}$ times while $(y_1, \ldots, y_s)$ runs through $V_s$ once. Hence $f(y_1, \ldots, y_s, x_1, \ldots, x_t)$ takes the value 0 and the value 1 both $2^{t+s-1}$ times while $(y_1, \ldots, y_s, x_1, \ldots, x_t)$ runs through $V_{s+t}$ once.

(ii) Let $f_1(x_1, \ldots, x_t, y_1, \ldots, y_s) = f(y_1, \ldots, y_s, x_1, \ldots, x_t) = g(y_1, \ldots, y_s)$.

Let $\xi$ be the sequence of $g$ hence $\eta = (\xi, \ldots, \xi)$ is the sequence of $f_1$, where $\eta$ is the concatenation of $2^t$ $\xi_s$.

Let $L$ be an affine sequence of length $2^{t+s}$. By Lemma 5, $L$ is a row of $\pm H_{t+s} = \pm H_t \times H_s$. Thus $L = \pm \ell' \times \ell''$ where $\ell'$ is a linear sequence of length $2^t$, a row of $H_t$ and $\ell''$ is a linear sequence of length $2^s$, a row of $H_s$. Write $\ell' = (a_1, \ldots, a_{2^t})$ thus $L = (a_1 \ell'', \ldots, a_{2^t} \ell'')$. Note that $\langle \eta, L \rangle = \sum_{j=1}^{2^t} a_j \langle \xi, \ell'' \rangle$. Let $\ell''$ be the sequence of a linear function on $V_s$, say $h$. Since $d(g, h) \geqq N_g$, by Lemma 6, $\langle \xi, \ell'' \rangle \leqq 2^s - 2N_g$. Note that $\sum_{j=1}^{2^t} a_j \leqq 2^t$ thus $\langle \eta, L \rangle \leqq 2^t(2^s - 2N_g)$. Let $L$ be the sequence of an affine function on $V_{t+s}$, say $h^*$. Hence by Lemma 6, $d(f_1, h^*) \geqq 2^t N_g$. Since $h^*$ is arbitrary $N_{f_1} \geqq 2^t N_g$. By (ii) of Lemma 8, $N_f = N_{f_1} \geqq 2^t N_g$. □

**Corollary 1** *Let $g(y_1, \ldots, y_s)$ be a function on $V_s$. Set $f(y_1, \ldots, y_s, x_1, \ldots, x_t) = g(y_1, \ldots, y_s)$, a function on $V_{s+t}$. Let $A = Q(s, t)$ where $Q(s, t)$ is defined as in Notation 2. Set $f(zA) = \psi(z)$ where $z = (y,\ x)$, $y = (y_1, \ldots, y_s)$, $x = (x_1, \ldots, x_t)$. If $g$ satisfies the $SAC$ then $\psi$ satisfies the $SAC$.*

*Proof.* Let $\gamma_i$ denote the $i$th row of $A$. Write $\gamma_i = (\sigma_i,\ \tau_i)$ where $\sigma_i \in V_s$, $\tau_i \in V_t$.

For $i = 1, \ldots, s$, $f(z) + f(z + \gamma_i) = g(y) + g(y + \sigma_i)$.

Since $g$ satisfies the $SAC$ $g(y) + g(y + \sigma_i)$ is balanced on $V_s$, by (i) of Lemma 9, $f(z) + f(z + \gamma_i)$ is balanced on $V_{s+t}$.

For $i = s + 1, \ldots, s + t$, $f(z) + f(z + \gamma_i) = g(y) + g(y + \sigma_i)$. By the same reasoning, $f(z) + f(z + \gamma_i)$ is balanced on $V_{s+t}$.
Note that $A$ is nonsingular. By Theorem 1, $\psi$, as a function on $V_{s+t}$, satisfies the $SAC$. □

## 3 Basic Construction

For $y \in V_s$, $x \in V_t$, write $y = (y_1, \ldots, y_s)$, $x = (x_1, \ldots, x_t)$.

$$f(y_1, \ldots, y_s, x_1, \ldots, x_t) = \sum_{(j_1,\ldots,j_s)\in V_s} D_{j_1,\ldots,j_s}(y) f_{j_1,\ldots,j_s}(x) + r(y) \tag{2}$$

where $D_{j_1,\ldots,j_s}$ is defined as in Notation 1, each $f_{j_1,\ldots,j_s}(x)$ is a function on $V_t$, $r(y)$ is a function on $V_s$.

**Lemma 10** *If each $f_{j_1,\ldots,j_s}(x)$ in (2) is balanced then $f$ is balanced.*

*Proof.* For any fixed $(j_1, \ldots, j_s) \in V_s$ $f(j_1, \ldots, j_s, x_1, \ldots, x_t) = D_{j_1,\ldots,j_s}(j_1, \ldots, j_s) f_{j_1,\ldots,j_s}(x) + r(j_1, \ldots, j_s) = f_{j_1,\ldots,j_s}(x) + r(j_1, \ldots, j_s)$, that is balanced. Thus $f$ is balanced. □

**Theorem 2** *Let $f$ be defined as in (2), where each $f_{j_1,\ldots,j_s}(x)$ is a nonzero linear function on $V_t$ then*

*(i) $f$ is balanced,*

*(ii) $N_f \geqq 2^{s+t-1} - 2^{t-1}$ if all $f_{j_1,\ldots,j_s}(x)$ are distinct linear functions on $V_t$,*

*(iii) $f(z) + f(z+\gamma)$ is balanced whenever $\beta \neq 0$, where $z = (y,\ x)$, $\gamma = (\beta,\ \alpha)$, $y, \beta \in V_s$, $x, \alpha \in V_t$, if $f_{j_1,\ldots,j_s}(x)$ are distinct linear functions on $V_t$.*

*Proof.* (i) Since any nonzero linear function is balanced, by Lemma 10, $f$ is balanced.

(ii) Let $\xi_{j_1,\ldots,j_s}$ be the sequence of $f(j_1, \ldots, j_s, x_1, \ldots, x_t) = f_{j_1,\ldots,j_s}(x) + r(j_1, \ldots, j_s)$. Thus $\xi_{j_1,\ldots,j_s}$ is a nonzero affine sequence. By Lemma 2, $\eta = (\xi_{0,\ldots,0},\ \xi_{0,\ldots,0,1}, \ldots, \xi_{1,\ldots,1,1})$ is the sequence of $f(y_1, \ldots, y_s, x_1, \ldots, x_t)$.

Let $L$ be an affine sequence of length $2^{s+t}$. By Lemma 5, $L$ is a row of $\pm H_{s+t} = \pm H_s \times H_t$. Thus $L = \pm\ell' \times \ell''$ where $\ell'$ is a linear sequence of length $2^s$, a row of $H_s$ and $\ell''$ is a linear sequence of length $2^t$, a row of $H_t$. Write $\ell' = (a_{0,\ldots,0},\ a_{0,\ldots,0,1}, \ldots, a_{1,\ldots,1,1})$, where the subscript $(j_1, \ldots, j_s) \in V_s$. Thus $L = (a_{0,\ldots,0}\ell'',\ a_{0,\ldots,0,1}\ell'', \ldots, a_{1,\ldots,1,1}\ell'')$. $\langle \eta, L\rangle = \sum_{j_1,\ldots,j_s} a_{j_1,\ldots,j_s} \langle \xi_{j_1,\ldots,j_s}, \ell''\rangle$. Note that each $\xi_{j_1,\ldots,j_s}$ is a nonzero affine sequence. Thus

$$\langle \xi_{j_1,\ldots,j_s}, \ell''\rangle = \begin{cases} \pm 2^t & \text{if } \xi_{j_1,\ldots,j_s} = \pm\ell'', \\ 0 & \text{otherwise.} \end{cases}$$

Since all the $\xi_{j_1,\ldots,j_s}$ are distinct there exists at most one $\xi_{j_1,\ldots,j_s}$ such that $\xi_{j_1,\ldots,j_s} = \pm\ell''$. Thus $\langle \eta, L\rangle = \pm 2^t$ or 0. Let $L$ be the sequence of an affine function, say $h^*$. By Lemma 6, $d(f, h^*) \geqq 2^{s+t-1} - 2^{t-1}$. Since $h^*$ is arbitrary $N_f \geqq 2^{s+t-1} - 2^{t-1}$.

(iii) Let $\beta = (b_1, \ldots, b_s)$. By Lemma 4,

$$D_{j_1,\ldots,j_s}(y_1 + b_1, \ldots, y_s + b_s) = D_{j_1+b_1,\ldots,j_s+b_s}(y_1, \ldots, y_s).$$

Hence

$$\begin{aligned} f(z+\gamma) &= \sum_{j_1,\ldots,j_s} D_{j_1,\ldots,j_s}(y+\beta) f_{j_1,\ldots,j_s}(x+\alpha) + r(y+\beta) \\ &= \sum_{j_1,\ldots,j_s} D_{j_1+b_1,\ldots,j_s+b_k}(y) f_{j_1,\ldots,j_s}(x+\alpha) + r(y+\beta) \\ &= \sum_{j_1+b_1,\ldots,j_s+b_s} D_{j_1+b_1,\ldots,j_s+b_k}(y) f_{j_1,\ldots,j_s}(x+\alpha) + r(y+\beta). \end{aligned} \tag{3}$$

Set $(j_1,\ldots,j_s) = (i_1+b_1,\ldots,i_s+b_s)$.

$$f(z+\gamma) = \sum_{i_1,\ldots,i_s} D_{i_1,\ldots,i_s}(y) f_{i_1+b_1,\ldots,i_s+b_s}(x+\alpha) + r(y+\beta).$$

$$f(z)+f(z+\gamma) = \sum_{i_1,\ldots,i_s} D_{i_1,\ldots,i_s}(y)(f_{j_1,\ldots,j_s}(x) + f_{j_1+b_1,\ldots,j_s+b_s}(x+\alpha)) + r(y) + r(y+\beta).$$

Note that $\beta = (b_1,\ldots,b_s) \neq 0$ $f_{j_1,\ldots,j_s}(x)+f_{j_1+b_1,\ldots,j_s+b_s}(x+\alpha) = f_{j_1,\ldots,j_s}(x)+f_{j_1+b_1,\ldots,j_s+b_s}(x)+f_{j_1+b_1,\ldots,j_s+b_s}(\alpha)$ is a non-constant affine function since all $f_{j_1,\ldots,j_s}(x)$ are distinct linear functions on $V_t$. By Lemma 10 $f(z)+f(z+\beta)$ is balanced.

□

# 4 A Group Generalised Hadamard Matrix

Let $G$ be a group, $\underline{p} = (p_1,\ldots,p_n)$, $\underline{q} = (q_1,\ldots,q_n)$ be two vectors of length $n$, whose entries $p_j$, $q_j$ come from $G$. Define the operation $\circ$ such that $\underline{p}\circ\underline{q} = (p_1q_1,\ldots,p_nq_n)$ and the inverse of $\underline{q}$ such that $\underline{q}^{-1} = (q_1^{-1},\ldots,q_k^{-1})$.

$\underline{p}$ and $\underline{q}$ are *s-orthogonal* if $\underline{p}\circ\underline{q}^{-1} = (p_1q_1^{-1},\ldots,p_nq_n^{-1})$ comprise $s$ times of all the elements of $G$.

*A generalised Hadamard matrix* ( [3], [4]) of type $s$ for group $G$ is a square matrix with entries from $G$ whose rows are mutually s-orthogonal.

A *group Hadamard matrix* [6] is a generalised Hadamard matrix whose rows form a group and whose columns form a group under the operation $\circ$. Note that in a group Hadamard matrix of type $s$ for $G$ there exists a row acting the role of identity. By the definition of generalised Hadamard matrix, each of other rows contains each element of $G$ $s$ times.

Let $\varepsilon$ be a primitive element of $GF(2^k)$, $G$ be the additive group of $GF(2^k)$. Set $X = (\varepsilon^{j-i+1(mod\ 2^k-1)})$, where $i,j = 1,2,\ldots,2^k-1$, and $D_1 = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & X & \\ 0 & & \end{bmatrix}$. Hence $D_1$ is a generalised Hadamard matrix of order $2^k$, type 1 (1-orthogonal) for $G$ also a group Hadamard matrix [3], [4], [6].

It is easy to find out that $D_2 = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & Y & \\ 0 & & \end{bmatrix}$, where $Y = (\varepsilon^{j+i-1(mod\ 2^k-1)})$, is also a generalised Hadamard matrix of order $2^k$, type 1 (1-orthogonal) for $G$ also a group Hadamard matrix.

Note that an entry of $Y$, an element of $G$, is a polynomial in $\varepsilon$, whose degree is no more than $k-1$, say $a_0 + a_1\varepsilon + \cdots + a_{k-1}\varepsilon^{k-1}$.

We now change $a_0 + a_1\varepsilon + \cdots + a_{k-1}\varepsilon^{k-1}$ into $a_0x_1 + a_1x_2 + \cdots + a_{k-1}x_k$, a linear function on $V_k$.

Note that all linear functions on $V_k$ form an additive group, denoted by $\Gamma_k$.

Correspondingly $D_2$ becomes a matrix $E$ with entries from $\Gamma_k$. Obviously $E$ is also a group Hadamard matrix of order $2^k$, type 1 (1-orthogonal) but for group $\Gamma_k$.

Write $E = (e_{i,j})$, where $i, j = 0, 1, \ldots, 2^k - 1$.

Let $y = (y_1, \ldots, y_k)$, $x = (x_1, \ldots, x_k)$. Set

$$f_i(y_1, \ldots, y_k, x_1, \ldots, x_k) = D_{0,\ldots,0}(y)e_{i,0}(x) + D_{0,\ldots,0,1}(y)e_{i,1}(x) + \cdots + D_{1,\ldots,1}(y)e_{i,2^k-1}(x) \quad (4)$$

where $i = 0, 1, \ldots, 2^k - 1$.

**Lemma 11** *For any fixed $s$, $1 \leqq s \leqq 2^k - 1$, $e_{1,s}, \ldots, e_{k,s}$ are linearly independent.*

*Proof.* Consider $\sum_{j=1}^{k} c_j f_j$ where $(c_1, \ldots, c_k) \neq (0, \ldots, 0)$. Note that $e_{1,1} = x_1$, $e_{2,1} = x_2$, $\ldots$, $e_{k,1} = x_k$. It is obvious that

$$\sum_{i=1}^{k} c_i e_{i,1} \neq 0. \quad (5)$$

Since $E$ is a group Hadamard matrix of type 1 (1-orthogonal) for $\Gamma_k$ there exists a row in $E$, say the $i_0$th row, such that $\xi_{i_0} = \sum_{i=1}^{k} c_i\xi_i$, where each $\xi_i$ denotes the $i$th row of $E$ and hence $\sum_{i=1}^{k} c_i e_{i,j} = e_{i_0,j}$, for every $j = 1, \ldots, 2^k - 1$. ¿From (5), the $i_0$th row of $E$ is not a zero row (i.e. $i_0 \neq 0$) and thus contains every linear function on $V_k$ since $E$ is a group Hadamard matrix of type 1 (1-orthogonal) for $\Gamma_k$. Thus $\sum_{i=1}^{k} c_i e_{i,s} = e_{i_0,s}$ is a nonzero linear function for every $s = 1, \ldots, 2^k - 1$. This proves that for any $s$, $1 \leqq s \leqq 2^k - 1$, $\sum_{i=1}^{k} c_i e_{i,s} = 0$ if and only if $(c_1, \ldots, c_k) = (0, \ldots, 0)$ thus $e_{1,s}, \ldots, e_{k,s}$ are linearly independent. □

## 5 A Set of Functions with Cryptographic Properties

Let $P$ be a permutation on $1, 2, \ldots, 2^k - 1$. Let $E'$ be the matrix obtained from $E$ by putting $P$ on the nonzero columns of $E$. Set $E' = (e'_{i,j})$, where $i, j = 0, 1, \ldots, 2^k - 1$.

Let $k < n < 2k$. Write $y = (y_1, \ldots, y_{n-k})$, $x = (x_1, \ldots, x_k)$, $z = (y,\ x)$. Note that $e'_{i,j}$ is nonzero linear function on $V_k$ for $i = 1, 2, \ldots, 2^k - 1$. Set

$$g_i(y,\ x) = D_{0,\cdots,0}(y)e'_{i,1}(x) + D_{0,\ldots,0,1}(y)e'_{i,2}(x) + \cdots + D_{1,\ldots,1}(y)e'_{i,2^{n-k}}(x) + r_i(y) \qquad (6)$$

where $i = 1, \ldots, 2^k - 1$, each subscript $(i_1, \ldots, i_{n-k}) \in V_{n-k}$ and each $r_i$ is a function on $V_{n-k}$.

Let $A = Q(n-k, k)$. Set

$$\psi_i(z) = g_i(zA),\ i = 1, \ldots, 2^k - 1. \qquad (7)$$

**Theorem 3** *For any nonzero linear combination of $\psi_1, \ldots, \psi_k$, defined as in (7), say $\psi = \sum_{j=1}^{k} c_j\psi_j$, where $(c_1, \ldots, c_k) \neq (0, \ldots, 0)$.*

*(i) $\psi$ is balanced,*

*(ii) $N_\psi \geqq 2^{n-1} - 2^{k-1}$,*

*(iii) $\psi$ satisfies the SAC,*

*(iv) the algebraic degree of $\psi$ can be $n - k + 1$.*

*Proof.* ¿From (6),

$$g = \sum_{j=1}^{k} c_j g_j = D_{0,\ldots,0}(y)\sum_{j=1}^{k} c_j e'_{j,1}(x) + D_{0,\ldots,0,1}(y)\sum_{j=1}^{k} c_j e'_{j,2}(x) + \cdots + D_{1,\ldots,1}(y)\sum_{j=1}^{k} c_j e'_{j,2^{n-k}}(x).$$

By Lemma 11, each of $\sum_{j=1}^{k} c_j e'_{j,1}(x)$, $\sum_{j=1}^{k} c_j e'_{j,2}(x), \ldots, \sum_{j=1}^{k} c_j e'_{j,2^{n-k}}(x)$ is a nonzero linear function on $V_k$. Since $E'$ is a group Hadamard matrix of type 1 for $\Gamma_k$, $\sum_{j=1}^{k} c_j e'_{j,1}(x)$, $\sum_{j=1}^{k} c_j e'_{j,2}(x)$, $\ldots$, $\sum_{j=1}^{k} c_j e'_{j,2^{n-k}}(x)$ are distinct linear functions. By Theorem 2, $g$ is balanced and $N_g \geqq 2^{n-1} - 2^{k-1}$. By Lemma 8, $\psi$ is balanced and $N_\psi \geqq 2^{n-1} - 2^{k-1}$.

Let $\gamma_i = (\beta_i,\ \alpha_i)$ be the $i$th row of $A = Q(n-k, k)$, where $\beta_i \in V_{n-k}$, $\alpha_i \in V_k$, $i = 1, \ldots, n$. Since all $\beta_i \neq 0$, by (iii) of Theorem 2, $g(z) + g(z + \gamma_i)$ is balanced, $i = 1, \ldots, n$. Note that $\psi(z) = g(zA)$. By Theorem 1, $\psi$ satisfies the $SAC$.

We can choose $E'$ such that $\sum_{j=1}^{2^{n-k}} e'_{1,j}$ is a nonzero linear function on $V_k$. Otherwise if $\sum_{j=1}^{2^{n-k}} e'_{1,j}$ is zero, we exchange the $2^{n-k}$th and the $(2^{n-k}+1)$th columns of $E'$. Correspondingly, $E'$ is changed into $E'' = (e''_{i,j})$. Since $e'_{1,2^{n-k}} \neq e'_{1,2^{n-k}+1}$, $\sum_{j=1}^{2^{n-k}} e''_{1,j}$ is a nonzero linear function on $V_k$. Hence it is reasonable to suppose $\sum_{j=1}^{2^{n-k}} e'_{1,j}$ is a nonzero linear function on $V_k$. Note that each $D_{j_1,\ldots,j_{n-k}}(y_1, \ldots, y_{n-k})$ contains the term $y_1 \cdots y_{n-k}$ and $y_1 \cdots y_{n-k} \sum_{j=1}^{2^{n-k}} e'_{1,j}$ cannot be deleted in

$$g_1(y,\ x) = D_{0,\cdots,0}(y)e'_{1,1}(x) + D_{0,\ldots,0,1}(y)e'_{1,2}(x) + \cdots + D_{1,\ldots,1}(y)e'_{1,2^{n-k}}(x) + r_1(y).$$

This proves that the degree of $g_1$ is $n - k + 1$.

Since $D_2(E)$ is symmetric the columns of $D_2(E)$ also form a group thus the columns of $E'$ form a group. Recall $\sum_{j=1}^{2^{n-k}} e'_{1,j}$ is a nonzero linear function on $V_k$. Thus $\sum_{j=1}^{2^{n-k}} e'_{i,j}$ is also a nonzero linear function on $V_k$, $i = 2, \ldots, 2^k - 1$.

To show this, note that the columns of $E'$ form a group thus the sum of the first, the second, ..., the $2^{n-k}$th columns of $E'$ is equal to a column of $E'$, say the $s_0$th column. Since $\sum_{j=1}^{2^{n-k}} e'_{1,j} = e'_{1,s_0}$ is a nonzero linear function on $V_k$ the $s_0$th column of $E'$ is a nonzero column (i.e. $s_0 \neq 0$). Thus the $s_0$th column contains all the linear functions on $V_k$ since the columns of $E'$ form a group.

This proves that $\sum_{j=1}^{2^{n-k}} e'_{i,j} = e'_{i,s_0}$ is a nonzero function if $i \neq 0$.

By the same reasoning, the degree of $g_i$ is $n - k + 1$, $i = 2, \ldots, 2^k - 1$.

Since the rows of $E'$ form a group there exists $i_0$th such that the $i_0$ row is equal to the linear combination of $g_1, \ldots, g_k$ corresponding to the coefficients $c_1, \ldots, c_k$. Thus $\sum_{i=1}^{k} c_i g_i = g_{i_0}$. Since the first, the second, ..., the $2^{n-k}$th rows of $E'$ are linearly independent (see Lemma 11) $g_{i_0}$ is a nonzero function (i.e. $i_0 \neq 0$). Thus the degree of $\sum_{i=1}^{2^{n-k}} c_i g_i = g_{i_0}$ is $n - k + 1$.

□

**Corollary 2** *$\Psi(z) = (\psi_1(z), \ldots, \psi_k(z))$, a mapping from $V_n$ to $V_k$, where each $\psi_j$ is defined as in Theorem 3, runs through all the $2^k$ vectors in $V_n$ each $2^{n-k}$ times while $z$ runs through $V_n$.*

*Proof.* By Theorem 1 of [13], this corollary is equivalent to (i) of Theorem 3. □

Since any matrix obtained by permuting the columns of a group Hadamard matrix is still a group Hadamard matrix, we can obtain an extremely large number of boolean function sets with the cryptographic properties mentioned in Theorem 3 and Corollary 2. These functions can be used in many cryptographic designs. In particular, results shown in this section have been successfully employed by the authors in systematically constructing cryptographically robust substitution boxes (S-boxes) [12].

# 6 Example

**Example 1** By using Theorem 3, we now construct 4 functions of 6 variables. Let $k = 4$ and $n = 6$ in Theorem 3. Choose $x^4 + x + 1$ as the primitive polynomial. Let $\varepsilon$ be a root of $x^4 + x + 1 = 0$. $\varepsilon^i$, $j = 0, 1, \ldots, 2^4 - 1$ form a sequence:

$$\begin{array}{llllllll} 1, & \varepsilon, & \varepsilon^2, & \varepsilon^3, & 1+\varepsilon, & \varepsilon+\varepsilon^2, & \varepsilon^2+\varepsilon^3, & 1+\varepsilon+\varepsilon^3, \\ 1+\varepsilon^2, & \varepsilon+\varepsilon^3, & 1+\varepsilon+\varepsilon^2, & \varepsilon+\varepsilon^2+\varepsilon^3, & 1+\varepsilon+\varepsilon^2+\varepsilon^3, & 1+\varepsilon^2+\varepsilon^3, & 1+\varepsilon^3, & \end{array}$$

that is the first row of $Y$ , where $D_2 = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & Y & \\ 0 & & \end{bmatrix}$ of order $2^k$ (see Section 4). We change $\varepsilon^i$ into $x_{i+1}$, $i = 0, 1, 2, 3$. The above sequence becomes

$$\begin{array}{cccccccc} x_1, & x_2, & x_3, & x_4, & x_1+x_2, & x_2+x_3, & x_3+x_4, \\ x_1+x_2+x_4, & x_1+x_3, & x_2+x_4, & x_1+x_2+x_3, & x_2+x_3+x_4, & x_1+x_2+x_3+x_4, & x_1+x_3+x_4, \\ x_1+x_4, & & & & & & \end{array}$$

that is the first row of $W$ , where $E = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & W & \\ 0 & & \end{bmatrix}$ (see Section 4).

We choose the submatrix of order $k \times 2^{k-2}$, that is the conjunction of the first four rows and the 4th, the 9th, the 12th, the 15th columns of $W$:

$$\begin{bmatrix} x_4 & x_1+x_3 & x_2+x_3+x_4 & x_1+x_4 \\ x_1+x_2 & x_2+x_4 & x_1+x_2+x_3+x_4 & x_1 \\ x_2+x_3 & x_1+x_2+x_3 & x_1+x_3+x_4 & x_2 \\ x_3+x_4 & x_2+x_3+x_4 & x_1+x_4 & x_3 \end{bmatrix}.$$

Using the above array we define (see (6))

$$g_1(y_1, y_2, x_1, x_2, x_3, x_4) = (1+y_1)(1+y_2)x_4 + (1+y_1)y_2(x_1+x_3) + y_1(1+y_2)(x_2+x_3+x_4) + y_1y_2(x_1+x_4),$$

$$g_2(y_1, y_2, x_1, x_2, x_3, x_4) = (1+y_1)(1+y_2)(x_1+x_2) + (1+y_1)y_2(x_2+x_4) + y_1(1+y_2)(x_1+x_2+x_3+x_4) + y_1y_2x_1,$$

$$g_3(y_1, y_2, x_1, x_2, x_3, x_4) = (1+y_1)(1+y_2)(x_2+x_3) + (1+y_1)y_2(x_1+x_2+x_3) + y_1(1+y_2)(x_1+x_3+x_4) + y_1y_2x_2,$$

$$g_4(y_1, y_2, x_1, x_2, x_3, x_4) = (1+y_1)(1+y_2)(x_3+x_4) + (1+y_1)y_2(x_2+x_3+x_4) + y_1(1+y_2)(x_1+x_3) + y_1y_2x_3,$$

Simplify the four functions

$$g_1(y_1, y_2, x_1, x_2, x_3, x_4) = x_4 + y_2x_4 + y_2x_1 + y_2x_3 + y_1x_2 + y_1x_3 + y_1y_2x_2 + y_1y_2x_4,$$

$$g_2(y_1, y_2, x_1, x_2, x_3, x_4) = x_1 + x_2 + y_2x_1 + y_2x_4 + y_1x_3 + y_1x_4 + y_1y_2x_1 + y_1y_2x_2 + y_1y_2x_3,$$

$$g_3(y_1, y_2, x_1, x_2, x_3, x_4) = x_2 + x_3 + y_1x_2 + y_2x_1 + y_1x_1 + y_1x_4 + y_1y_2x_2 + y_1y_2x_3 + y_1y_2x_4,$$

$$g_4(y_1, y_2, x_1, x_2, x_3, x_4) = x_3 + x_4 + y_1y_1 + y_2x_2 + y_1x_4 + y_1y_2x_1 + y_1y_2x_2.$$

Let

$$A = Q(2,4) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and $g_i(zA) = \psi(z)$, where $z = (y_1, y_2, x_1, x_2, x_3, x_4)$, $j = 1, 2, 3, 4$. Hence $\psi_i(y_1, y_2, x_1, x_2, x_3, x_4) = g_i(y_1 + x_1 + x_2 + x_3 + x_4, y_2, x_1, x_2, x_3, x_4)$, $i = 1, 2, 3, 4$. Let $\psi$ be a nonzero linear combination of $\psi_1$, $\psi_2$, $\psi_3$, $\psi_4$ i.e. $\psi = \sum_1^4 c_j\psi_j$, $(c_1,\ c_2,\ c_3,\ c_4) \neq (0,\ 0,\ 0,\ 0)$. By Theorem 3 and Corollary 2

(i) $\psi$ is balanced,

(ii) $N_\psi \geqq 2^5 - 2^3 = 24$,

(iii) $\psi$ satisfies the $SAC$,

(iv) the degree of $\psi$ is 3,

(v) $\Psi(z) = (\psi_1(z), \psi_2(z), \psi_3(z), \psi_4(z))$, a mapping from $V_6$ to $V_4$, runs through all the $2^4$ vectors in $V_4$ each $2^2$ times while $z$ runs through $V_6$ once.

Note that the upper bound of nonlinearities of a balanced function on $V_6$ is 26 (see Corollary 3 of [11]). Thus the nonlinearity 24 of any nonzero linear combination of the these functions in this S-box is very high.

## References


[1] C. M. Adams and S. E. Tavares. Generating and counting binary bent sequences. *IEEE Transactions on Information Theory*, IT-36 No. 5:1170–1173, 1990.

[2] C. M. Adams and S. E. Tavares. The use of bent sequences to achieve higher-order strict avalanche criterion. Technical Report, TR 90-013, Department of Electrical Engineering, Queen's University, 1990.

[3] A. T. Butson. Generalized Hadamard matrices. *Proc. Amer. Math. Soc.*, 3:894–898, 1962.

[4] A. T. Butson. Relations among generalized Hadamard matrices, relative difference sets, and maximal length recurring sequences. *Canad. Math.*, 15:42–48, 1963.

[5] M. H. Dawson and S. E. Tavares. An expanded set of S-box design criteria based on information theory and its relation to differential-like attacks. In *Advances in Cryptology - EUROCRYPT'91*, volume 547, Lecture Notes in Computer Science, pages 352–367. Springer-Verlag, Berlin, Heidelberg, New York, 1991.

[6] W. de Launey. Generalised Hadamard matrices whose rows and columns form a group, volume 1036 of *Combinatorial Mathematics X, Lecture Notes in Mathematics*. Springer-Verlag, Berlin, Heidelberg, New York *154-176*, 1983.

[7] J. Detombe and S. Tavares. Constructing large cryptographically strong S-boxes. In *Advances in Cryptology - AUSCRYPT'92*, volume 718, Lecture Notes in Computer Science, pages 165–181. Springer-Verlag, Berlin, Heidelberg, New York, 1993.

[8] R. Forré. The strict avalanche criterion: Special properties of boolean functions and extended definition. In *Advances in Cryptology - CRYPTO'88*, volume 403, Lecture Notes in Computer Science, pages 450–468. Springer-Verlag, Berlin, Heidelberg, New York, 1989.

[9] S. Lloyd. Counting functions satisfying a higher order strict avalanche criterion. In *Advances in Cryptology - EUROCRYPT'89*, volume 434, Lecture Notes in Computer Science, pages 64–74. Springer-Verlag, Berlin, Heidelberg, New York, 1990.

[10] W. Meier and O. Staffelbach. Nonlinearity criteria for cryptographic functions. In *Advances in Cryptology - EUROCRYPT'89*, volume 434, Lecture Notes in Computer Science, pages 549–562. Springer-Verlag, Berlin, Heidelberg, New York, 1990.

[11] J. Seberry, X. M. Zhang, and Y. Zheng. Highly nonlinear 0-1 balanced functions satisfying strict avalanche criterion. In *Advances in Cryptology - AUSCRYPT'92*, volume 718, Lecture Notes in Computer Science, pages 145–155. Springer-Verlag, Berlin, Heidelberg, New York, 1993.

[12] J. Seberry, X. M. Zhang, and Y. Zheng. Systematic generation of cryptographically robust S-boxes. In *Proceedings of the first ACM Conference on Computer and Communications Security*, pages 172 – 182. The Association for Computing Machinery, New York, 1993.

[13] J. Seberry, X. M. Zhang, and Y. Zheng. Relationships among nonlinearity criteria. In *Advances in Cryptology - EUROCRYPT'94*, volume 950, Lecture Notes in Computer Science, pages 376–388. Springer-Verlag, Berlin, Heidelberg, New York, 1995.

[14] S. E. Tavares, M. Sivabalan, and L. E. Peppard. On the designs of SP networks from an information theoretic point of view. In *Advances in Cryptology - CRYPTO'92*, volume 740, Lecture Notes in Computer Science, pages 260–279. Springer-Verlag, Berlin, Heidelberg, New York, 1993.

[15] W. D. Wallis, A. Penfold Street, and J. Seberry Wallis. *Combinatorics: Room Squares, sum-free sets, Hadamard Matrices*, volume 292 of Lecture Notes in Mathematics. Springer-Verlag, Berlin, Heidelberg, New York, 1972.

[16] A. F. Webster. Plaintext/ciphertext bit dependencies in cryptographic system. Master's Thesis, Department of Electrical Engineering, Queen's University, Ontario, 1985.

[17] A. F. Webster and S. E. Tavares. On the design of S-boxes. In *Advances in Cryptology - CRYPTO'85*, volume 219, Lecture Notes in Computer Science, pages 523–534. Springer-Verlag, Berlin, Heidelberg, New York, 1986.