Random Correlation Matrices

Jovan Dj. Golić*

Faculty of Electrical Engineering, University of Belgrade
Bulevar Revolucije 73, 11001 Belgrade, Yugoslavia
golic@galeb.etf.bg.ac.yu

Abstract

Given a bijective vectorial Boolean function $Z_2^n \to Z_2^n$, define the correlation matrix P as an $N \times N$ matrix, $N = 2^n - 1$, whose entries are given as the squares of the correlation coefficients between nonzero linear combinations of the component Boolean functions of F and nonzero linear Boolean functions of the same n variables. Let Λ denote the number of nonzero entries in P. When F is chosen uniformly at random, the expected value and variance of Λ are determined. As a consequence, it is shown that for any $\gamma_N = o(N)$, the fraction of all F such that $\Lambda \leq N\gamma_N$ is $o(N^{-1})$.

Similar results are also obtained for partially linear F. When F is such that $(1-\varepsilon_n)n$ component functions of F are necessarily linear, where $\varepsilon_n n \to \infty$ as $n \to \infty$, it is derived that for any $\gamma_N = o(N)$, the fraction of all F such that $\Lambda \leq N\gamma_N$ is $o(N^{\varepsilon_n - 2})$.

1 Introduction

If f and g are two Boolean functions $Z_2^n \to Z_2$, $Z_2 = \{0,1\}$, then the correlation coefficient between f and g is defined as

$$\begin{aligned} c(f,g) &= \Pr(f(X) = g(X)) - \Pr(f(X) \neq g(X)) && (1)\\ &= 2^{-n} \sum_{X \in Z_2^n} (-1)^{f(X)}(-1)^{g(X)} && (2) \end{aligned}$$

where in (1) the argument values $X = (x_1, \ldots, x_n)$ are assumed to be uniformly distributed, see [4]. If one of the functions is linear, say $g = l_W \stackrel{\text{def}}{=} X \cdot W$, where

*This work was carried out while the author was with the Information Security Research Centre, Queensland University of Technology, Brisbane, Australia. This research was supported in part by the Science Fund of Serbia, grant #04M02, through the Mathematical Institute, Serbian Academy of Science and Arts.

$W = (w_1, \ldots, w_n) \in Z_2^n$ and $X \cdot W = x_1 w_1 + \cdots + x_n w_n$[1] is the *dot product* of the binary n-tuples X and W, then clearly

$$c(f, l_W) \;=\; 2^{-n} \hat{f}(W) \tag{3}$$

where

$$\hat{f}(W) \;=\; \sum_{X \in Z_2^n} (-1)^{f(X)} (-1)^{X \cdot W}, \quad W \in Z_2^n \tag{4}$$

is the well-known Walsh transform of f. Note that f can be recovered by the inverse Walsh transform

$$(-1)^{f(X)} \;=\; 2^{-n} \sum_{W \in Z_2^n} \hat{f}(W)(-1)^{X \cdot W}, \quad X \in Z_2^n. \tag{5}$$

As a result, it is noted in [4] that Parseval's theorem yields that

$$\sum_{W \in Z_2^n} c^2(f, l_W) \;=\; 1. \tag{6}$$

Let $F = (f_1, \ldots, f_n)$ denote a bijective vectorial Boolean function $Z_2^n \to Z_2^n$. Then for any $V = (v_1, \ldots, v_n) \in Z_2^n$, one can define the correlation coefficient, $c(F \cdot V, l_W)$, between a linear combination of the component Boolean functions of F determined by V and a linear Boolean function l_W determined by $W = (w_1, \ldots, w_n) \in Z_2^n$. A Boolean function is called *balanced* if it takes each of the values, 0 and 1, an equal number of times. Since F is bijective, it follows that every nonzero linear combination $F \cdot V$, $V \neq (0, \ldots, 0)$, is balanced. Also, every nonzero linear function l_W, $W \neq (0, \ldots, 0)$, is balanced too. Therefore, $c(F \cdot V, l_W) = 0$ if V or W is equal to $(0, \ldots, 0)$.

Accordingly, letting $N = 2^n - 1$, the *correlation matrix* P is defined as an $N \times N$ matrix $P = [P(i,j)], 1 \leq i, j \leq N$, where $i = \sum_{k=1}^{n} i_k 2^{n-k}$ and $j = \sum_{k=1}^{n} j_k 2^{n-k}$ are integer represenations of binary n-tuples $(i_1, \ldots, i_n)$ and $(j_1, \ldots, j_n)$, respectively, and

$$P(i,j) \;=\; c^2 \left(\sum_{k=1}^{n} j_k f_k, \sum_{k=1}^{n} i_k x_k \right). \tag{7}$$

This means that $P(i,j)$ is the square of the correlation coefficient between a nonzero linear combination, specified by j, of the component Boolean functions of F and a linear Boolean function of X specified by i. The trivial values $i = 0$ and $j = 0$ are both excluded. In view of (6), it then follows that each column of P sums up to one. On the other hand, since F is bijective, input variables can be expressed as a bijective function, F^{-1}, of the output variables, so that each row of P also sums up to one. Hence, for a bijective function F, the correlation matrix P is doubly stochastic.

Let Λ denote the number of nonzero entries in P (note that Λ is a function of P and P is a function of F). When F is chosen uniformly at random, Λ becomes

[1] The summation of Boolean functions is modulo 2 throughout.

an integer-valued random variable whose probability distribution is determined by the fractions of all F giving rise to particular values of Λ. Our objective is to derive the expected value and variance of Λ which will by Chebyshev's inequality enable us to show that for any $\gamma_N = o(N)$, the fraction of all F such that $\Lambda \leq N\gamma_N$ is $o(N^{-1})$. Similar results will also be established if F is partially linear, that is, if a given fraction of the component functions of F are necessarily linear.

2 Application

The problems considered are motivated by the Markov chain approach [5] to the so-called linear cryptanalysis [3] of product block ciphers. Given a block size n, a product block cipher is composed of a bijective round function $F : Z_2^n \to Z_2^n$ which is iterated a number of times/rounds to produce the ciphertext block from a given plaintext block, used as the input to the first round. The secret key is combined with the outputs of intermediate rounds (typically by a linear function) in such a way that the product block cipher remains bijective for any particular value of the secret key. In the case of the so-called Feistel block ciphers like the well-known DES which swap ciphertext halves at each round, the round function F is partially linear. The linear cryptanalysis is based on mutually correlated linear functions of the ciphertext and plaintext bits where the correlation coefficient should approximately be bigger than $2^{-n/2}$. It is shown in [5] that the square of this correlation coefficient can be upper-bounded by the Markov chain whose transition matrix is given as the correlation matrix P corresponding to the round function F. If the Markov chain (or simply, the transition matrix P) is ergodic (finite, aperiodic, and irreducible), then the powers of P, which is doubly stochastic, converge to the matrix whose all entries are equal to N^{-1} (for example, see [2]). As a result [5], it then follows that for a sufficiently large number of rounds, the absolute value of the correlation coefficient between any two linear functions of the ciphertext and plaintext bits, respectively, is at most $N^{-1/2} \approx 2^{-n/2}$, which renders the product block cipher immune to the linear cryptanalysis.

Unfortunately, for most practical round functions F, due to a large value of N, it is generally difficult to check whether the correlation matrix P is irreducible. However, it is shown in [5] that the ergodicity of P for a random round function F can be studied by using some results from the random graph theory, see [1] and [6]. More precisely, given P, the associated directed graph $G = (V, E)$ on a set of N vertices $V = \{v_1, v_2, \ldots, v_N\}$ is defined in such a way that there is a directed edge from v_i to v_j if and only if $P(i, j) > 0$. It then follows that the matrix P is irreducible if and only if the associated graph G is strongly connected. As a consequence of a well-known result from [6], it is then pointed out in [5] that if G is selected uniformly from all graphs with N vertices and m edges, where $m = N(\log N + \delta_N)$ and $\delta_N \to \infty$ as $N \to \infty$, then $\Pr(G$ is both aperiodic and strongly connected$) \to 1$ as N increases. When the round function F is selected uniformly at random, this is then used to argue that if $\delta_N \to \infty$ and $\Pr(\Lambda \geq N(\log N + \delta_N)) \to 1$ as $N \to \infty$, then $\Pr(P$ is ergodic$) \to 1$ as N increases. This means that most product block ciphers become immune to linear cryptanalysis after a sufficient number of rounds. Our results imply that for any $\delta_N = o(N)$ such that $\delta_N \to \infty$ as $N \to \infty$, $\Pr(\Lambda \geq N(\log N + \delta_N)) \to 1$ as $N \to \infty$, as desired.

3 Random Bijective Functions

Our main objective in this section is to derive the expected value and variance of Λ when a bijective function F is randomly chosen according to the uniform distribution. Λ can be expressed as an integer sum $\Lambda = \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_{ij}$ of binary random variables where λ_{ij} is equal to 1 if $P(i,j) > 0$ and to 0 otherwise. By using the well-known expressions [2] we first obtain

Lemma 1 The expected value and variance of Λ are respectively given by

$$\mathbf{E}[\Lambda] = \sum_{1\leq i,j\leq N} (1 - p_{ij}) \tag{8}$$

$$\mathbf{Var}[\Lambda] = \sum_{1\leq i,j\leq N} (p_{ij} - p_{ij}^2) + \sum_{1\leq i,j\leq N} \sum_{\substack{1\leq i'j'\leq N \\ (i',j')\neq(i,j)}} (p_{ij,i'j'} - p_{ij}p_{i'j'}) \tag{9}$$

where $p_{ij} = \Pr(\lambda_{ij} = 0)$ and $p_{ij,i'j'} = \Pr(\lambda_{ij} = 0, \lambda_{i'j'} = 0)$. □

We now determine these probabilities and their asymptotic behavior when 2^n is large, as is the case in cryptographic applications. To this end, we need the following two simple results. Say that a vectorial Boolean function $Z_2^n \to Z_2^m$, $m \leq n$, is *balanced* if it takes every value from Z_2^m an equal number of times. If $F = (f_1, \ldots, f_m)$ is a vectorial Boolean function $Z_2^n \to Z_2^m$, then any vectorial Boolean function $F' = (f_{i_1}, \ldots, f_{i_k})$, $1 \leq i_1 < \cdots < i_k \leq m$, $k \leq m$, is called a *subfunction* of F.

Lemma 2 If F is a bijective (balanced) vectorial Boolean function $Z_2^n \to Z_2^n$, then every subfunction of F and every balanced function (e.g., every nonzero linear combination) of the component functions of F are both balanced. If F is uniformly distributed among all bijective functions $Z_2^n \to Z_2^n$, then each fixed balanced function $Z_2^n \to Z_2^m$ (e.g., each fixed nonzero linear combination or each fixed subfunction) of the component functions of F is uniformly distributed among all functions $Z_2^n \to Z_2^m$. □

Lemma 3 The correlation coefficient between any two balanced Boolean functions f and g is equal to zero if and only if the vectorial Boolean function (f, g) is balanced.

Proof. Let f and g be Boolean functions of n variables and let

$$m(i,j) = |\{X : f(X) = i,\ g(X) = j,\ X \in Z_2^n\}|,\quad (i,j) \in Z_2^2.$$

Then in view of (1) we get

$$c(f,g) = 2^{-n}(m(0,0)+m(1,1)-m(0,1)-m(1,0)) \tag{10}$$
$$= 2^{1-n}(m(0,0)-m(0,1)) \tag{11}$$

because $m(0,0)+m(1,0)=m(0,1)+m(1,1)$, as g is balanced. Hence, $c(f,g)=0$ if and only if $m(0,0)=m(0,1)$. Since f is balanced, this is further equivalent to $m(i,j)=2^{n-2}$, $(i,j)\in Z_2^2$, that is, to (f,g) being balanced. □

Lemma 4 For any i,j, we have

$$p_{ij} \;=\; p_1 \stackrel{\text{def}}{=} \frac{\binom{2^{n-1}}{2^{n-2}}^2}{\binom{2^n}{2^{n-1}}} \;\sim\; \sqrt{\frac{8}{\pi 2^n}}\,. \tag{12}$$

Proof. According to Lemma 2, we obtain that each nonzero linear combination of the component functions of F is uniformly distributed among all $\binom{2^n}{2^{n-1}}$ balanced Boolean functions of n variables. The number of such functions that are not correlated to any given balanced Boolean function, such as a nonzero linear function, is by virtue of Lemma 3 equal to $\binom{2^{n-1}}{2^{n-2}}^2$. So, (12) directly follows. The asymptotics is easily obtained by using Stirling's formula $m! \sim \sqrt{2\pi}\, m^{m+1/2}\, e^{-m}$. □

It is more difficult to derive the pairwise probabilities $p_{ij,i'j'}$. There are three possible cases: (1) $i'=i$ and $j'\neq j$, (2) $i'\neq i$ and $j'=j$ and (3) $i'\neq i$ and $j'\neq j$. They are settled by the following two lemmas.

Lemma 5 For any $i'=i$, $j'\neq j$ or $i'\neq i$, $j'=j$, we have

$$p_{ij,i'j'} \;=\; p_2 \stackrel{\text{def}}{=} \frac{\sum_{k=0}^{2^{n-2}} \binom{2^{n-2}}{k}^4}{\binom{2^n}{2^{n-1}}} \;\sim\; \frac{8}{\pi 2^n}\,. \tag{13}$$

Proof. Let (g_1,g_2) be any balanced pair of Boolean functions of n variables. In view of Lemmas 2 and 3, it follows that for $i'\neq i$ and $j'=j$ the probability $p_{ij,i'j'}$ is equal to the relative number, p_2, of balanced Boolean functions, g_3, of n variables such that (g_1,g_3) and (g_2,g_3) are both balanced. Likewise, letting g_3 be any balanced Boolean function of n variables, for $i'=i$ and $j'\neq j$ the probability $p_{ij,i'j'}$ is equal to the relative number of balanced pairs, (g_1,g_2), of Boolean functions of n variables, such that (g_1,g_3) and (g_2,g_3) are both balanced. As the relative numbers are independent of the choice of (g_1,g_2) and g_3, respectively, they are both equal to the relative number of triples, (g_1,g_2,g_3), of Boolean functions of n variables such that (g_1,g_3) and (g_2,g_3) are both balanced, provided that (g_1,g_2) and g_3 are both balanced.

In order to derive p_2, let $m(l_1,l_2,l_3)=|\{X : (g_1(X),g_2(X),g_3(X))=(l_1,l_2,l_3), X\in Z_2^n\}|$, $(l_1,l_2,l_3)\in Z_2^3$. Then the condition that (g_1,g_2), (g_1,g_3), and (g_2,g_3) are all balanced (which implies that g_3 is balanced) can be expressed by the following system of 12 linear equations

$$\sum_{l_{s_3} \in Z_2} m(l_1, l_2, l_3) = 2^{n-2}, \qquad (l_{s_1}, l_{s_2}) \in Z_2^2 \tag{14}$$

for each $(s_1, s_2) = (1, 2)$, $(1, 3)$, and $(2, 3)$, where (s_1, s_2, s_3) is a permutation of (1, 2, 3). Now, suppose that an arbitrary pair (g_1, g_2) is fixed. Our objective is to derive the number of Boolean functions g_3 such that the system (14) is satisfied. A simple algebraic manipulation yields that $m(0,0,0) = m(0,1,1) = m(1,0,1) = m(1,1,0)$ must hold. Consequently, it follows that an 8-tuple of nonnegative integers $m(l_1, l_2, l_3)$, $(l_1, l_2, l_3) \in Z_2^3$, is a solution to the system (14), whose rank is equal to 7, if and only if $0 \leq m(0,0,0) \leq 2^{n-2}$, $m(0,0,1) = 2^{n-2} - m(0,0,0)$, $m(0,1,1) = m(1,0,1) = m(1,1,0) = m(0,0,0)$, and $m(0,1,0) = m(1,0,0) = m(1,1,1) = m(0,0,1)$. The number of different g_3 such that the system (14) is satisfied is then equal to $\sum_{k=0}^{2^{n-2}} \binom{2^{n-2}}{k}^4$, so that p_2 is given by (13).

The asymptotics can be proved by using the normal approximation to the binomial coefficients, obtained by Stirling's formula. Namely, for any fixed integer k as $\nu \to \infty$, we have

$$\binom{2\nu}{\nu+k} 2^{-2\nu} \sim \frac{1}{\sqrt{\pi\nu}} e^{-k^2/\nu} = \sqrt{\frac{2}{\nu}}\, \eta\left(\sqrt{\frac{2}{\nu}} k\right) \tag{15}$$

where $\eta(x) = \sqrt{2\pi}^{-1} e^{-x^2/2}$ is the normal probability density function, see [2]. Accordingly, for any fixed integer k, we obtain

$$\frac{\binom{2^{n-2}}{2^{n-3}+k}^4}{\binom{2^n}{2^{n-1}}} \sim \frac{64}{\pi 2^{3n/2}}\, \eta\left(2^{-n/2+3} k\right). \tag{16}$$

More generally, the approximation (15) also holds uniformly in k on any interval $-k_\nu \leq k \leq k_\nu$ where $k_\nu^3/\nu^2 \to 0$ as $\nu \to \infty$, see [2]. As a consequence, we get

$$\begin{aligned} p_2 &\sim \frac{8}{\pi 2^n} \sum_{k=-\infty}^{\infty} 2^{-n/2+3}\, \eta\left(2^{-n/2+3} k\right) \\ &\sim \frac{8}{\pi 2^n} \int_{-\infty}^{\infty} \eta(x) dx = \frac{8}{\pi 2^n}. \end{aligned} \tag{17}$$

□

Lemma 6 For any $i' \neq i$ and $j' \neq j$, we have

$$p_{ij,i'j'} = p_3 \stackrel{\text{def}}{=} \frac{1}{\binom{2^n}{2^{n-1}}^2 \binom{2^{n-1}}{2^{n-2}}^4} \sum_{\mathbf{m} \in \mathcal{M}_n} \frac{2^n!}{\prod_{l=0}^{15} m(l)} \sim \frac{8}{\pi 2^n} \tag{18}$$

where $\mathcal{M}_n$ is the set of all the nonnegative integer solutions $\mathbf{m} = (m(0), \ldots, m(15))$, where $l = \sum_{k=1}^{4} l_k 2^{4-k}$ is an integer representation of a binary 4-tuple (l_1, l_2, l_3, l_4), to the following system of 16 linear equations

$$\sum_{(l_{s_3}, l_{s_4}) \in Z_2^2} m(l_1, l_2, l_3, l_4) \quad = \quad 2^{n-2}, \qquad (l_{s_1}, l_{s_2}) \in Z_2^2 \tag{19}$$

for each $(s_1, s_2) = (1, 2), (3, 4), (1, 3)$, and $(2, 4)$, where (s_1, s_2, s_3, s_4) is a permutation of $(1, 2, 3, 4)$ such that $s_3 < s_4$.

Proof. Let (g_1, g_2, g_3, g_4) be a 4-tuple of balanced Boolean functions of n variables. In view of Lemmas 2 and 3, it follows that for $i' \neq i$ and $j' \neq j$ the probability $p_{ij,i'j'}$ is given as Q_1/Q_2 where Q_1 is the number of 4-tuples (g_1, g_2, g_3, g_4) such that (g_1, g_2), (g_3, g_4), (g_1, g_3), and (g_2, g_4) are all balanced, and Q_2 is the total number of the 4-tuples such that (g_1, g_2) and (g_3, g_4) are both balanced. It directly follows that $Q_2 = \binom{2^n}{2^{n-1}}^2 \binom{2^{n-1}}{2^{n-2}}^4$, see Lemma 4. In order to determine Q_1, let $m(l_1, l_2, l_3, l_4) = |\{X : (g_1(X), g_2(X), g_3(X), g_4(X)) = (l_1, l_2, l_3, l_4), X \in Z_2^n\}|$, $(l_1, l_2, l_3, l_4) \in Z_2^4$. Then the required condition that the four given pairs of Boolean functions are balanced can be expressed by the system of 16 linear equations (19) specified as above. Hence, $p_{ij,i',j'} = p_3$.

The asymptotics is not straightforward to prove. Let $k(l) = m(l) - 2^{n-4}$, $0 \leq l \leq 15$, and let $\mathcal{K}_n$ denote the set of all the vectors $\mathbf{k} = (k(0), \ldots, k(15))$ corresponding to vectors $\mathbf{m}$ from $\mathcal{M}_n$. For every (large) n choose a vector $\mathbf{k}$ from $\mathcal{K}_n$ in such a way that every component $k(l)$ of $\mathbf{k}$ satisfies $k(l)^3 2^{-2n} \to 0$ as 2^n increases. Then the multivariate normal approximation to the multinomial coefficients in (18) yields

$$\frac{1}{\binom{2^n}{2^{n-1}}^2 \binom{2^{n-1}}{2^{n-2}}^4} \frac{2^n!}{\prod_{l=0}^{15}(2^{n-4} + k(l))} \sim \frac{64}{(\pi 2^{n-3})^{\frac{9}{2}}} \exp\left(-\frac{1}{2}\frac{\mathbf{k}\mathbf{k}^t}{2^{n-4}}\right) \tag{20}$$

where $\mathbf{k}\mathbf{k}^t$ denotes the matrix product of the vector $\mathbf{k}$ and its transpose $\mathbf{k}^t$ (i.e., the dot product of $\mathbf{k}$ with itself). The system of linear equations has rank equal to 9, so that every $\mathbf{k}$ from $\mathcal{K}_n$ can be linearly expressed in terms of just 7 linearly independent components, e.g., $\mathbf{k}^t = \mathbf{A}\hat{\mathbf{k}}^t$, $\mathbf{k} \in \mathcal{K}_n$, where $\hat{\mathbf{k}} = (k(0), k(1), k(2), k(4), k(6), k(8), k(9))$ and $\mathbf{A}$ is the corresponding matrix given as

$$\mathbf{A} \quad = \quad \begin{bmatrix} \mathbf{I}_7 \\ \mathbf{B} \end{bmatrix} \tag{21}$$

where $\mathbf{I}_7$ is the identity matrix of dimensions 7×7 and

$$\mathbf{B} \quad = \quad \begin{bmatrix} -1 & -1 & -1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & -1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & -1 & 0 & 0 \\ -1 & 0 & -1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & -1 & 0 & -1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}. \tag{22}$$

In view of $\mathbf{kk^t} = \hat{\mathbf{k}}\mathbf{A}^t\mathbf{A}\hat{\mathbf{k}}^t$, where $\mathbf{A}^t$ denotes the transpose of $\mathbf{A}$, (20) then reduces to

$$\frac{1}{\binom{2^n}{2^{n-1}}^2 \binom{2^{n-1}}{2^{n-2}}^4} \frac{2^n!}{\prod_{l=0}^{15}(2^{n-4}+k(l))} \sim \frac{64}{(\pi 2^{n-3})^{\frac{9}{2}}} \exp\left(-\frac{1}{2}\frac{\hat{\mathbf{k}}\mathbf{A}^t\mathbf{A}\hat{\mathbf{k}}^t}{2^{n-4}}\right) \tag{23}$$

provided that every component $\hat{k}(l)$ of $\hat{\mathbf{k}}$ satisfies $\hat{k}(l)^3 2^{-2n} \to 0$ as 2^n increases. The corresponding set $\hat{\mathcal{K}}_n$ of all possible values of $\hat{\mathbf{k}}$ is then the set of all integer-valued 7-tuples such that each component $k(l)$ of $\mathbf{k}$ satisfies $k(l) \geq -2^{n-4}$ (i.e., $m(l) \geq 0$). Note that the approximation (23) also holds uniformly in $\hat{\mathbf{k}}$ on any set $-\hat{k}_n(l) \leq \hat{k}(l) \leq \hat{k}_n(l)$, $1 \leq l \leq 7$, where $\hat{k}_n(l)^3 2^{-2n} \to 0$ as 2^n increases. Consequently, letting $\eta(\mathbf{x}) = (2\pi)^{-7/2}\sqrt{\det \mathbf{A}^t\mathbf{A}}\ \exp\left(-\frac{1}{2}\mathbf{x}\mathbf{A}^t\mathbf{A}\mathbf{x}^t\right)$ denote a multivariate normal distribution of seven variables, we obtain

$$\begin{aligned} p_3 &\sim \frac{8}{\pi 2^n}\frac{64}{(2\pi)^{\frac{7}{2}}} \sum_{\hat{\mathbf{k}} \in Z^7} 2^{-(n-4)\frac{7}{2}} \exp\left(-\frac{1}{2}\frac{\hat{\mathbf{k}}\mathbf{A}^t\mathbf{A}\hat{\mathbf{k}}^t}{2^{n-4}}\right) \\ &\sim \frac{8}{\pi 2^n}\frac{64}{\sqrt{\det \mathbf{A}^t\mathbf{A}}} \int_{\mathbf{x}} \eta(\mathbf{x})d\mathbf{x} = \frac{8}{\pi 2^n}, \end{aligned} \tag{24}$$

as direct computation yields that $\det \mathbf{A}^t\mathbf{A} = 64^2$. □

Lemmas 4–6 essentially show the asymptotic pairwise independence of the zero entries in the correlation matrix when 2^n is large. Consequently, in light of Lemma 1 and Chebyshev's inequality $\Pr(|\Lambda - \mathbf{E}[\Lambda]| \geq \varepsilon) \leq \mathbf{Var}[\Lambda]/\varepsilon^2$, Lemmas 4–6 result in

Theorem 1 For a random bijective function F chosen according to the uniform distribution, the expected value and variance of the number Λ of nonzero entries in the correlation matrix P satisfy

$$\mathbf{E}[\Lambda] = N^2(1-p_1) \sim N^2\left(1-\sqrt{\frac{8}{\pi N}}\right) \sim N^2 \tag{25}$$

$$\begin{aligned} \mathbf{Var}[\Lambda] &= N^2 p_1(1-p_1) + 2N^2(N-1)(p_2-p_1^2) + N^2(N-1)^2(p_3-p_1^2) \\ &= o(N^3) \end{aligned} \tag{26}$$

where $N = 2^n - 1$. For any $\gamma_N = o(N)$,

$$\Pr(\Lambda \leq N\gamma_N) = o\left(\frac{1}{N}\right). \tag{27}$$

□

4 Random Partially Linear Bijective Functions

For Feistel block ciphers like DES the probabilistic model of a random round function F is not quite appropriate because one half of the component Boolean functions of F are identity mappings. Our objective in this section is to derive the expected value and variance of Λ when a partially linear bijective function F is randomly chosen according to the uniform distribution. Recall that F is called partially linear if a given fraction of the component functions of F are necessarily linear.

Lemma 7 Let F be a partially linear bijective function $Z_2^n \to Z_2^n$ consisting of a set F_1 of n_1 linear component functions and a set F_2 of n_2 arbitrary component functions, where $n = n_1 + n_2$. Then the number Λ of nonzero entries in the correlation matrix P is given as

$$\Lambda \quad = \quad N_1\Lambda_2 + \Lambda_2 + N_1 \qquad (28)$$

where $N_1 = 2^{n_1} - 1$ and Λ_2 is the number of nonzero entries in the columns of P corresponding to all $N_2 = 2^{n_2} - 1$ nonzero linear combinations of the n_2 functions from F_2.

Proof. Since the number of nonzero entries in the columns of P corresponding to all N_1 nonzero linear combinations of the n_1 linear functions from F_1 is N_1, it remains to show that $N_1\Lambda_2$ is the number of nonzero entries in the N_1N_2 columns of P corresponding to all the linear combinations that necessarily include at least one linear function from F_1 and at least one function from F_2. Let $\sum l_i + \sum f_j$ denote any such linear combination. From (1) it follows that for each nonzero linear function l of n variables, $c(\sum l_i + \sum f_j, l) = c(l + \sum l_i + \sum f_j, 0)$. As l ranges through all nonzero linear functions of n variables, $l + \sum l_i$ ranges through all linear functions of n variables including the constant zero one which is substituted for $\sum l_i$. According to Lemma 2, as F is balanced, then $c(\sum f_j, \sum l_i) = 0$, and on the other hand, as $\sum f_j$ is balanced, then $c(\sum f_j, 0) = 0$ too. Consequently, the column of P corresponding to $\sum l_i + \sum f_j$ is just a permutation of the column corresponding to $\sum f_j$. Hence for any given $\sum l_i$, the number of nonzero entries in all the columns of P corresponding to $\sum l_i + \sum f_j$ for all N_2 nonzero linear combinations $\sum f_j$ is equal to Λ_2. The total number of the considered nonzero entries, for all N_1 different $\sum l_i$, is then $N_1\Lambda_2$. □

Lemma 7 shows that for a partially linear bijective function F, Λ does not depend on the particular choice of linear functions but only on the number of them. So, when one picks such F uniformly at random, the probability distribution of Λ is determined by the probability distribution of Λ_2. Accordingly, by using Lemma 1 we obtain the expressions for the expected value and variance of Λ similar to those given in Theorem 1, except that they depend on both $N_1 = 2^{n_1} - 1$ and $N_2 = 2^{n_2} - 1$ (note that $N = 2^n - 1 = N_1N_2 + N_1 + N_2$). If we assume that n_1 and n_2 are given as functions of n such that $n_2 \to \infty$ as $n \to \infty$, then we get

Theorem 2 Let F be a partially linear bijective function $Z_2^n \to Z_2^n$ with $n_1 = (1-\varepsilon_n)n$ linear component functions and let $n_2 = \varepsilon_n n \to \infty$ as $n \to \infty$. Then for a random F chosen according to the uniform distribution, the expected value and variance of the number Λ of nonzero entries in the correlation matrix P satisfy

$$\mathbf{E}[\Lambda] \sim N^2 \tag{29}$$
$$\mathbf{Var}[\Lambda] = o\left(N^{2+\varepsilon_n}\right) \tag{30}$$

where $N = 2^n - 1$. For any $\gamma_N = o(N)$

$$\Pr\left(\Lambda \le N\gamma_N\right) = o\left(\frac{1}{N^{2-\varepsilon_n}}\right). \tag{31}$$

□

Comparing Theorems 1 and 2 we see that the expected value of Λ for partially linear F is asymptotically the same as for arbitrary F, whereas the variance is reduced.

References

[1] B. Bollobás, *Random Graphs*. London: Academic Press, 1985.

[2] W. Feller, *An Introduction to Probability Theory and its Applications*. New York: Wiley, 3. edition, vol. 1, 1968.

[3] M. Matsui, "Linear cryptanalysis method for DES cipher," Advances in Cryptology - Eurocrypt '93, *Lecture Notes in Computer Science*, vol. 765, T. Helleseth ed., Springer-Verlag, pp. 386-397, 1994.

[4] W. Meier and O. Staffelbach, "Nonlinearity criteria for cryptographic functions," Advances in Cryptology - Eurocrypt '89, *Lecture Notes in Computer Science*, vol. 434, J.-J. Quisquater and J. Vandewalle eds., Springer-Verlag, pp. 549-562, 1990.

[5] L. O'Connor and J. Dj. Golić, "A unified Markov approach to differential and linear cryptanalysis," Advances in Cryptology - Asiacrypt '94, *Lecture Notes in Computer Science*, vol. 917, J. Pieprzyk and R. Safavi-Naini eds., Springer-Verlag, pp. 387-397, 1995.

[6] I. Palásti, "On the strong connectedness of random graphs," *Studia Sci. Math. Hungar.*, vol. 1, pp. 205-214, 1966.

(Received 3/2/97)