Correlations Between (Nonlinear) Combiners of Input and Output of Random Functions and Permutations

Subhabrata Samajder and Palash Sarkar
Applied Statistics Unit
Indian Statistical Institute
203, B.T.Road, Kolkata, India - 700108.
subhabrata.samajder@gmail.com, palash@isical.ac.in

May 15, 2019

Abstract

Linear cryptanalysis considers correlations between linear input and output combiners for block ciphers and stream ciphers. Daemen and Rijmen (2007) had obtained the distributions of the correlations between linear input and output combiners of uniform random functions and uniform random permutations. The present work generalises these results to obtain the distributions of the correlations between arbitrary input and output combiners of uniform random functions and uniform random permutations.

Keywords: correlation, uniform random function, uniform random permutation.
Mathematics Subject Classification (2010): 94A60, 68P25, 62P99.

1 Introduction

One of the basic tools for analysing symmetric key ciphers is a possible correlation between linear combinations of the input and output of a primitive. If this correlation is different from that of an idealised version of the primitive, then a distinguishing attack becomes possible. Determining whether a distinguishing attack is indeed possible requires the knowledge of the distributions of correlations for the idealised primitives. Two kinds of idealised primitives are usually considered, namely uniform random functions and uniform random permutations. For example, a uniform random permutation is an idealisation of a block cipher while a uniform random function is an idealisation of the state to keystream map in a stream cipher.

The distributions of the correlations between linear combinations of input and output for uniform random functions and uniform random permutations were derived in [1]. For the case of uniform random permutations, the distribution was earlier stated without proof in [3].

Our Contributions

This work extends the results of Daemen and Rijmen [1] by considering the correlation between arbitrary combiners of the input and output of uniform random functions and uniform random permutations. For any input combiner and any output combiner, the complete distributions of the correlations in the two cases are derived. The results are more conveniently stated in terms of the weight of the XOR of the input and the output combiners. For the case of a uniform random function, we show that the distribution of the weight is given by the convolution of two binomial distributions; if the output combiner is balanced, then the weight distribution is given by a binomial distribution. For the case of a uniform random permutation, we show that the distribution of the weight is given by a hypergeometric distribution.

Our approach to proving the results is different from that in [1]. The proofs in [1] consist essentially of counting Boolean functions. Instead we have used direct probability arguments. This yields proofs which are simple and at the same time work for arbitrary combiners.

2 Preliminaries

An m-variable Boolean function f is a map $f : \{0,1\}^m \rightarrow \{0,1\}$. The weight $\mathsf{wt}(f)$ of f is defined to be the cardinality of the support of f, i.e.,

$$\mathsf{wt}(f) = \#\{\alpha \in \{0,1\}^m : f(\alpha) = 1\}.$$

The function f is said to be balanced if $\mathsf{wt}(f) = 2^{m-1}$.

Let $f, g : \{0,1\}^m \rightarrow \{0,1\}$ be two Boolean functions. By $f \oplus g$ we denote the Boolean function $h : \{0,1\}^m \rightarrow \{0,1\}$ where $h(\alpha) = f(\alpha) \oplus g(\alpha)$ for all $\alpha \in \{0,1\}^m$. The correlation between f and g is denoted as $C(f,g)$ and is defined to be

$$C(f,g) = 1 - \frac{\mathsf{wt}(f \oplus g)}{2^{m-1}}.$$

An (m,n) function S is a map $S : \{0,1\}^m \rightarrow \{0,1\}^n$. Let $\phi : \{0,1\}^m \rightarrow \{0,1\}$ and $\psi : \{0,1\}^n \rightarrow \{0,1\}$. Given S, ϕ and ψ, we define a Boolean function

$$f_S[\phi,\psi] : \{0,1\}^m \rightarrow \{0,1\}, \text{ where } f_S[\phi,\psi](\alpha) = \phi(\alpha) \oplus \psi(S(\alpha)). \tag{1}$$

The function ϕ is a combiner of the input of S while the function ψ is a combiner of the output of S. Both $\phi(\cdot)$ and $\psi(S(\cdot))$ are m-variable Boolean functions. So, it is meaningful to talk about the correlation between these two functions. This correlation will be denoted as $C_S(\phi,\psi)$ and is equal to

$$C_S(\phi,\psi) \quad = \quad = 1 - \frac{\mathsf{wt}(f_S[\phi,\psi])}{2^{m-1}}. \tag{2}$$

So, $C_S(\phi,\psi)$ measures the correlation between the combiner of the input as given by ϕ and the combiner of the output as given by ψ. From (2), determining $C_S(\phi,\psi)$ essentially boils down to determining $\mathsf{wt}(f_S[\phi,\psi])$.

Probability distributions: $\mathsf{Ber}(p)$ denotes the Bernoulli distribution with probability of success p; $\mathsf{Bin}(k,p)$ denotes the binomial distribution with k trials and probability of success p; $\mathsf{HG}(k,k_1,s)$ denotes the hypergeometric distribution corresponding to a population of size k of which k_1 are of a specified type and $k - k_1$ are of a different type and a sample of size s is drawn without repetition.

3 Case of Uniform Random Function

Let ρ be a function picked uniformly at random from the set of all functions from $\{0,1\}^m$ to $\{0,1\}^n$. Such a ρ is a uniform random (m,n) function. An equivalent way to view ρ is the following. Let $\alpha_0, \ldots, \alpha_{2^m-1}$ be an enumeration of $\{0,1\}^m$. Let $X_i = \rho(\alpha_i)$, $i = 0, \ldots, 2^m - 1$. Then the random variables $X_0, \ldots, X_{2^m-1}$ are independent and uniformly distributed over $\{0,1\}^n$.

Theorem 1. *Let ρ be a uniform random (m, n) function. Let ϕ and ψ be m and n-variable Boolean functions respectively. Let $\alpha_0, \ldots, \alpha_{2^m-1}$ be an enumeration of $\{0,1\}^m$. For $0 \leq i \leq 2^m - 1$, define $W_i = f_\rho[\phi, \psi](\alpha_i)$. Then $W_i \sim \mathsf{Ber}(p_i)$, where*

$$p_i \quad = \quad \frac{\mathsf{wt}(\psi) + \phi(\alpha_i)(2^n - 2\mathsf{wt}(\psi))}{2^n}. \tag{3}$$

If ψ is a balanced Boolean function, then $W_i \sim \mathsf{Ber}(1/2)$.

Proof. Let $X_i = \rho(\alpha_i)$. Since ρ is a uniform random function, X_i is uniformly distributed over $\{0,1\}^n$. We have

$$W_i \quad = \quad f_\rho[\phi, \psi](\alpha_i) = \phi(\alpha_i) \oplus \psi(\rho(\alpha_i)) = \phi(\alpha_i) \oplus \psi(X_i).$$

Let $Y_i = \psi(X_i)$. Then Y_i is a binary valued random variable where Y_i takes the value 1 if and only if X_i lies in the support of ψ. Since X_i is uniformly distributed over $\{0,1\}^n$, the probability that X_i lies in the support of ψ is $\mathsf{wt}(\psi)/2^n$. So, $\Pr[Y_i = 1] = \mathsf{wt}(\psi)/2^n$ and $\Pr[Y_i = 0] = (2^n - \mathsf{wt}(\psi))/2^n$. Consequently,

$$\begin{aligned}
\Pr[W_i = 1] &= \Pr[\phi(\alpha_i) \oplus \psi(X_i) = 1] \\
&= \Pr[Y_i = 1 \oplus \phi(\alpha_i)] \\
&= \frac{(1 - \phi(\alpha_i))\mathsf{wt}(\psi) + \phi(\alpha_i)(2^n - \mathsf{wt}(\psi))}{2^n} \\
&= \frac{\mathsf{wt}(\psi) + \phi(\alpha_i)(2^n - 2\mathsf{wt}(\psi))}{2^n} \\
&= p_i.
\end{aligned}$$

This shows that W_i follows $\mathsf{Ber}(p_i)$.

If ψ is a balanced Boolean function, then $\mathsf{wt}(\psi) = 2^{n-1}$ in which case $p_i = 1/2$ and so W_i follows $\mathsf{Ber}(1/2)$. □

We are interested in the weight of the function $f_\rho[\phi, \psi]$.

Proposition 1. *Let ρ be a uniform random (m, n) function. Let ϕ and ψ be m and n-variable Boolean functions respectively. Let $\alpha_0, \ldots, \alpha_{2^m-1}$ be an enumeration of $\{0,1\}^m$ and $W_i = f_\rho[\phi, \psi](\alpha_i)$. Let $W = \mathsf{wt}(f_\rho[\phi, \psi])$. Then $W = \sum_{i=0}^{2^m-1} W_i$.*

Proof. The following calculation shows the result.

$$W \quad = \quad \mathsf{wt}(f_\rho[\phi, \psi]) = \#\{\alpha_i : f_\rho[\phi, \psi](\alpha_i) = 1\} = \#\{i : W_i = 1\} = \sum_{i=0}^{2^m-1} W_i.$$

□

Theorem 2. *Let ρ be a uniform random (m, n) function. Let ϕ and ψ be m and n-variable Boolean functions respectively. Then*

$$\Pr\left[\mathsf{wt}(f_\rho[\phi, \psi]) = w\right] \quad = \quad \sum_{t=0}^{w} \binom{w_0}{t} \binom{2^m - w_0}{w - t} \left(\frac{w_1}{2^n}\right)^{2^m - w - w_0 + 2t} \left(1 - \frac{w_1}{2^n}\right)^{w_0 + w - 2t} \tag{4}$$

where $w_0 = \mathsf{wt}(\phi)$ and $w_1 = \mathsf{wt}(\psi)$.

Further, if ψ is a balanced Boolean function, i.e., $w_1 = 2^{n-1}$, then $\mathsf{wt}(f_\rho[\phi, \psi]) \sim \mathsf{Bin}(2^m, 1/2)$.

Proof. Let $\alpha_0, \ldots, \alpha_{2^m-1}$ be an enumeration of $\{0,1\}^m$ and $X_i = \rho(\alpha_i)$ as in Theorem 1. Note

$$W_i \quad = \quad f_\rho[\phi, \psi](\alpha_i) = \phi(\alpha_i) \oplus \psi(X_i).$$

Since the random variables $X_0, \ldots, X_{2^m-1}$ are independent, so are the random variables $W_0, \ldots, W_{2^m-1}$.

From Proposition 1, $\mathsf{wt}(f_\rho[\phi, \psi]) = W = \sum_{i=0}^{2^m-1} W_i$ where $W_i \sim \mathsf{Ber}(p_i)$ with p_i given by (3). Note that p_i takes either the value $\mathsf{wt}(\psi)/2^n$ or $(2^n - \mathsf{wt}(\psi))/2^n$ according as $\phi(\alpha_i)$ equals 0 or 1. So, $W_0, \ldots, W_{2^m-1}$ is a sequence of 2^m Poisson trials, where each W_i either follows $\mathsf{Ber}\left(\frac{\mathsf{wt}(\psi)}{2^n}\right)$ or follows $\mathsf{Ber}\left(\frac{2^n - \mathsf{wt}(\psi)}{2^n}\right)$. Thus, W can be written as the sum of two binomially distributed random variables Z_1 and Z_2, i.e., $W = Z_1 + Z_2$, where

$$Z_1 \sim \mathsf{Bin}\left(\mathsf{wt}(\phi), \frac{\mathsf{wt}(\psi)}{2^n}\right) \quad \text{and} \quad Z_2 \sim \mathsf{Bin}\left(2^m - \mathsf{wt}(\phi), \frac{2^n - \mathsf{wt}(\psi)}{2^n}\right).$$

Consequently, the distribution of W is given by the convolution of these two binomial distributions. Simplifying the expression for the convolution, we obtain the stated result. □

The special case of Theorem 2 where ϕ and ψ are non-trivial linear functions was given in [1]. The proof of this result in [1] is a counting argument which uses the fact that when ψ is a non-trivial balanced function, $\mathsf{wt}(\psi) = 2^{n-1}$. So, the proof in [1] covers the case of ψ being a balanced function. Theorem 2 provides the general result without any conditions on ψ (or ϕ).

4 Case of Uniform Random Permutation

Let $m = n$ and we consider the set of all bijections from $\{0,1\}^n$ to itself, i.e., the set of all permutations of $\{0,1\}^n$. There are $2^n!$ such permutations.

Proposition 2. *Let S be any permutation of $\{0,1\}^n$; let ϕ and ψ be n-variable Boolean functions. Let x be an integer such that $0 \leq x \leq \min(\mathsf{wt}(\phi), \mathsf{wt}(\psi))$. Then*

$$\#\{\alpha : \phi(\alpha) = 1 \text{ and } \psi(S(\alpha)) = 1\} = x$$

if and only if

$$\mathsf{wt}(f_S[\phi, \psi]) = \mathsf{wt}(\phi) + \mathsf{wt}(\psi) - 2x.$$

Proof. Define

$$\begin{aligned} A_{0,0} &= \{\alpha : \phi(\alpha) = 0, \psi(S(\alpha)) = 0\}; \\ A_{0,1} &= \{\alpha : \phi(\alpha) = 0, \psi(S(\alpha)) = 1\}; \\ A_{1,0} &= \{\alpha : \phi(\alpha) = 1, \psi(S(\alpha)) = 0\}; \\ A_{1,1} &= \{\alpha : \phi(\alpha) = 1, \psi(S(\alpha)) = 1\}. \end{aligned}$$

The sets $A_{0,0}, A_{0,1}, A_{1,0}$ and $A_{1,1}$ are mutually disjoint; $A_{0,0} \cup A_{0,1} = \{\alpha : \phi(\alpha) = 0\}$; $A_{1,0} \cup A_{1,1} = \{\alpha : \phi(\alpha) = 1\}$ and so

$$\begin{aligned} \#A_{0,0} + \#A_{0,1} &= 2^n - \mathsf{wt}(\phi), \\ \#A_{1,0} + \#A_{1,1} &= \mathsf{wt}(\phi). \end{aligned} \tag{5}$$

Further, $A_{0,0} \cup A_{1,0} = \{\alpha : \psi(S(\alpha)) = 0\}$. Since S is a permutation, $\{\alpha : \psi(S(\alpha)) = 0\} = \{\beta : \psi(\beta) = 0\}$. So, $A_{0,0} \cup A_{1,0} = \{\beta : \psi(\beta) = 0\}$ and similarly, $A_{0,1} \cup A_{1,1} = \{\beta : \psi(\beta) = 1\}$ leading to

$$\begin{array}{rcl} \#A_{0,0} + \#A_{1,0} & = & 2^n - \mathsf{wt}(\psi), \\ \#A_{0,1} + \#A_{1,1} & = & \mathsf{wt}(\psi). \end{array} \tag{6}$$

Equations (5) and (6) imply that $\#A_{1,1} = x$ if and only if $\#A_{0,1} + \#A_{1,0} = \mathsf{wt}(\phi) + \mathsf{wt}(\psi) - 2x$.

Note that the support of $f_S[\phi, \psi]$ is $A_{0,1} \cup A_{1,0}$ and $A_{1,1} = \{\alpha : \phi(\alpha) = 1, \psi(S(\alpha)) = 1\}$. So, $\#\{\alpha : \phi(\alpha) = 1, \psi(S(\alpha)) = 1\} = x$ if and only if $\mathsf{wt}(f_S[\phi, \psi]) = \mathsf{wt}(\phi) + \mathsf{wt}(\psi) - 2x$. □

From Proposition 2, given the functions ϕ and ψ, the possible weights that $f_S[\phi, \psi]$ can take for any permutation S of $\{0,1\}^n$ are the elements of the set

$$\{\mathsf{wt}(\phi) + \mathsf{wt}(\psi) - 2x : 0 \leq x \leq \min(\mathsf{wt}(\phi), \mathsf{wt}(\psi))\}. \tag{7}$$

Suppose π is picked uniformly from the set of all permutations of $\{0,1\}^n$. We are interested in the probability that $f_\pi[\phi, \psi]$ takes a value from the set given by (7).

Theorem 3. *Let π be a uniform random permutation of $\{0,1\}^n$; let ϕ and ψ be n-variable Boolean functions. Then for $0 \leq x \leq \min(\mathsf{wt}(\phi), \mathsf{wt}(\psi))$,*

$$\Pr[\mathsf{wt}(f_\pi[\phi, \psi]) = \mathsf{wt}(\phi) + \mathsf{wt}(\psi) - 2x] \quad = \quad \frac{\binom{\mathsf{wt}(\phi)}{x}\binom{2^n - \mathsf{wt}(\phi)}{\mathsf{wt}(\psi) - x}}{\binom{2^n}{\mathsf{wt}(\psi)}}. \tag{8}$$

If both ϕ and ψ are balanced functions, then

$$\Pr[\mathsf{wt}(f_\pi[\phi, \psi]) = \mathsf{wt}(\phi) + \mathsf{wt}(\psi) - 2x] \quad = \quad \frac{\binom{2^{n-1}}{x}^2}{\binom{2^n}{2^{n-1}}}. \tag{9}$$

Proof. Let $\alpha_0, \ldots, \alpha_{2^n-1}$ be an enumeration of $\{0,1\}^n$ and let $X_i = \pi(\alpha_i)$. Unlike the case where π is a uniform random function, the random variables $X_0, \ldots, X_{2^n-1}$ are not independent. Instead, it is more convenient to view these random variables in the following manner. Consider an urn containing balls labelled $\alpha_0, \ldots, \alpha_{2^n-1}$. Balls are picked one by one from the urn *without replacement* and we number the trials from 0 to $2^n - 1$. Then the random variable X_i is the label of the ball picked in trial number i.

Consider the random Boolean function $g(\alpha) = \psi(\pi(\alpha))$. A Boolean function is defined by its support. So, it is sufficient to choose $\mathsf{wt}(\psi)$ balls from the urn and let the labels of these balls define the support of g. From Proposition 2, the probability that $\mathsf{wt}(f_\pi[\phi, \psi]) = \mathsf{wt}(\phi) + \mathsf{wt}(\psi) - 2x$ is equal to the probability that the cardinality of the set

$$A_{1,1} = \{\alpha : \phi(\alpha) = 1 \text{ and } \psi(\pi(\alpha)) = 1\} = \{\alpha : \phi(\alpha) = 1 \text{ and } g(\alpha) = 1\}$$

is x.

To obtain this probability, we consider the following equivalent random experiment. As before, consider the urn containing balls labelled $\alpha_0, \ldots, \alpha_{2^n-1}$. Further, say that a ball labelled α_i is 'red' if $\phi(\alpha_i) = 1$ and otherwise it is 'black'. Now, consider that $\mathsf{wt}(\psi)$ balls are drawn from this urn which defines the support of g. The event that we are interested in is that x of these $\mathsf{wt}(\psi)$ are 'red' while the other $\mathsf{wt}(\psi) - x$ are 'black'. The probability of this event is the probability that $\#A_{1,1} = x$ which is given by the right hand side of (8). Then (8) follows from Proposition 2.

In the case where both ϕ and ψ are balanced functions, both their weights are equal to 2^{n-1}. So, substituting 2^{n-1} for $\mathsf{wt}(\phi)$ and $\mathsf{wt}(\psi)$ in (8) and using $\binom{2^{n-1}}{2^{n-1}-x} = \binom{2^{n-1}}{x}$ yields (9). □

The expression given on the right hand side of (8) is the probability mass function of the hypergeometric distribution. In the special case where ϕ and ψ are non-trivial linear functions, the distribution given by (9) was proved in [1].

5 Conclusion

In this paper, we have obtained the distributions of the correlations between arbitrary input and output combiners of uniform random functions and uniform random permutations. These generalise earlier results by Daemen and Rijmen [1] who had considered only linear combiners.

References

[1] Joan Daemen and Vincent Rijmen. Probability distributions of correlation and differentials in block ciphers. *J. Mathematical Cryptology*, 1(3):221–242, 2007.

[2] Mitsuru Matsui. Linear Cryptanalysis Method for DES Cipher. In *Advances in Cryptology–EUROCRYPT'93*, pages 386–397. Springer, 1993.

[3] Luke O'Connor. Properties of linear approximation tables. In Bart Preneel, editor, *Fast Software Encryption: Second International Workshop. Leuven, Belgium, 14-16 December 1994, Proceedings*, volume 1008 of *Lecture Notes in Computer Science*, pages 131–136. Springer, 1994.