Distinguishing Error of Nonlinear Invariant Attacks

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

October 1, 2018

Abstract

Linear cryptanalysis considers correlations between linear input and output combiners for block ciphers and stream ciphers. Daeman and Rijmen (2007) had obtained the distributions of the correlations between linear input and output combiners of uniform random functions and uniform random permutations. Our first contribution is to generalise these results to obtain the distributions of the correlations between arbitrary input and output combiners of uniform random functions and uniform random permutations. Recently, Todo et al. (2018) have proposed nonlinear invariant attacks which consider correlations between nonlinear input and output combiners for a key-alternating block cipher. In its basic form, a nonlinear invariant attack is a distinguishing attack. The second and the main contribution of this paper is to obtain precise expressions for the errors of nonlinear invariant attacks in distinguishing a key-alternating cipher from either a uniform random function or a uniform random permutation.
Keywords: correlation, uniform random function, uniform random permutation, block cipher, nonlinear invariant attack, distinguishing attack, error probability.

1 Introduction

Let $S : \{0,1\}^m \rightarrow \{0,1\}^n$ be a function arising in a context of symmetric key cryptography. Two important examples are the state to keystream map of a stream cipher, and the encryption function of a block cipher, for which $m = n$. The goal of a distinguishing attack is to be able to distinguish a real cryptographic primitive from an idealised primitive. The idealised primitive could be a uniform random function ρ from $\{0,1\}^m$ to $\{0,1\}^n$ or, for $m = n$, it could be a uniform random permutation π of $\{0,1\}^n$.

A distinguishing attack based on correlation between input and output combiners proceeds as follows. Let $\phi : \{0,1\}^m \rightarrow \{0,1\}$ and $\psi : \{0,1\}^n \rightarrow \{0,1\}$ be two functions. The function ϕ serves as a combiner of the input of S while the function ψ serves as a combiner of the output of S. The correlation between input and output combiners is the correlation between ϕ and $\psi \circ S$. This correlation is captured by considering the weight of the function $f_S : \{0,1\}^m \rightarrow \{0,1\}$ defined by $f_S(\alpha) = \phi(\alpha) \oplus \psi(S(\alpha))$. Suppose it is possible to find some property of S such that the function f_S has a nature which is different from f_ρ or f_π. Then such a property forms the basis of distinguishing S from either ρ or π.

Obtaining the nature of f_S requires a considerable amount of ingenuity, and is obtained by carefully studying the overall design and the internal structure of S. On the other hand, the nature of f_ρ and f_π are obtained mathematically. To determine the success probability of an attack, it is important to have sufficient information about both f_S and either f_ρ or f_π. In this paper, we will be concerned with properties of f_ρ and f_π.

Linear cryptanalysis: Distinguishing attacks based on linear cryptanalysis [7] is the classical example of the above scenario. For such an attack, the functions ϕ and ψ are linear functions. Linear cryptanalysis has an extensive history and has been successfully applied to both block and stream ciphers. When ϕ and ψ are linear functions, precise distributions of the weights of f_ρ and f_π have been obtained by Daeman and Rijmen [4]. For the case of f_π, the distribution was earlier stated without proof in [9]. The results of [4] have formed the basis for an alternative formulation of the wrong key randomisation hypothesis in linear cryptanalysis [3] and has been followed up in later works [2, 1].

Nonlinear invariant attack: Nonlinear combiners of inputs and outputs of a key alternating cipher arise in the context of nonlinear invariant attack which has been introduced by Todo et al. [10]. Suppose $n = m$ and S is an r-round key alternating cipher $E_K : \{0,1\}^n \to \{0,1\}^n$. The crux of a nonlinear invariant attack is that there may exist an n-variable Boolean function g and a class of weak keys K such that for any plaintext P, $g(P) \oplus g(E_K(P))$ is a constant which is independent of P. Such a g is called a nonlinear invariant. The existence of nonlinear invariants and weak keys have been shown for practical block ciphers SCREAM, iSCREAM and Midori64 [10]. Nonlinear approximations have been previously studied by Herpes et al. [5] and Knudsen and Robshaw [6].

Our Contributions

This work makes two contributions.

The first contribution is to extend the results of Daemen and Rijmen [4] by considering correlation between arbitrary combiners of the input and output of uniform random functions and uniform random permutations. In other words, we allow ϕ and ψ to be arbitrary Boolean functions and obtain the distributions of the weights of f_ρ and f_π. For the case of a uniform random function ρ, if the output combiner ψ is balanced, then we prove that this weight follows the binomial distribution; on the other hand, if the output combiner is not balanced, then we derive bounds on the probability that the weight deviates from its expected value. In the case of a uniform random permutation π, we show that the distribution of the weights of f can be expressed in terms of the hypergeometric distribution.

Our approach to proving the results is different from that in [4]. The proofs in [4] are counting arguments and essentially consist of counting Boolean functions under certain restrictions. While this approach works when the input and output combiners are linear functions, we found it difficult to extend this approach for arbitrary Boolean functions. Instead we have used direct probability arguments. This yields proofs which are simple and at the same time work for arbitrary combiners.

The second and the main contribution of this work is to perform an analysis of the distinguishing error of nonlinear invariant attacks. The goal is to be able to distinguish E_K from a uniform random permutation π of $\{0,1\}^n$ (or, from a uniform random function ρ). Suppose g is a nonlinear invariant for E_K. Further, suppose that distinct plaintexts $P_1, \ldots, P_N$ are used by the distinguisher. Then if K is a weak key, $g(P_1) \oplus g(E_K(P_1)) = \cdots = g(P_N) \oplus g(E_K(P_N))$. To be able to construct a distinguisher it is required to determine the probability ε that $g(P_1) \oplus g(\pi(P_1)) = \cdots = g(P_N) \oplus g(\pi(P_N))$. The distinguisher can make one-sided error and the probability of this error is precisely ε.

We consider the following more general problem. (This generalisation has been mentioned in Section 7 of [10].) Let g_0 and g_r be any two n-variable Boolean functions. We determine the probability that $g_0(P_1) \oplus g_r(\pi(P_1)) = \cdots = g_0(P_N) \oplus g_r(\pi(P_N))$. This is done in two cases, namely, when $P_1, \ldots, P_N$ are chosen under uniform random sampling without replacement and when $P_1, \ldots, P_N$ are distinct n-bit values without any randomness. Further, these probabilities are also computed when π is replaced by a uniform random function ρ. Our analysis provides expressions for the error probabilities of the corresponding distinguishers. Such an analysis was not performed in [10]. Some of the consequences of our analysis are as follows.

1. It turns out that the error probability considered in [10] is that of distinguishing E_K from a uniform random function. The error probability of distinguishing E_K from a uniform random permutation is obtained here for the first time.
2. The general form of the error probabilities are derived without any restriction on g_0 and g_r. When g_0 and g_r are balanced functions, we prove the following two results.
   (a) The error in distinguishing from a uniform random function is $1/2^{N-1}$.
   (b) The error in distinguishing from a uniform random permutation is at least as large as the error in distinguishing from a uniform random function. This is a consequence of Jensen's inequality. For moderate values of N, the error in distinguishing from a uniform random permutation is almost the same as the error in distinguishing from a uniform random function.

Structure of the Paper

In Section 2, we provide the generalisation of the results of Daemen and Rijmen which appear in [4]. Section 3 provides a background of nonlinear invariant attacks as distinguishing attacks and defines the relevant distinguishing errors. Section 4 provides the analysis of the error in distinguishing from a uniform random permutation while Section 5 provides the analysis of error in distinguishing from a uniform random permutation. Appendix B provides an alternative expression for the later error. Some computational results are provided in Section 6.

2 Correlation Between Input and Output Combiners

In this section, we consider the distribution of correlation between input and output combiners of uniform random functions and uniform random permutations. The case of uniform random function is analysed in Section 2.1 and the case of uniform random permutation is analysed in Section 2.2. Before proceeding, we introduce some basic concepts and notation.

For two binary strings α and β of the same length, $\alpha \oplus \beta$ will denote a binary string obtained by bitwise XOR of α and β. An m-variable Boolean function f is a map $f : \{0,1\}^m \to \{0,1\}$. The support of f, denoted $\mathsf{supp}(f)$, is defined as follows.

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

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 $\mathsf{wt}(f) = 2^{m-1}$.

The imbalance of f will be denoted as $\mathsf{Imb}(f)$ and is defined as follows.

$$\mathsf{Imb}(f) = \frac{1}{2}\left(\#\{\alpha \in \{0,1\}^m : f(\alpha) = 0\} - \#\{\alpha \in \{0,1\}^m : f(\alpha) = 1\}\right) = 2^{m-1} - \mathsf{wt}(f).$$

Let $f, g : \{0,1\}^m \to \{0,1\}$ be two Boolean functions. By $f \oplus g$ we denote the Boolean function $h : \{0,1\}^m \to \{0,1\}$ 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

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

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

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

The function ϕ is a combiner of the input of S while the function ψ is a combiner of the output of S. There are no restrictions on ϕ and ψ and in particular, they are not required to be linear combiners. 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

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

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.

2.1 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 an 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 $\{0,1\}^n$.

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 $\{0,1\}^m$. 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*

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

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

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

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 2. *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.

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

□

Theorem 1. *Let ρ be a uniform random (m,n) function. Let ϕ and ψ be m and n-variable Boolean functions respectively. If ψ is a balanced Boolean function, then $\mathsf{wt}(f_\rho[\phi,\psi]) \sim \mathsf{Bin}(2^m, 1/2)$.*

Proof. From Proposition 2, $\mathsf{wt}(f_\rho[\phi,\psi]) = W = \sum_{i=0}^{2^m-1} W_i$ where $W_i \sim \mathsf{Ber}(p_i)$ with p_i given by (3). If ψ is a balanced Boolean function, then $p_i = 1/2$ and $W_i \sim \mathsf{Ber}(1/2)$. Let $\alpha_0, \ldots, \alpha_{2^m-1}$ be an enumeration of $\{0,1\}^m$ and $X_i = \rho(\alpha_i)$ as in Proposition 1. Note

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

Since the random variables $X_0, \ldots, X_{2^m-1}$ are independent, so are the random variables $W_0, \ldots, W_{2^m-1}$. As a result, W is a sum of 2^m independent random variables each of which follows $\mathsf{Ber}(1/2)$. So, $W \sim \mathsf{Bin}(2^m, 1/2)$. □

The special case of Theorem 1 where ϕ and ψ are non-trivial linear functions was proved in [4].

In the case where ψ is not a balanced function, 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, the W_i's are not identically distributed and hence W does not follow the binomial distribution. In this case, $W_0, \ldots, W_{2^m-1}$ is a sequence of 2^m Poisson trials. It is possible to use the Chernoff bound to get an estimate of the probability that W stays close to the mean.

Theorem 2. *Let ρ be a uniform random (m,n) function. Let ϕ and ψ be m and n-variable Boolean functions respectively. Then the expected value of $\mathsf{wt}(f_\rho[\phi,\psi])$ is*

$$\mu = \frac{2^m \mathsf{wt}(\psi) + 2^n \mathsf{wt}(\phi) - 2\mathsf{wt}(\phi)\mathsf{wt}(\psi)}{2^n}. \tag{4}$$

Further, for any $0 < \delta < 1$

$$\Pr\left[|\mathsf{wt}(f_\rho[\phi,\psi]) - \mu| \leq \delta\mu\right] \leq 2e^{-\mu\delta^2/3}. \tag{5}$$

Proof. Let W_i be as in Proposition 1 so that $\mathsf{wt}(f_\rho[\phi,\psi]) = \sum_{i=0}^{2^m-1} W_i$. From Proposition 1, $W_i \sim \mathsf{Ber}(p_i)$ and so the expected value of W_i is p_i. By linearity of expectation, the expected value of $\mathsf{wt}(f_\rho[\phi,\psi])$ equals

$$\begin{aligned}
\sum_{i=0}^{2^m-1} p_i &= \sum_{i=0}^{2^m-1} \frac{\mathsf{wt}(\psi) + \phi(\alpha_i)(2^n - 2\mathsf{wt}(\psi))}{2^n} \\
&= \frac{2^m \mathsf{wt}(\psi) + \mathsf{wt}(\phi)(2^n - 2\mathsf{wt}(\psi))}{2^n} \\
&= \frac{2^m \mathsf{wt}(\psi) + 2^n \mathsf{wt}(\phi) - 2\mathsf{wt}(\phi)\mathsf{wt}(\psi)}{2^n}.
\end{aligned}$$

As in the proof of Theorem 1, $W_0, \ldots, W_{2^m-1}$ are independent and since $W_i \sim \mathsf{Ber}(p_i)$, these random variables form a sequence of Poisson trials. The Chernoff bound applies (see Section A) leading to (5). □

2.2 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 3. *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

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

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

$$\begin{aligned} \#A_{0,0} + \#A_{1,0} &= 2^n - \mathsf{wt}(\psi), \\ \#A_{0,1} + \#A_{1,1} &= \mathsf{wt}(\psi). \end{aligned} \tag{7}$$

Equations (6) and (7) 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 3, 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

$$\{\mathsf{wt}(\phi) + \mathsf{wt}(\psi) - 2x : 0 \leq x \leq \min(\mathsf{wt}(\phi), \mathsf{wt}(\psi))\}. \tag{8}$$

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 (8).

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

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

If both ϕ and ψ are balanced functions, then

$$\Pr[\mathsf{wt}(f_\pi[\phi,\psi]) = \mathsf{wt}(\phi) + \mathsf{wt}(\psi) - 2x] = \frac{\binom{2^{n-1}}{x}^2}{\binom{2^n}{2^{n-1}}}. \tag{10}$$

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 3, 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 (9). From Proposition 3, it follows that $\mathsf{wt}(f_\pi[\phi, \psi]) = \mathsf{wt}(\phi) + \mathsf{wt}(\psi) - 2x$ if and only if $\#A_{1,1} = x$. This shows (9).

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 (9) and using $\binom{2^{n-1}}{2^{n-1}-x} = \binom{2^{n-1}}{x}$ yields (10). □

The expression given on the right hand side of (9) 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 (10) was proved in [4].

3 Nonlinear Invariant Attack

We provide a brief description of the nonlinear invariant attack for key alternating ciphers. Our description follows the suggestion in Section 7 of [10] where the nonlinear invariants are allowed to be different for the different rounds. Let $E_K : \{0,1\}^n \to \{0,1\}^n$ be a key alternating block cipher which iterates a round function $R : \{0,1\}^n \to \{0,1\}^n$ over r rounds. For an n-bit string L, define $R_L : \{0,1\}^n \to \{0,1\}^n$ as $R_L(\alpha) = R(\alpha \oplus L)$. For a plaintext P, let the ciphertext C be $C = E_K(P)$ which is obtained in the following manner. The secret key K is used to obtain the round keys $K_0, \ldots, K_{r-1}$. Then

$$C = (R_{K_{r-1}} \circ R_{K_{r-2}} \circ \cdots \circ R_{K_0})(P).$$

Suppose there are functions $g_0, \ldots, g_r : \{0,1\}^n \to \{0,1\}$ and constants $c_0, \ldots, c_{r-1} \in \{0,1\}$, such that there are round keys $K_0, \ldots, K_{r-1}$ for which

$$g_{i+1}(R(\alpha \oplus K_i)) = g_i(\alpha \oplus K_i) \oplus c_i = g_i(\alpha) \oplus g_i(K_i) \oplus c_i \tag{11}$$

for all $\alpha \in \{0,1\}^n$. Then $g_0, \ldots, g_r$ are called nonlinear invariants with associated constants $c_0, \ldots, c_{r-1}$. The round keys $K_0, \ldots, K_{r-1}$ are called weak keys.

The primary requirement in a key invariant attack is the property given in the following proposition. This property has been derived in [10] for the case where the functions $g_0, \ldots, g_r$ are all equal. The extension to possibly different $g_0, \ldots, g_r$ is quite straightforward and is given by the following result.

Proposition 4. *Let $E_K : \{0,1\}^n \to \{0,1\}^n$ be an r-round key alternating cipher. Suppose $g_0, \ldots, g_r$ are nonlinear invariants with associated constants $c_0, \ldots, c_{r-1}$ such that there are weak round keys $K_0, \ldots, K_{r-1}$ obtained from a key K. Then for any $\alpha \in \{0,1\}^n$,*

$$f_{E_K}[g_0, g_r](\alpha) \quad = \quad g_0(\alpha) \oplus g_r(E_K(\alpha)) \tag{12}$$

is a constant which is independent of α.

Proof. There are n-bit strings $\alpha_1, \ldots, \alpha_{r-1}$ such that $\alpha_0 = \alpha$; $\alpha_{i+1} = R_{K_i}(\alpha_i) = R(\alpha_i \oplus K_i)$ for $i = 0, \ldots, r-1$; and $\beta = \alpha_r = E_K(\alpha)$. The following holds.

$$\begin{aligned}
g_r(\beta) &= g_r(R(\alpha_{r-1} \oplus K_{r-1})) \\
&= g_{r-1}(\alpha_{r-1}) \oplus g_{r-1}(K_{r-1}) \oplus c_{r-1} \\
&= (g_{r-1}(R(\alpha_{r-2} \oplus K_{r-2}))) \oplus g_{r-1}(K_{r-1}) \oplus c_{r-1} \\
&= g_{r-2}(\alpha_{r-2}) \oplus (g_{r-2}(K_{r-2}) \oplus g_{r-1}(K_{r-1})) \oplus (c_{r-2} \oplus c_{r-1}) \\
&\vdots \\
&= g_0(P) \oplus \left(\bigoplus_{i=0}^{r-1} g_i(K_i) \right) \oplus \left(\bigoplus_{i=0}^{r-1} c_i \right).
\end{aligned}$$

So,

$$g_0(\alpha) \oplus g_r(\beta) \quad = \quad \left(\bigoplus_{i=0}^{r-1} g_i(K_i) \right) \oplus \left(\bigoplus_{i=0}^{r-1} c_i \right). \tag{13}$$

The right hand side of (13) is determined by the functions $g_0, \ldots, g_{r-1}$, the constants $c_0, \ldots, c_{r-1}$ and the round keys $K_0, \ldots, K_{r-1}$. In particular, it is independent of α. □

Proposition 4 shows that if $g_0, \ldots, g_r$ are nonlinear invariants for some weak keys $K_0, \ldots, K_{r-1}$, then for all 2^n n-bit strings α, $g_0(\alpha) \oplus g_r(E_K(\alpha))$ is a constant. We next consider the following question. Suppose g_0 and g_r are any two n-variable Boolean functions, $\alpha_1, \ldots, \alpha_N$ and $\beta_1, \ldots, \beta_N$ are arbitrary n-bit strings, what is the maximum value of N such that $g_0(\alpha_1) \oplus g_r(\beta_1) = \cdots = g_0(\alpha_N) \oplus g_r(\beta_N)$ holds?

Proposition 5. *Let $g_0, g_r : \{0,1\}^n \to \{0,1\}$. Let $\alpha_1, \ldots, \alpha_N$ and $\beta_1, \ldots, \beta_N$ be n-bit strings such that*

$$g_0(\alpha_1) \oplus g_r(\beta_1) = \cdots = g_0(\alpha_N) \oplus g_r(\beta_N).$$

Then $N \leq \mathfrak{N}$, where

$$\mathfrak{N} \quad = \quad \max\left(\min(2^n + w_0 - w_r, 2^n - w_0 + w_r), \min(w_0 + w_r, 2^{n+1} - w_0 - w_r)\right). \tag{14}$$

Here $w_0 = \mathsf{wt}(g_0)$ and $w_r = \mathsf{wt}(g_r)$. If $w_0 = w_r$, then the right hand side of (14) is equal to 2^n.

Proof. The condition $g_0(\alpha_1) \oplus g_r(\beta_1) = \cdots = g_0(\alpha_N) \oplus g_r(\beta_N)$ can occur in two ways, namely that all of the individual expressions are equal to 0 or, all of these are equal to 1.

Consider the maximum possible value of N such that $g_0(\alpha_1) \oplus g_r(\beta_1) = \cdots = g_0(\alpha_N) \oplus g_r(\beta_N) = 0$. An individual relation $g_0(\alpha_i) \oplus g_r(\beta_i)$ can be 0 in two possible ways, either $g_0(\alpha_i) = g_r(\beta_i) = 0$ or $g_0(\alpha_i) = g_r(\beta_i) = 1$. Suppose there are N_0 α_i's such that $g_0(\alpha_i) = g_r(\beta_i) = 0$ and there are N_1 α_i's such that $g_0(\alpha_i) = g_r(\beta_i) = 1$. Since $g_0(\alpha_i) = 1$ for N_1 i's, it follows that $N_1 \leq w_0$ and similarly, $N_1 \leq w_r$ so that $N_1 \leq \min(w_0, w_r)$. A similar argument shows that $N_0 \leq \min(2^n - w_0, 2^n - w_r)$. Since $N = N_0 + N_1$, we have $N \leq \min(w_0, w_r) + \min(2^n - w_0, 2^n - w_r)$.

Now consider the maximum possible value of N such that $g_0(\alpha_1) \oplus g_r(\beta_1) = \cdots = g_0(\alpha_N) \oplus g_r(\beta_N) = 1$. An argument similar to the above shows that $N \leq \min(2^n - w_0, w_r) + \min(w_0, 2^n - w_r)$.

The maximum value of N such that $g_0(\alpha_1) \oplus g_r(\beta_1) = \cdots = g_0(\alpha_N) \oplus g_r(\beta_N)$ is either the maximum value of N such that $g_0(\alpha_1) \oplus g_r(\beta_1) = \cdots = g_0(\alpha_N) \oplus g_r(\beta_N) = 0$ or the maximum value of N such that $g_0(\alpha_1) \oplus g_r(\beta_1) = \cdots = g_0(\alpha_N) \oplus g_r(\beta_N) = 1$. This shows that

$$N \leq \max\left(\min(w_0, w_r) + \min(2^n - w_0, 2^n - w_r), \min(w_0, 2^n - w_r) + \min(2^n - w_0, w_r)\right). \tag{15}$$

A simple argument shows that the right hand side of (15) is equal to the right hand side of (14). $\square$

Remark: Consider Propositions 4 and 5 together. If $g_0, \ldots, g_r$ are nonlinear invariants, then for all 2^n n-bit strings α, $g_0(\alpha) \oplus g_r(E_K(\alpha))$ is a constant. So, if $\mathfrak{N} < 2^n$, then there are no choices of Boolean functions $g_1, \ldots, g_{r-1}$, such that $g_0, g_1, \ldots, g_{r-1}, g_r$ are nonlinear invariants.

Notation: For the convenience of the ensuing description, we introduce some notation.

- For a Boolean function f and $\overline{\alpha} = (\alpha_1, \ldots, \alpha_N)$ where $\alpha_i \in \{0,1\}^n$ for $i = 1, \ldots, N$, define $\Psi(f, \overline{\alpha}) = (f(\alpha_1), \ldots, f(\alpha_N))$.
- For $0 \leq w \leq 2^n$, let $\mathcal{F}_w$ be the set of all n-variable Boolean functions having weight w.
- Given g_0, for $0 \leq \ell \leq N$, let $\mathcal{P}_\ell[g_0]$ be the set of all $\overline{\alpha} = (\alpha_1, \ldots, \alpha_N)$, $\alpha_i \in \{0,1\}^n$ such that $g_0(\alpha_i) = 1$ for exactly ℓ of the α_i's, i.e., $\mathcal{P}_\ell = \{\overline{\alpha} = (\alpha_1, \ldots, \alpha_N) : \#\{i : g_0(\alpha_i) = 1\} = \ell\}$. When g_0 is clear from the context we will simply write $\mathcal{P}_\ell$ instead of $\mathcal{P}_\ell[g_0]$.

Lemma 1. *Let $\overline{P} = (P_1, \ldots, P_N)$ where $P_1, \ldots, P_N$ are chosen from $\{0,1\}^n$ under uniform random sampling without replacement. Then*

$$\Pr[\overline{P} \in \mathcal{P}_\ell[g_0]] = \frac{\binom{w_0}{\ell}\binom{2^n - w_0}{N-\ell}}{\binom{2^n}{N}}, \tag{16}$$

where $w_0 = \mathsf{wt}(g_0)$.

Proof. The event $\overline{P} \in \mathcal{P}_\ell$ occurs if exactly ℓ of the P_i's fall in the support of g_0 while the other $N - \ell$ of the P_i's fall outside the support of g_0. Let us call strings in the support of g_0 to be red and the strings outside the support of g_0 to be black. So, there are w_0 red strings and $2^n - w_0$ black strings. The random experiment consists of choosing N distinct strings from 2^n strings such that ℓ are red and $N - \ell$ are black. This is the setting of hypergeometric distribution and the required probability is given by the right hand side of (16). $\square$

3.1 Building Distinguishers

Proposition 4 provides a structural property for a key alternating cipher E_K. Suppose $g_0, \ldots, g_r$ are nonlinear invariants (with associated constants $c_0, \ldots, c_{r-1}$) and K is such that $K_0, \ldots, K_{r-1}$ are weak keys, then for any plaintext P, $g_0(P) \oplus g_r(E_K(P))$ is a constant. To be able to distinguish E_K from a uniform random permutation π (resp. a uniform random function ρ), it is required to obtain the probability that $g_0(P) \oplus g_r(\pi(P))$ (resp. $g_0(P) \oplus g_r(\rho(P))$) is a constant.

The availability of a single plaintext is not sufficient to construct a meaningful distinguisher. So, suppose plaintexts $P_1, \ldots, P_N$ are used by the distinguishing algorithm. Since it is not useful to repeat plaintexts, without loss of generality, we may assume $P_1, \ldots, P_N$ to be distinct. From Proposition 4, we have that

$$f_{E_K}[g_0, g_r](P_1) = f_{E_K}[g_0, g_r](P_2) = \cdots = f_{E_K}[g_0, g_r](P_N). \tag{17}$$

Distinguishing from a uniform random permutation: Since a block cipher E_K is a bijective map, the appropriate goal would be to distinguish E_K from a uniform random permutation π of $\{0,1\}^n$. To build a distinguisher, it is required to know the probability of the following event.

$$\mathcal{E}^\pi : f_\pi[g_0, g_r](P_1) = f_\pi[g_0, g_r](P_2) = \cdots = f_\pi[g_0, g_r](P_N).$$

The event $\mathcal{E}^\pi$ can be written as the disjoint union of two events $\mathcal{E}_0^\pi$ and $\mathcal{E}_1^\pi$, i.e., $\mathcal{E}^\pi = \mathcal{E}_0^\pi \cup \mathcal{E}_1^\pi$, where

$$\begin{array}{ll} \mathcal{E}_0^\pi : & f_\pi[g_0, g_r](P_1) = 0, f_\pi[g_0, g_r](P_2) = 0, \ldots, f_\pi[g_0, g_r](P_N) = 0; \\ \mathcal{E}_1^\pi : & f_\pi[g_0, g_r](P_1) = 1, f_\pi[g_0, g_r](P_2) = 1, \ldots, f_\pi[g_0, g_r](P_N) = 1. \end{array} \tag{18}$$

So,

$$\Pr[\mathcal{E}^\pi] = \Pr[\mathcal{E}_0^\pi] + \Pr[\mathcal{E}_1^\pi]. \tag{19}$$

Suppose $\mathcal{D}^{\mathcal{O}}$ be a distinguisher which distinguishes E_K from π using a nonlinear invariant attack. On input $P_1, \ldots, P_N$, $\mathcal{D}^{\mathcal{O}}$ returns either real indicating that its oracle $\mathcal{O}$ is E_K; or it returns rnd indicating that its oracle is a uniform random permutation π. The distinguisher $\mathcal{D}^{\mathcal{O}}$ invokes $\mathcal{O}$ on inputs $P_1, \ldots, P_N$ obtaining in return $C_1 = \mathcal{O}(P_1), \ldots, C_N = \mathcal{O}(P_N)$. If $g_0(P_1) \oplus g_r(C_1) = \cdots = g_0(P_N) \oplus g_r(C_N)$, then $\mathcal{D}^{\mathcal{O}}$ returns real, else $\mathcal{D}^{\mathcal{O}}$ returns rnd.

If $\mathcal{O} = E_K$ for a weak key K, then $\mathcal{D}$ always returns real and hence makes no error. On the other hand, if $\mathcal{O} = \pi$, then the correct answer should be rnd, but, it is possible that $\mathcal{D}$ makes an error and returns real. So, the error that $\mathcal{D}$ can make is one-sided and the probability that $\mathcal{D}$ returns real when $\mathcal{O} = \pi$ is exactly $\Pr[\mathcal{E}^\pi]$.

Uniform random function: Considering a block cipher to be a map from n-bit strings to n-bit strings, a weaker goal would be to distinguish E_K from a uniform random function ρ from $\{0,1\}^n$ to $\{0,1\}^n$. The events $\mathcal{E}^\rho, \mathcal{E}_0^\rho$ and $\mathcal{E}_1^\rho$ are defined in a manner similar to $\mathcal{E}^\pi, \mathcal{E}_0^\pi$ and $\mathcal{E}_1^\pi$ respectively with π replaced by ρ. To build a distinguisher, it is required to obtain the probability of $\mathcal{E}^\rho$. As in the case of uniform random permutation, a distinguisher can make only one-sided error and the probability of this error is $\Pr[\mathcal{E}^\rho]$.

Choice of plaintexts: On being provided with distinct plaintexts $P_1, \ldots, P_N$, the distinguisher can make an error. The error probability depends on the manner in which $P_1, \ldots, P_N$ are chosen. We will analyse the error probability under the following two possible scenarios.

Uniform random sampling without replacement: In this analysis, we assume that $P_1, \ldots, P_N$ are chosen from $\{0,1\}^n$ using uniform random sampling without replacement.

Fixed values: In this analysis, it is assumed that $P_1, \ldots, P_N$ are fixed n-bit strings, i.e., there is no randomness in the plaintexts. Suppose $(P_1, \ldots, P_N) \in \mathcal{P}_\ell[g_0]$, i.e., there are exactly ℓ P_i's such that $g_0(P_i) = 1$. We show that the probability of error depends on ℓ.

We introduce the following notation to denote the four different kinds of error probabilities that can occur.

- $\varepsilon_{\pi,\$}$ is the error probability of distinguishing E_K from a uniform random permutation π when $P_1, \ldots, P_N$ are chosen under uniform random sampling without replacement, i.e., $\varepsilon_{\pi,\$} = \Pr[\mathcal{E}^\pi]$ when $P_1, \ldots, P_N$ are chosen under uniform random sampling without replacement.

- $\varepsilon_{\pi,\ell}$ is the error probability of distinguishing E_K from a uniform random permutation π when $(P_1, \ldots, P_N) \in \mathcal{P}_\ell$, i.e., $\varepsilon_{\pi,\ell} = \Pr[\mathcal{E}^\pi]$ when $(P_1, \ldots, P_N) \in \mathcal{P}_\ell$.

- $\varepsilon_{\rho,\$}$ is the error probability of distinguishing E_K from a uniform random function ρ when $P_1, \ldots, P_N$ are chosen under uniform random sampling without replacement, i.e., $\varepsilon_{\rho,\$} = \Pr[\mathcal{E}^\rho]$ when $P_1, \ldots, P_N$ are chosen under uniform random sampling without replacement.

- $\varepsilon_{\rho,\ell}$ is the error probability of distinguishing E_K from a uniform random function ρ when $(P_1,\ldots,P_N) \in \mathcal{P}_\ell$, i.e., $\varepsilon_{\rho,\ell} = \Pr[\mathcal{E}^\rho]$ when $(P_1,\ldots,P_N) \in \mathcal{P}_\ell$.

4 Error Probability for Uniform Random Function

In this section, we obtain expressions for $\varepsilon_{\rho,\$}$ and $\varepsilon_{\rho,\ell}$. The expression for $\varepsilon_{\rho,\$}$ is given in Theorem 5 with Lemma 2 leading up to it. Corollary 4 to Lemma 2 provides the expression for $\varepsilon_{\rho,\ell}$.

Lemma 2. *Let g_0 and g_r be two n-variable Boolean functions. Let ρ be a uniform random function and $F = f_\rho[g_0, g_r] = g_0 \oplus (g_r \circ \rho)$. Let $\overline{\alpha} = (\alpha_1,\ldots,\alpha_N)$ where $\alpha_1,\ldots,\alpha_N$ are distinct n-bit strings. Then*

$$\Pr[\Psi(F,\overline{\alpha}) = (0,\ldots,0)] = \prod_{i=1}^{N}\left(\frac{w_r}{2^n}g_0(\alpha_i) + \frac{2^n - w_r}{2^n}(1 - g_0(\alpha_i))\right) = \left(\frac{w_r}{2^n}\right)^{\ell}\left(\frac{2^n - w_r}{2^n}\right)^{N-\ell}; \tag{20}$$

$$\Pr[\Psi(F,\overline{\alpha}) = (1,\ldots,1)] = \prod_{i=1}^{N}\left(\frac{2^n - w_r}{2^n}g_0(\alpha_i) + \frac{w_r}{2^n}(1 - g_0(\alpha_i))\right) = \left(\frac{w_r}{2^n}\right)^{N-\ell}\left(\frac{2^n - w_r}{2^n}\right)^{\ell}. \tag{21}$$

where $w_r = \mathsf{wt}(g_r)$ and ℓ is such that $\overline{\alpha} \in \mathcal{P}_\ell$.

Further, if g_r is balanced, then $\Pr[\Psi(F,\overline{\alpha}) = (0,\ldots,0)] = \Pr[\Psi(F,\overline{\alpha}) = (1,\ldots,1)] = 1/2^N$.

Proof. Consider $\Psi(F,\overline{\alpha}) = (0,\ldots,0)$ which is the following event:

$$g_r(\rho(\alpha_1)) = g_0(\alpha_1), \ldots, g_r(\rho(\alpha_N)) = g_0(\alpha_N).$$

Since $\alpha_1,\ldots,\alpha_N$ are distinct and ρ is a uniform random function, the n-bit strings $X_1 = \rho(\alpha_1),\ldots,X_N = \rho(\alpha_N)$ are independent and uniformly distributed over $\{0,1\}^n$. Let $p_i = \Pr[g_r(\rho(\alpha_i)) = g_0(\alpha_i)] = \Pr[g_r(X_i) = g_0(\alpha_i)]$ for $i = 1,\ldots,N$. Since $X_1,\ldots,X_N$ are independent, so are the events $g_r(X_1) = g_0(\alpha_1),\ldots,g_r(X_N) = g_0(\alpha_N)$. Consequently,

$$\begin{aligned}\Pr[\Psi(F,\overline{\alpha}) = (0,\ldots,0)] &= \Pr[g_r(X_1) = g_0(\alpha_1),\ldots,g_r(X_N) = g_0(\alpha_N)] \\ &= p_1 \cdots p_N.\end{aligned}$$

Since X_i is uniformly distributed over $\{0,1\}^n$, the event $g_r(X_i) = 1$ occurs if and only if X_i falls within the support of g_r and the probability of this is $w_r/2^n$. Similarly, the event $g_r(X_i) = 0$ occurs with probability $(2^n - w_r)/2^n$.

$$p_i = \Pr[g_r(X_i) = g_0(\alpha_i)] = \begin{cases} \Pr[g_r(X_i) = 1] = w_r/2^n & \text{if } g_0(\alpha_i) = 1; \\ \Pr[g_r(X_i) = 0] = (2^n - w_r)/2^n & \text{if } g_0(\alpha_i) = 0. \end{cases}$$

This can be compactly written as

$$p_i = \frac{w_r}{2^n}g_0(\alpha_i) + \frac{2^n - w_r}{2^n}(1 - g_0(\alpha_i).$$

Let $\overline{\alpha} \in \mathcal{P}_\ell$. Then for exactly ℓ of the α_i's we have $g_0(\alpha_i) = 1$ while for the other $N - \ell$ of the α_i's, we have $g_0(\alpha_i) = 0$. This consideration leads to (20).

The proof for (21) is similar. If g_r is balanced, then $w_r = 2^{n-1}$ which shows the last part of the theorem. □

Corollary 4. *Let g_0 and g_r be two n-variable Boolean functions. Let ρ be a uniform random function and $F = f_\rho[g_0, g_r] = g_0 \oplus (g_r \circ \rho)$. Let $\overline{P} = (P_1,\ldots,P_N) \in \mathcal{P}_\ell$. Then*

$$\varepsilon_{\rho,\ell} = \Pr[\mathcal{E}^\rho] = \left(\frac{w_r}{2^n}\right)^{\ell}\left(\frac{2^n - w_r}{2^n}\right)^{N-\ell} + \left(\frac{w_r}{2^n}\right)^{N-\ell}\left(\frac{2^n - w_r}{2^n}\right)^{\ell}. \tag{22}$$

Further, if g_r is balanced, then $\varepsilon_{\rho,\ell} = 1/2^{N-1}$.

Theorem 5. *Let g_0 and g_r be two n-variable Boolean functions. Let ρ be a uniform random function from $\{0,1\}^n$ to $\{0,1\}^n$ and $F = f_\rho[g_0, g_r] = g_0 \oplus (g_r \circ \rho)$. Let $\overline{P} = (P_1, \ldots, P_N)$ where $P_1, \ldots, P_N$ are chosen from $\{0,1\}^n$ under uniform random sampling without replacement and these are independent of F. Then*

$$\begin{aligned}
\Pr[\mathcal{E}_0^\rho] = \Pr[\Psi(F,\overline{P}) = (0,\ldots,0)] &= \sum_{\ell=0}^{N} \left(\frac{w_r}{2^n}\right)^{\ell} \left(\frac{2^n - w_r}{2^n}\right)^{N-\ell} \cdot \frac{\binom{w_0}{\ell}\binom{2^n-w_0}{N-\ell}}{\binom{2^n}{N}}; \\
\Pr[\mathcal{E}_1^\rho] = \Pr[\Psi(F,\overline{P}) = (1,\ldots,1)] &= \sum_{\ell=0}^{N} \left(\frac{2^n - w_r}{2^n}\right)^{\ell} \left(\frac{w_r}{2^n}\right)^{N-\ell} \cdot \frac{\binom{w_0}{\ell}\binom{2^n-w_0}{N-\ell}}{\binom{2^n}{N}}.
\end{aligned} \tag{23}$$

Here $w_0 = \mathsf{wt}(g_0)$ and $w_r = \mathsf{wt}(g_r)$.

Consequently,

$$\varepsilon_{\rho,\$} \;=\; \Pr[\mathcal{E}^\rho] = \sum_{\ell=0}^{N} \left(\left(\frac{w_r}{2^n}\right)^{\ell} \left(\frac{2^n - w_r}{2^n}\right)^{N-\ell} + \left(\frac{2^n - w_r}{2^n}\right)^{\ell} \left(\frac{w_r}{2^n}\right)^{N-\ell} \right) \cdot \frac{\binom{w_0}{\ell}\binom{2^n-w_0}{N-\ell}}{\binom{2^n}{N}}. \tag{24}$$

Further, if g_r is balanced, then $\varepsilon_{\rho,\$} = 1/2^{N-1}$.

Proof. Consider the event $\mathcal{E}_0^\rho$.

$$\begin{aligned}
&\Pr[\mathcal{E}_0^\rho] = \Pr[\Psi(F,\overline{P}) = (0,\ldots,0)] \\
&= \sum_{\ell=0}^{N} \Pr[\Psi(F,\overline{P}) = (0,\ldots,0), \overline{P} \in \mathcal{P}_\ell] \\
&= \sum_{\ell=0}^{N} \sum_{\overline{\alpha} \in \mathcal{P}_\ell} \Pr[\Psi(F,\overline{P}) = (0,\ldots,0), \overline{P} = \overline{\alpha}] \\
&= \sum_{\ell=0}^{N} \sum_{\overline{\alpha} \in \mathcal{P}_\ell} \Pr[\Psi(F,\overline{\alpha}) = (0,\ldots,0), \overline{P} = \overline{\alpha}] \\
&= \sum_{\ell=0}^{N} \sum_{\overline{\alpha} \in \mathcal{P}_\ell} \Pr[\Psi(F,\overline{\alpha}) = (0,\ldots,0)] \cdot \Pr[\overline{P} = \overline{\alpha}] \quad \text{(since } F \text{ and } \overline{P} \text{ are independent)} \\
&= \sum_{\ell=0}^{N} \sum_{\overline{\alpha} \in \mathcal{P}_\ell} \left(\frac{w_r}{2^n}\right)^{\ell} \left(\frac{2^n - w_r}{2^n}\right)^{N-\ell} \cdot \Pr[\overline{P} = \overline{\alpha}] \quad \text{(from Lemma 2)} \\
&= \sum_{\ell=0}^{N} \left(\frac{w_r}{2^n}\right)^{\ell} \left(\frac{2^n - w_r}{2^n}\right)^{N-\ell} \sum_{\overline{\alpha} \in \mathcal{P}_\ell} \Pr[\overline{P} = \overline{\alpha}] \\
&= \sum_{\ell=0}^{N} \left(\frac{w_r}{2^n}\right)^{\ell} \left(\frac{2^n - w_r}{2^n}\right)^{N-\ell} \Pr[\overline{P} \in \mathcal{P}_\ell] \\
&= \sum_{\ell=0}^{N} \left(\frac{w_r}{2^n}\right)^{\ell} \left(\frac{2^n - w_r}{2^n}\right)^{N-\ell} \cdot \frac{\binom{w_0}{\ell}\binom{2^n-w_0}{N-\ell}}{\binom{2^n}{N}} \quad \text{(from Lemma 1).}
\end{aligned}$$

The probability of the event $\mathcal{E}_1^\rho$ is similarly obtained. Since $\mathcal{E}^\rho$ is the disjoint union of $\mathcal{E}_0^\rho$ and $\mathcal{E}_1^\rho$, we obtain (24).

If g_r is balanced, $w_r = 2^{n-1}$ and we have

$$\varepsilon_{\rho,\$} \;=\; \frac{1}{2^{N-1}} \sum_{\ell=0}^{N} \frac{\binom{w_0}{\ell}\binom{2^n-w_0}{N-\ell}}{\binom{2^n}{N}} = \frac{1}{2^{N-1}}.$$

The last equality holds since $\binom{w_0}{\ell}\binom{2^n-w_0}{N-\ell}/\binom{2^n}{N}$ is the probability that a random variable X equals ℓ where X follows $\mathsf{HG}(2^n, w_0, N)$ and so $\sum_{\ell=0}^{N}\Pr[X=\ell]=1$. □

Remarks:

1. From Corollay 4 and Theorem 5, we have that if g_r is balanced, then $\varepsilon_{\rho,\ell} = \varepsilon_{\rho,\$} = 1/2^{N-1}$, i.e., the error probability of the distinguisher is determined only by the number of distinct plaintexts that are used and not on whether these are fixed or chosen using uniform random sampling without replacement.
2. It has been mentioned in [10] that the distinguishing error of a nonlinear invariant attack is $1/2^{N-1}$. The above analysis shows that this is the error in distinguishing from a uniform random function.

5 Error Probability for Uniform Random Permutation

In this section, we obtain expressions for $\varepsilon_{\pi,\$}$ and $\varepsilon_{\pi,\ell}$. The expression for $\varepsilon_{\pi,\$}$ is given in Theorem 7. Lemmas 3 and 1 are intermediate steps to proving the theorem. Corollary 6 (to Lemma 3) provides the expression for $\varepsilon_{\pi,\ell}$. Using the results of Section 2.2, it is possible to obtain a different expression for $\varepsilon_{\pi,\$}$. This expression is derived in Appendix B.

Lemma 3. *Let g_0 and g_r be n-variable Boolean functions. Let π be a uniform random permutation and $F = f_\pi[g_0, g_r] = g_0 \oplus (g_r \circ \pi)$. Let $\alpha_1, \ldots, \alpha_N$ be distinct n-bit strings such that $\#\{i : g_0(\alpha_i) = 1\} = \ell$. Denote $\overline{\alpha} = (\alpha_1, \ldots, \alpha_N)$. Then*

$$\Pr[\Psi(F,\overline{\alpha}) = (0,\ldots,0)] = \frac{\binom{2^n-w_r}{N-\ell}\binom{w_r}{\ell}}{\binom{2^n}{N}\binom{N}{\ell}} \quad \textit{and} \ \Pr[\Psi(F,\overline{\alpha}) = (1,\ldots,1)] = \frac{\binom{w_r}{N-\ell}\binom{2^n-w_r}{\ell}}{\binom{2^n}{N}\binom{N}{\ell}}, \tag{25}$$

where $w_r = \mathsf{wt}(g_r)$.

Proof. Consider the first statement. It is given that $g_0(\alpha_i) = 1$ for exactly ℓ of the α_i's.

Let us start with the special case where $g_0(\alpha_1) = \cdots = g_0(\alpha_\ell) = 1$ and $g_0(\alpha_{\ell+1}) = \cdots = g_0(\alpha_N) = 0$. Then the event $\Psi(F,\overline{\alpha}) = (0,\ldots,0)$ holds if and only if $g_r(\pi(\alpha_1)) = \cdots = g_r(\pi(\alpha_\ell)) = 1$ and $g_r(\pi(\alpha_{\ell+1})) = \cdots = g_r(\pi(\alpha_N)) = 0$. Since $\alpha_1,\ldots,\alpha_N$ are distinct n-bit strings and π is a uniform random permutation of $\{0,1\}^n$, the random quantities $\pi(\alpha_1),\ldots,\pi(\alpha_N)$ can be thought of as being chosen from $\{0,1\}^n$ using uniform random sampling without replacement. Further, $g_r(\pi(\alpha_i)) = 1$ (resp. 0) if and only if $\pi(\alpha_i)$ falls within (resp. outside) the support of g_r.

From the above considerations, the probability that $g_r(\pi(\alpha_1)) = 1$ is $w_r/2^n$; the probability that $g_r(\pi(\alpha_2)) = 1$ given that $g_r(\pi(\alpha_1)) = 1$ is $(w_r-1)/(2^n-1)$; continuing, the probability that $g_r(\pi(\alpha_\ell)) = 1$ given that $g_r(\pi(\alpha_1)) = 1, \ldots, g_r(\pi(\alpha_{\ell-1})) = 1$ is $(w_r-\ell+1)/(2^n-\ell+1)$; the probability that $g_r(\pi(\alpha_{\ell+1})) = 0$ given that $g_r(\pi(\alpha_1)) = 1,\ldots,g_r(\pi(\alpha_\ell)) = 1$ is $(2^n-w_r)/(2^n-\ell)$; the probability that $g_r(\pi(\alpha_{\ell+2})) = 0$ given that $g_r(\pi(\alpha_1)) = 1,\ldots,g_r(\pi(\alpha_\ell)) = 1$ and $g_r(\pi(\alpha_{\ell+1})) = 0$ is $(2^n-w_r-1)/(2^n-\ell-1)$; continuing, the probability that $g_r(\pi(\alpha_N)) = 0$ given that $g_r(\pi(\alpha_1)) = 1,\ldots,g_r(\pi(\alpha_\ell)) = 1$ and $g_r(\pi(\alpha_{\ell+1})) = 0,\ldots,g_r(\pi(\alpha_{N-1})) = 0$ is $(2^n-w_r-(N-\ell)+1)/(2^n-N-1)$. So,

$$\begin{aligned}&\Pr[\Psi(F,\overline{\alpha}) = (0,\ldots,0)]\\ &= \frac{w_r(w_r-1)\cdots(w_r-\ell+1)(2^n-w_r)(2^n-w_r-1)\cdots(2^n-w_r-(N-\ell)-1)}{2^n(2^n-1)\cdots(2^n-N+1)}.\end{aligned} \tag{26}$$

Consider now the general case where there are exactly ℓ values of i such that $g_0(\alpha_i) = 1$ and these are not necessarily the first ℓ α_i's. Following the argument given above for the special case, it is not difficult to see that the probability of $\Psi(F,\overline{\alpha}) = (0,\ldots,0)$ in the general case is also given by (26). In particular, the argument shows that the numerator of the probability in the general case is a reordering of the numerator of (26) while the denominator remains the same. So, in all cases the probability of $\Psi(F,\overline{\alpha}) = (0,\ldots,0)$ is given by (26). Multiplying the numerator and denominator of (26) by $\ell!(N-\ell)!N!$ and some simplifications, we obtain

$$\Pr[\Psi(F,\overline{\alpha}) = (0,\ldots,0)] \;=\; \frac{\binom{2^n-w_r}{N-\ell}\binom{w_r}{\ell}}{\binom{2^n}{N}\binom{N}{\ell}}.$$

This shows the first statement. The other statement is obtained similarly. □

Corollary 6. *Let g_0 and g_r be two n-variable Boolean functions. Let π be a uniform random permutation and $F = f_\pi[g_0, g_r] = g_0 \oplus (g_r \circ \pi)$. Let $\overline{P} = (P_1,\ldots,P_N) \in \mathcal{P}_\ell$. Then*

$$\varepsilon_{\pi,\ell} \;=\; \Pr[\mathcal{E}^\pi] = \frac{\binom{2^n-w_r}{N-\ell}\binom{w_r}{\ell}}{\binom{2^n}{N}\binom{N}{\ell}} + \frac{\binom{w_r}{N-\ell}\binom{2^n-w_r}{\ell}}{\binom{2^n}{N}\binom{N}{\ell}}, \tag{27}$$

where $w_r = \mathsf{wt}(g_r)$.

Theorem 7. *Let g_0 and g_r be two n-variable Boolean functions. Let π be a uniform random permutation of $\{0,1\}^n$ and $F = f_\pi[g_0, g_r] = g_0 \oplus (g_r \circ \pi)$. Let $\overline{P} = (P_1,\ldots,P_N)$ where $P_1,\ldots,P_N$ are chosen from $\{0,1\}^n$ under uniform random sampling without replacement and these are independent of F. Then*

$$\begin{aligned}
\Pr[\mathcal{E}_0^\pi] = \Pr[\Psi(F,\overline{P}) = (0,\ldots,0)] &= \sum_{\ell=0}^{N} \frac{\binom{2^n-w_r}{N-\ell}\binom{w_r}{\ell}}{\binom{2^n}{N}\binom{N}{\ell}} \cdot \frac{\binom{w_0}{\ell}\binom{2^n-w_0}{N-\ell}}{\binom{2^n}{N}}; \\
\Pr[\mathcal{E}_1^\pi] = \Pr[\Psi(F,\overline{P}) = (1,\ldots,1)] &= \sum_{\ell=0}^{N} \frac{\binom{w_r}{N-\ell}\binom{2^n-w_r}{\ell}}{\binom{2^n}{N}\binom{N}{\ell}} \cdot \frac{\binom{w_0}{\ell}\binom{2^n-w_0}{N-\ell}}{\binom{2^n}{N}}.
\end{aligned} \tag{28}$$

Here $w_0 = \mathsf{wt}(g_0)$ and $w_r = \mathsf{wt}(g_r)$. Consequently,

$$\varepsilon_{\pi,\$} \;=\; \Pr[\mathcal{E}^\pi] = \sum_{\ell=0}^{N} \frac{\binom{2^n-w_r}{N-\ell}\binom{w_r}{\ell} + \binom{w_r}{N-\ell}\binom{2^n-w_r}{\ell}}{\binom{2^n}{N}\binom{N}{\ell}} \cdot \frac{\binom{w_0}{\ell}\binom{2^n-w_0}{N-\ell}}{\binom{2^n}{N}}. \tag{29}$$

If both g_0 and g_r are balanced, then $\varepsilon_{\pi,\$}$ is the expectation of $2p(X)/\binom{N}{X}$, i.e.,

$$\varepsilon_{\pi,\$} \;=\; \mathbf{E}\left[\frac{2p(X)}{\binom{N}{X}}\right] \tag{30}$$

where X follows $\mathsf{HG}(2^n, 2^{n-1}, N)$ and for $\ell = 0,\ldots,N$, $p(\ell)$ is the probability that $X = \ell$.

Proof. Consider $\Pr[\mathcal{E}_0^\pi]$.

$$
\begin{aligned}
&\Pr[\Psi(F,\overline{P}) = (0,\ldots,0)] \\
&= \sum_{\ell=0}^{N} \sum_{\overline{\alpha}\in\mathcal{P}_\ell} \Pr[\Psi(F,\overline{P}) = (0,\ldots,0), \overline{P} = \overline{\alpha}] \\
&= \sum_{\ell=0}^{N} \sum_{\overline{\alpha}\in\mathcal{P}_\ell} \Pr[\Psi(F,\overline{\alpha}) = (0,\ldots,0), \overline{P} = \overline{\alpha}] \\
&= \sum_{\ell=0}^{N} \sum_{\overline{\alpha}\in\mathcal{P}_\ell} \Pr[\Psi(F,\overline{\alpha}) = (0,\ldots,0)] \cdot \Pr[\overline{P} = \overline{\alpha}] \quad \text{(since } F \text{ and } \overline{P} \text{ are independent)} \\
&= \sum_{\ell=0}^{N} \sum_{\overline{\alpha}\in\mathcal{P}_\ell} \frac{\binom{2^n-w_r}{N-\ell}\binom{w_r}{\ell}}{\binom{2^n}{N}\binom{N}{\ell}} \cdot \Pr[\overline{P} = \overline{\alpha}] \quad \text{(from Lemma 3)} \\
&= \sum_{\ell=0}^{N} \frac{\binom{2^n-w_r}{N-\ell}\binom{w_r}{\ell}}{\binom{2^n}{N}\binom{N}{\ell}} \sum_{\overline{\alpha}\in\mathcal{P}_\ell} \Pr[\overline{P} = \overline{\alpha}] \\
&= \sum_{\ell=0}^{N} \frac{\binom{2^n-w_r}{N-\ell}\binom{w_r}{\ell}}{\binom{2^n}{N}\binom{N}{\ell}} \Pr[\overline{P} \in \mathcal{P}_\ell] \\
&= \sum_{\ell=0}^{N} \frac{\binom{2^n-w_r}{N-\ell}\binom{w_r}{\ell}}{\binom{2^n}{N}\binom{N}{\ell}} \cdot \frac{\binom{w_0}{\ell}\binom{2^n-w_0}{N-\ell}}{\binom{2^n}{N}} \quad \text{(from Lemma 1).}
\end{aligned}
$$

$\Pr[\mathcal{E}_1^\pi]$ is obtained similarly. Further, the probability of $\mathcal{E}^\pi$ is obtained from (19).

If both g_0 and g_r are balanced, then $w_0 = w_r = 2^{n-1}$ and we have

$$
\varepsilon_{\pi,\$} = \sum_{\ell=0}^{N} \frac{2\binom{2^{n-1}}{N-\ell}\binom{2^{n-1}}{\ell}}{\binom{2^n}{N}\binom{N}{\ell}} \cdot \frac{\binom{2^{n-1}}{\ell}\binom{2^{n-1}}{N-\ell}}{\binom{2^n}{N}} = \sum_{\ell=0}^{N} \frac{2p(\ell)}{\binom{N}{\ell}} \cdot \frac{\binom{2^{n-1}}{\ell}\binom{2^{n-1}}{N-\ell}}{\binom{2^n}{N}} = \mathbf{E}\left[\frac{2p(X)}{\binom{N}{X}}\right].
$$

□

The next result shows that when g_0 and g_r are balanced, the distinguishing error for uniform random permutations is at least as large as that for uniform random functions.

Theorem 8. *Let g_0 and g_r be two balanced n-variable Boolean functions. Let π be a uniform random permutation of $\{0,1\}^n$ and ρ be a uniform random function from $\{0,1\}^n$ to $\{0,1\}^n$. Define $F_\pi = f_\pi[g_0, g_r] = g_0 \oplus (g_r \circ \pi)$ and $F_\rho = f_\rho[g_0, g_r] = g_0 \oplus (g_r \circ \rho)$. Let $\overline{P} = (P_1,\ldots,P_N)$ where $P_1,\ldots,P_N$ are chosen from $\{0,1\}^n$ under uniform random sampling without replacement and these are independent of F_ρ or F_π. Let*

$$
\begin{aligned}
\varepsilon_{\pi,\$} = \Pr[\mathcal{E}^\pi] &= \Pr[\mathcal{E}^\pi] + \Pr[\mathcal{E}_0^\pi] = \Pr[\Psi(F_\pi,\overline{P}) = (0,\ldots,0)] + \Pr[\Psi(F_\pi,\overline{P}) = (1,\ldots,1)]; \\
\varepsilon_{\rho,\$} = \Pr[\mathcal{E}^\rho] &= \Pr[\mathcal{E}^\rho] + \Pr[\mathcal{E}_0^\rho] = \Pr[\Psi(F_\rho,\overline{P}) = (0,\ldots,0)] + \Pr[\Psi(F_\rho,\overline{P}) = (1,\ldots,1)].
\end{aligned}
$$

Then $\varepsilon_{\pi,\$} \geq \varepsilon_{\rho,\$}$.

Proof. It is given that g_0 and g_r are both balanced. From Theorem 5, it follows that $\varepsilon_{\rho,\$} = 1/2^{N-1}$. From Theorem 7, we have that $\varepsilon_{\pi,\$}$ is the expectation of $2p(X)/\binom{N}{X}$, i.e., $\varepsilon_{\pi,\$} = \mathbf{E}[2p(X)/\binom{N}{X}]$, where X follows $\mathsf{HG}(2^n, 2^{n-1}, N)$ and for $\ell = 0,\ldots,N$, $p(\ell)$ is the probability that $X = \ell$.

Let $Y = 2p(X)/\binom{N}{X}$. Using Jensen's inequality, we obtain

$$\begin{aligned}
\frac{1}{\mathbf{E}[Y]} &\leq \mathbf{E}\left[\frac{1}{Y}\right] \\
&= \mathbf{E}\left[\frac{\binom{N}{X}}{2p(X)}\right] \\
&= \sum_{\ell=0}^{N} \frac{\binom{N}{\ell}}{2p(\ell)} \cdot \Pr[X = \ell] \\
&= \sum_{\ell=0}^{N} \frac{\binom{N}{\ell}}{2p(\ell)} \cdot p(\ell) \\
&= 2^{N-1}.
\end{aligned}$$

Noting $\varepsilon_{\pi,\$} = \mathbf{E}[Y]$ and $\varepsilon_{\rho,\$} = 1/2^{N-1}$ gives the desired result. □

6 Computational Results

This section gives a summary of the computations done with the expressions of the error probabilities of nonlinear invariant attack presented in Section 4 and 5. For computing $\varepsilon_{\pi,\$}$ which is the error probability for distinguishing from a uniform random permutation, we have used the expression given by (29).

In our computations we have used the following Stirling's approximation to compute the binomial coefficients.

$$\binom{\mathfrak{k}}{i} \approx \frac{1}{\sqrt{2\pi\mathfrak{k}}(i/\mathfrak{k})^{i+\frac{1}{2}}(1-i/\mathfrak{k})^{\mathfrak{k}-i+\frac{1}{2}}}.$$

The computations were done for $n = 16, 32, 48$ and 64; and $N = 2^{\mathfrak{n}}$ for $\mathfrak{n} = 2, 4, 8$ and 16, except that the case $N = 2^{16}$ was not considered when $n = 16$. Further, we have considered balanced g_0 and g_r, i.e., $\mathsf{wt}(g_0) = \mathsf{wt}(g_r) = 2^{n-1}$. As a result, $\varepsilon_{\rho,\$}$, which is the error probability of distinguishing from a uniform random function, is equal to $1/2^{N-1}$.

Comparison between $\varepsilon_{\pi,\$}$ and $\varepsilon_{\rho,\$}$. Table 1 gives the value of $\varepsilon_{\pi,\$}$ and the ratio $\varepsilon_{\pi,\$}/\varepsilon_{\rho,\$} = 2^{N-1}\varepsilon_{\pi,\$}$ for different values of n and $\mathfrak{n}$. It may be noted that the last column of the table confirms Theorem 8 which shows that for balanced g_0 and g_r, $\varepsilon_{\pi,\$} \geq \varepsilon_{\rho,\$} = 1/2^{N-1}$. Further, the ratio is close to 1. This may be explained by referring to the proof of Theorem 8. The result $\varepsilon_{\pi,\$} \geq 1/2^{N-1}$ is obtained using Jensen's inequality to the convex function $f(x) = 1/x$. It is known that Jensen's inequality is tight when the convex function is a straight line. In the range of x where Jensen's inequality is applied, it turns out that $f(x)$ behaves almost like a straight line. Consequently, the inequality is almost tight in this range of applicability.

7 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 Daeman and Rijmen [4] who had considered only linear combiners. Correlation between nonlinear input and output combiners arise in the context of nonlinear invariant attacks. We have performed a detailed analysis of the distinguishing error of such attacks.

n	$\mathfrak{n}$	$\varepsilon_{\pi,\$}$	$2^{N-1} \times \varepsilon_{\pi,\$}$
16	2	0.133739	1.069910
	4	0.000031	1.017414
	8	1.728943×10^{-77}	1.000990
32	2	0.133739	1.069908
	4	0.000031	1.017415
	8	1.728930×10^{-77}	1.000982
	16	$9.982420\times10^{-19729}$	1.000004
48	2	0.133739	1.069908
	4	0.000031	1.017415
	8	1.728930×10^{-77}	1.000982
	16	$9.982420\times10^{-19729}$	1.000004
64	2	0.133739	1.069908
	4	0.000031	1.017415
	8	1.728930×10^{-77}	1.000982
	16	$9.982420\times10^{-19729}$	1.000004

Table 1: Comparison between $\varepsilon_{\pi,\$}$ and $\varepsilon_{\rho,\$} = 2^{-(N-1)}$.

Acknowledgement

An earlier version of this paper was entitled "Correlation Between (Nonlinear) Combiners of Input and Output of Random Functions and Permutations and Analysis of Nonlinear Invariant Attacks" and consisted only of the material in Section 2. A reviewer of the earlier version had suggested applying the techniques to the study of nonlinear invariant attack. We have been successful in doing so and are grateful to the reviewer for having provided this suggestion.

References

[1] Tomer Ashur, Tim Beyne, and Vincent Rijmen. Revisiting the wrong-key-randomization hypothesis. *IACR Cryptology ePrint Archive*, 2016:990, 2016.

[2] Céline Blondeau and Kaisa Nyberg. Joint data and key distribution of simple, multiple, and multidimensional linear cryptanalysis test statistic and its impact to data complexity. *Des. Codes Cryptography*, 82(1-2):319–349, 2017.

[3] Andrey Bogdanov and Elmar Tischhauser. On the Wrong Key Randomisation and Key Equivalence Hypotheses in Matsui's Algorithm 2. In *Fast Software Encryption*, pages 19–38. Springer, 2014.

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

[5] Carlo Harpes, Gerhard G. Kramer, and James L. Massey. A generalization of linear cryptanalysis and the applicability of matsui's piling-up lemma. In Louis C. Guillou and Jean-Jacques Quisquater, editors,

Advances in Cryptology - EUROCRYPT '95, International Conference on the Theory and Application of Cryptographic Techniques, Saint-Malo, France, May 21-25, 1995, Proceeding, volume 921 of *Lecture Notes in Computer Science*, pages 24–38. Springer, 1995.

[6] Lars R. Knudsen and Matthew J. B. Robshaw. Non-linear approximations in linear cryptanalysis. In Ueli M. Maurer, editor, *Advances in Cryptology - EUROCRYPT '96, International Conference on the Theory and Application of Cryptographic Techniques, Saragossa, Spain, May 12-16, 1996, Proceeding*, volume 1070 of *Lecture Notes in Computer Science*, pages 224–236. Springer, 1996.

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

[8] Rajeev Motwani and Prabhakar Raghavan. *Randomized Algorithms.* Chapman & Hall/CRC, 2010.

[9] 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.

[10] Yosuke Todo, Gregor Leander, and Yu Sasaki. Nonlinear invariant attack: Practical attack on full SCREAM, iSCREAM and Midori64. *Journal of Cryptology*, 2018. `https://doi.org/10.1007/s00145-018-9285-0`.

A Chernoff Bound

We briefly recall the Chernoff bound. This result can be found in standard texts [8].

Theorem 9. *Let $X_1, X_2, \ldots, X_\lambda$ be a sequence of independent Poisson trials such that for $1 \le i \le \lambda$, $\Pr[X_i = 1] = p_i$. Then for $X = \sum_{i=1}^{\lambda} X_i$ and $\mu = E[X] = \sum_{i=1}^{\lambda} p_i$ the following bounds hold:*

$$\text{For any } 0 < \delta \le 1,\ \Pr[X \ge (1+\delta)\mu] \le e^{-\mu\delta^2/3}. \tag{31}$$

$$\text{For any } 0 < \delta < 1,\ \Pr[X \le (1-\delta)\mu] \le e^{-\mu\delta^2/2}. \tag{32}$$

B Alternative Expression for $\varepsilon_{\pi,\$}$

Lemma 4. *Let $\overline{P} = (P_1, \ldots, P_N)$ where $P_1, \ldots, P_N$ are chosen from $\{0,1\}^n$ under uniform random sampling without replacement. Let $f : \{0,1\}^n \to \{0,1\}$ be of weight w. Then*

$$\Pr[\Psi(f,\overline{P}) = (0,\ldots,0)] = \frac{\binom{2^n-w}{N}}{\binom{2^n}{N}} \quad \text{and} \quad \Pr[\Psi(f,\overline{P}) = (1,\ldots,1)] = \frac{\binom{w}{N}}{\binom{2^n}{N}}. \tag{33}$$

Proof. Consider the first statement. We need to consider $f(P_1) = 0, \ldots, f(P_N) = 0$. This holds if and only if all of $P_1, \ldots, P_N$ fall outside the support of f. The probability that P_1 falls outside the support of f is $(2^n - w)/2^n$; given that P_1 falls outside the support of f, the probability that P_2 falls outside the support of f is $(2^n - w - 1)/(2^n - 1)$; given that P_1, P_2 falls outside the support of f, the probability that P_3 falls outside the support of f is $(2^n - w - 2)/(2^n - 2)$ and so on. As a result we obtain

$$\begin{aligned}
\Pr[\Psi(f,\overline{P}) = (0,\ldots,0)] &= \frac{2^n-w}{2^n} \cdot \frac{2^n-w-1}{2^n-1} \cdot \frac{2^n-w-2}{2^n-2} \cdots \frac{2^n-w-N+1}{2^n-N+1} \\
&= \frac{(2^n-w)(2^n-w-1)\cdots(2^n-w-N+1))}{2^n(2^n-1)\cdots(2^n-N+1)} \cdot \frac{(2^n-N)!}{(2^n-N)!} \cdot \frac{(2^n-w-N)!}{(2^n-w-N)!} \\
&= \frac{\binom{2^n-w}{N}}{\binom{2^n}{N}}.
\end{aligned}$$

The other statement is obtained similarly. $\square$

Lemma 5. *Let F be a random (but, not necessarily uniform random) Boolean function. Let $\overline{P} = (P_1, \ldots, P_N)$ where $P_1, \ldots, P_N$ are chosen from $\{0,1\}^n$ under uniform random sampling without replacement and these are independent of F. Then*

$$\begin{aligned}
\Pr[\Psi(F,\overline{P}) = (0,\ldots,0)] &= \sum_{w=0}^{2^n} \frac{\binom{2^n-w}{N}}{\binom{2^n}{N}} \cdot \Pr[F \in \mathcal{F}_w]; \\
\Pr[\Psi(F,\overline{P}) = (1,\ldots,1)] &= \sum_{w=0}^{2^n} \frac{\binom{w}{N}}{\binom{2^n}{N}} \cdot \Pr[F \in \mathcal{F}_w].
\end{aligned} \tag{34}$$

Proof. Consider the first statement.

$$\begin{aligned}
&\Pr[\Psi(F,\overline{P}) = (0,\ldots,0)] \\
&= \sum_{w=0}^{2^n} \sum_{f\in\mathcal{F}_w} \Pr[\Psi(F,\overline{P}) = (0,\ldots,0) \wedge F = f] \\
&= \sum_{w=0}^{2^n} \sum_{f\in\mathcal{F}_w} \Pr[\Psi(f,\overline{P}) = (0,\ldots,0) \wedge F = f] \\
&= \sum_{w=0}^{2^n} \sum_{f\in\mathcal{F}_w} \Pr[\Psi(f,\overline{P}) = (0,\ldots,0)] \Pr[F = f] \quad \text{(since F and $\overline{P}$ are independent)} \\
&= \sum_{w=0}^{2^n} \sum_{f\in\mathcal{F}_w} \frac{\binom{2^n-w}{N}}{\binom{2^n}{N}} \cdot \Pr[F = f] \quad \text{(from Lemma 4)} \\
&= \sum_{w=0}^{2^n} \frac{\binom{2^n-w}{N}}{\binom{2^n}{N}} \sum_{f\in\mathcal{F}_w} \Pr[F = f] \\
&= \sum_{w=0}^{2^n} \frac{\binom{2^n-w}{N}}{\binom{2^n}{N}} \cdot \Pr[F \in \mathcal{F}_w].
\end{aligned}$$

The other statement is obtained similarly. $\square$

Theorem 10. *Let g_0 and g_r be two n-variable Boolean functions. Let π be a uniform random permutation of $\{0,1\}^n$ and $F = f_\pi[g_0, g_r] = g_0 \oplus (g_r \circ \pi)$. Let $\overline{P} = (P_1, \ldots, P_N)$ where $P_1, \ldots, P_N$ are chosen from $\{0,1\}^n$ under uniform random sampling without replacement and these are independent of F. Then*

$$\begin{aligned}
\Pr[\mathcal{E}_0^\pi] = \Pr[\Psi(F,\overline{P}) = (0,\ldots,0)] &= \sum_{x=0}^{\mathfrak{m}} \frac{\binom{2^n-w_0-w_r+2x}{N}}{\binom{2^n}{N}} \cdot \frac{\binom{w_0}{x}\binom{2^n-w_0}{w_r-x}}{\binom{2^n}{w_r}}; \\
\Pr[\mathcal{E}_1^\pi] = \Pr[\Psi(F,\overline{P}) = (1,\ldots,1)] &= \sum_{x=0}^{\mathfrak{m}} \frac{\binom{w_0+w_r-2x}{N}}{\binom{2^n}{N}} \cdot \frac{\binom{w_0}{x}\binom{2^n-w_0}{w_r-x}}{\binom{2^n}{w_r}}.
\end{aligned} \tag{35}$$

Here $w_0 = \mathsf{wt}(g_0)$, $w_r = \mathsf{wt}(g_r)$ and $\mathfrak{m} = \min(w_0, w_r)$. Consequently,

$$\varepsilon_{\pi,\$} = \Pr[\mathcal{E}^\pi] = \sum_{x=0}^{\mathfrak{m}} \frac{\binom{2^n-w_0-w_r+2x}{N} + \binom{w_0+w_r-2x}{N}}{\binom{2^n}{N}} \cdot \frac{\binom{w_0}{x}\binom{2^n-w_0}{w_r-x}}{\binom{2^n}{w_r}}. \tag{36}$$

Proof. From Theorem 3, the possible values of the weight of F are $w_0 + w_r - 2x$ for $x = 0, \ldots, \mathfrak{m}$ and for $w = w_0 + w_r - 2x$, $\Pr[F \in \mathcal{F}_w] = \binom{w_0}{x}\binom{2^n - w_0}{w_r - x}/\binom{2^n}{w_r}$.

Consider $\Pr[\mathcal{E}_0^\pi]$. From Lemma 5,

$$\begin{aligned}\Pr[\mathcal{E}_0^\pi] = \Pr[\Psi(F, \overline{P}) = (0, \ldots, 0)] &= \sum_{w=0}^{2^n} \frac{\binom{2^n - w}{N}}{\binom{2^n}{N}} \cdot \frac{\binom{w_0}{x}\binom{2^n - w_0}{w_r - x}}{\binom{2^n}{w_r}} \\ &= \sum_{x=0}^{\mathfrak{m}} \frac{\binom{2^n - w_0 - w_r + 2x}{N}}{\binom{2^n}{N}} \cdot \frac{\binom{w_0}{x}\binom{2^n - w_0}{w_r - x}}{\binom{2^n}{w_r}}.\end{aligned}$$

The other statement is obtained similarly. Further, the probability of $\mathcal{E}^\pi$ is obtained from (19). $\square$