IACR Transactions on Symmetric Cryptology
ISSN 2519-173X, Vol. 2021, No. 4, pp. 124–137. DOI:10.46586/tosc.v2021.i4.124-137

Statistical Model of Correlation Difference and Related-Key Linear Cryptanalysis

Kaisa Nyberg

Aalto University School of Science, Espoo, Finland
kaisa.nyberg@aalto.fi

Abstract. The goal of this work is to propose a related-key model for linear cryptanalysis. We start by giving the mean and variance of the difference of sampled correlations of two Boolean functions when using the same sample of inputs to compute both correlations. This result is further extended to determine the mean and variance of the difference of correlations of a pair of Boolean functions taken over a random data sample of fixed size and over a random pair of Boolean functions. We use the properties of the multinomial distribution to achieve these results without independence assumptions. Using multivariate normal approximation of the multinomial distribution we obtain that the distribution of the difference of related-key correlations is approximately normal. This result is then applied to existing related-key cryptanalyses. We obtain more accurate right-key and wrong-key distributions and remove artificial assumptions about independence of sampled correlations. We extend this study to using multiple linear approximations and propose a χ^2-type statistic, which is proven to be χ^2 distributed if the linear approximations are independent. We further examine this statistic for multidimensional linear approximation and discuss why removing the assumption about independence of linear approximations does not work in the related-key setting the same way as in the single-key setting.

Keywords: block cipher · linear cryptanalysis · related-key attack · statistical model.

1 Introduction

Linear cryptanalysis is one of the main standard statistical methods for analysing the strength of a symmetric-key block cipher. It is mostly used in the single-key setting. Applications to related-key setting are much more rare in comparison, for example, with differential cryptanalysis. The only works so far seem to be [RN13] and [BBR+13]. They consider difference of correlations under a fixed difference in related key pairs.

Linear cryptanalysis exploits biased linear expressions computed from cipher data called as linear approximations. Given a sample of plaintext-ciphertext pairs, the cryptanalyst computes the sampled correlation of the linear approximation. Statistical modelling of sampled correlations are needed to determine the data requirements of the attack.

Statistical distributions of sampled correlations of linear approximations of block ciphers are well established in the single-key setting, see e.g. [BN17]. The goal of this paper is to derive statistical distributions of the difference of the sampled correlations of Boolean functions. Such differences of correlations emerge in related-key linear cryptanalysis when the correlation of a linear approximation of a block cipher is analysed for two different keys. In previous works mentioned above, the distributions are modelled under the assumption that the sampled correlations computed for two different keys are statistically independent. Considering the fact that the related-key cryptanalysis exploits some nonrandom behaviour of a block cipher that becomes observable when analysing data obtained from the cipher with two different keys, the assumption about statistical independence is somewhat contradictory.

Received: 2021-09-01 Accepted: 2021-11-01 Published: 2021-12-03

Another approach to argue for independency has been to use two independent data samples to compute the correlations [RN13]. In this paper we establish the distribution of the correlation difference where the two correlations are computed for the same set of known plaintexts.

Our contributions. The main technical result of this paper gives the mean and the variance of the difference of sampled correlations of a given pair of Boolean functions when using the same sample of inputs to compute both correlations. This result is further extended to determine the mean and variance of the difference of correlations of a pair of Boolean functions taken over a random data sample of fixed size and over a random pair of Boolean functions. We use properties of multinomial distribution to achieve these results without independence assumptions on the pair of functions. Using multivariate normal approximation of the multinomial distribution we obtain that the distribution of the correlation difference is approximately normal.

We then discuss the impact of this result on the existing works on related-key cryptanalysis [RN13] and [BBR+13]. We propose a statistic for analysing the difference of sampled correlations of a set of linear approximations applied to the cipher with two different keys. In particular, we establish the distributions of this statistic for KDIB cryptanalysis for both right and wrong keys without assuming independence of the sampled correlations computed for related keys.

While the new wrong-key model is essentially the same as the one derived under the independence assumption, the right-key model is more detailed and may potentially lead to improvements in practical applications.

2 Sampling of the Difference of Correlations

Let f and f' be two Boolean functions from $\mathbb{F}_2^n$ to $\mathbb{F}_2$ with correlations $\mathrm{cor}(f) = 2p - 1$ and $\mathrm{cor}(f') = 2p' - 1$, respectively, where we denoted by p and p' the probabilities that f and f' take the value 0.

A random sample $S \subset \mathbb{F}_2^n$ of size N is composed of triplets $(x, f(x), f'(x))$ where each $x \in S$ is picked equiprobably from $\mathbb{F}_2^n$. The sampling is done either with or without replacement. The difference between these sampling methods in our applications is well captured by using the finite population correction factor given by

$$\frac{2^n - N}{2^n - 1},$$

see e.g. [RT89]. The variances of the binomial and hypergeometric distributions differ by this factor. When considering these two alternative sampling methods in parallel we will use the following constant

$$B = \begin{cases} \frac{2^n - N}{2^n - 1}, & \text{if the sample is drawn without replacement, or} \\ 1, & \text{if the sample is drawn with replacement.} \end{cases} \tag{1}$$

Let $S \subset \mathbb{F}_2^n$ be a sample of inputs to f and f' of size N. We denote by

$$X_{\gamma,\delta} = |\{x \in S \,|\, f(x) = \gamma,\ f'(x) = \delta\}|, \text{ where } \gamma, \delta \in \{0, 1\},$$

the value distribution of the pairs $(f(x), f'(x))$, $x \in S$, and by

$$\begin{aligned} \widehat{\mathrm{cor}}(f) &= \frac{1}{N}(X_{0,0} + X_{0,1} - X_{1,0} - X_{1,1}) \text{ and} \\ \widehat{\mathrm{cor}}(f') &= \frac{1}{N}(X_{0,0} + X_{1,0} - X_{0,1} - X_{1,1}) \end{aligned}$$

the sampled correlations of f and f'. The next theorem gives the parameters of the probability distribution of the difference of the sampled correlations over a random sample. We denote by $\widehat{c}$ the difference $\widehat{\mathrm{cor}}(f) - \widehat{\mathrm{cor}}(f')$.

Theorem 1. *Taken over a random sample of size N, the mean of $\widehat{c}$ is equal to $2(p-p') = \mathrm{cor}(f) - \mathrm{cor}(f')$ and the variance is equal to*

$$\frac{4B}{N}\left(q - (p-p')^2\right), \tag{2}$$

where q is the probability that $f(x) \neq f'(x)$ over $x \in \mathbb{F}_2^n$, and B is the constant (1) determined by the sampling method.

Proof. With the notation defined above we have

$$N\widehat{c} = 2\left(X_{0,1} - X_{1,0}\right).$$

If sampling is with replacement, then $(X_{0,0}, X_{0,1}, X_{1,0}, X_{1,1})$ follows the multinomial distribution with sample size N, where $N = X_{0,0} + X_{0,1} + X_{1,0} + X_{1,1}$, and probabilities $(p_{0,0}, p_{0,1}, p_{1,0}, p_{1,1})$ where

$$p_{\gamma,\delta} = 2^{-n}|\{x \in \mathbb{F}_2^n \mid f(x) = \gamma, f'(x) = \delta\}|, \ \gamma, \delta \in \{0,1\}. \tag{3}$$

Thus $p_{0,0} + p_{0,1} = p$, $p_{0,0} + p_{1,0} = p'$, and $q = p_{0,1} + p_{1,0}$. It follows that the mean of $\widehat{c}$ is equal to

$$2(p_{0,1} - p_{1,0}) = 2(p - p').$$

To compute the variance of $\widehat{c}$, let us first compute the expected value of $(\frac{N}{2}\widehat{c})^2 = (X_{0,1} - X_{1,0})^2$. We get

$$\begin{aligned}
&\mathrm{Exp}\left(\left(X_{0,1} - X_{1,0}\right)^2\right) \\
&= \mathrm{Exp}\left(X_{0,1}^2\right) + \mathrm{Exp}\left(X_{1,0}^2\right) - 2\mathrm{Exp}\left(X_{0,1}X_{1,0}\right) \\
&= Np_{0,1} + N(N-1)p_{0,1}^2 + Np_{1,0} + N(N-1)p_{1,0}^2 - 2N(N-1)p_{0,1}p_{1,0} \\
&= Np_{0,1} + Np_{1,0} + N(N-1)\left(p_{0,1} - p_{1,0}\right)^2
\end{aligned}$$

using the properties of the multinomial distribution, see Appendix A.

From this we get that the variance of $\widehat{c}$ is equal to

$$\begin{aligned}
&\mathrm{Exp}\left(\widehat{c}^2\right) - \left(\mathrm{Exp}\left(\widehat{c}\right)\right)^2 \\
&= \frac{4}{N^2}\left(Np_{0,1} + Np_{1,0} + N(N-1)\left(p_{0,1} - p_{1,0}\right)^2\right) - 4\left(p_{0,1} - p_{1,0}\right)^2 \\
&= \frac{4}{N}\left(p_{0,1} + p_{1,0} - \left(p_{0,1} - p_{1,0}\right)^2\right),
\end{aligned}$$

which was the claim in case the sampling is done with replacement. If sampling is done without replacement, then $(X_{0,0}, X_{0,1}, X_{1,0}, X_{1,1})$ follows the multivariate hypergeometric distribution, which means that the variance must be multiplied by the finite population correction factor. □

In models where $\widehat{\mathrm{cor}}(f)$ and $\widehat{\mathrm{cor}}(f')$ are assumed to be independent, the variance of $\widehat{c}$ is two times the variance of one sampled correlation, that is, approximately equal to $2B/N$ using the binomial or hypergeometric distribution, see e.g. [BN17]. This is only a rough estimate of the actual value given by Theorem 1 which hides the cases where $q \neq \frac{1}{2}$ and $p \neq p'$. In particular, assuming independence, it is not possible to distinguish between the cases where just $\mathrm{cor}(f) = \mathrm{cor}(f')$, or actually $f = f'$. If $f = f'$ the variance given by (2) is equal to zero for all N.

Independence of $\widehat{\mathrm{cor}}(f)$ and $\widehat{\mathrm{cor}}(f')$ can be formally established if two independently drawn sets S and S' of inputs and $\widehat{\mathrm{cor}}(f)$ is computed from pairs $(x, f(x))$, $x \in S$, and $\widehat{\mathrm{cor}}(f')$ is computed from pairs $(x, f'(x))$, $x \in S'$, repectively. Assuming $q = \frac{1}{2}$ and $(p - p')^2 \approx 0$ we get that the variance of $\widehat{c}$ is equal to $2B/N$ independently of whether one sample of N inputs or two samples, each one consisting of N inputs, is used. Given N inputs x it takes $2N$ oracle calls to get $f(x)$ and $f'(x)$, while the corresponding numbers in the two-sample case are $2N$ inputs and $2N$ oracle calls.

In the next section, we will consider the distribution of $\widehat{c}$ by randomizing, in addition to the data sample, also over a pair (f, f') of Boolean functions, where the functions f and f' cannot, in general, be considered independent. In our applications, the pair (f, f') originates from a linear approximation formed for a cipher and a pair of related keys, see Section 4. Theorem 5 gives an example of the distribution of $\widehat{c}$ in such a situation without assuming independence of f and f'.

3 Difference of Correlations of Random Boolean Functions

Given two Boolean functions f and f', let us denote by c the difference of their correlations, that is, $c = \mathrm{cor}(f) - \mathrm{cor}(f')$. Let us now examine the probability distribution of c for a random pair of Boolean functions drawn equiprobably among all pairs of Boolean functions. Using the notation $p_{\gamma,\delta}$ given by (3), let us denote

$$N_{\gamma,\delta} = 2^n p_{\gamma,\delta}, \text{ for } \gamma, \delta \in \{0, 1\}.$$

Then the 4-tuple $(N_{0,0}, N_{0,1}, N_{1,0}, N_{1,1})$ follows the multinomial distribution such that $N_{0,0} + N_{0,1} + N_{1,0} + N_{1,1} = 2^n$ with probabilities $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$. Then $c = 2^{1-n}(N_{0,1} - N_{1,0})$. By using the same arguments as in the proof of Theorem 1 we can prove the following result.

Theorem 2. *The mean of c taken over a random pair (f, f') of Boolean functions from $\mathbb{F}_2^n$ to $\mathbb{F}_2$ is equal to zero. The variance of c is equal to 2^{1-n}.*

Proof. We substitute $N = 2^n$, $q = p = p' = \frac{1}{2}$, and $B = 1$ to the expression (2). □

Now, we can derive the parameters of the probability distribution of the difference $\widehat{c}$ of the sampled correlations taken over a random sample $S \subset \mathbb{F}_2^n$ of size N and a random pair (f, f') of Boolean functions from $\mathbb{F}_2^n$ to $\mathbb{F}_2$. This is achieved by combining the results of Theorems 1 and 2.

Theorem 3. *The mean of the difference of the sampled correlations of two Boolean functions taken over a random sample of size N and a random pair of Boolean functions from $\mathbb{F}_2^n$ to $\mathbb{F}_2$ is equal to zero. Its variance is equal to $\frac{2}{N}\left(1 + \frac{N-1}{2^n}\right)$ if sampling is with replacement. For sampling without replacement, the variance is equal to $\frac{2}{N}$.*

Proof. We get the variance of $\widehat{c}$ taken over the samples and pairs of Boolean functions by taking the expected value of the variance of $\widehat{c}$ from Theorem 1 and adding to it the variance of the mean c of $\widehat{c}$ as given by Theorem 2. We get

$$\mathrm{Exp}\left(\frac{4B}{N}\left(q - \frac{c^2}{4}\right)\right) + \mathrm{Var}\,(c) = \frac{4B}{N}\left(\frac{1}{2} - \frac{2^{1-n}}{4}\right) + 2^{1-n}.$$

For sampling with replacement, we take $B = 1$ to get

$$\frac{4}{N}\left(\frac{1}{2} - \frac{2^{1-n}}{4}\right) + 2^{1-n} = \frac{2}{N}\left(1 - \frac{1}{2^n} + \frac{N}{2^n}\right) = \frac{2}{N}\left(1 + \frac{N-1}{2^n}\right).$$

For sampling without replacemet, we substitute $B = \frac{2^n-N}{2^n-1}$ to obtain

$$\frac{2B}{N}\left(1-\frac{1}{2^n}\right)+2^{1-n} = \frac{2}{N}\frac{2^n-N}{2^n-1}\frac{2^n-1}{2^n}+2^{1-n} = \frac{2}{N}\left(\frac{2^n-N}{2^n}+\frac{N}{2^n}\right) = \frac{2}{N}.$$

□

Let us recall the corresponding result for the sampled correlation $\widehat{\mathrm{cor}}(f)$ of a random Boolean function f from [BN17]: the variance of $\widehat{\mathrm{cor}}(f)$ is equal to $\frac{1}{N}\left(1+N2^{-n}\right)$ for sampling with replacement and equal to $\frac{1}{N}$ for sampling without replacement. Hence for the difference of correlations of two equiprobably and independently drawn Boolean functions, these values are twice as large, that is, equal to the variances given by Theorem 3. While it is straightforward to derive these results by assuming independence of f and f', we have obtained them without any independence assumptions on the randomly and equiprobably chosen pair (f, f') by exploiting the properties of the multinomial distribution.

4 Applications to Linear Cryptanalysis

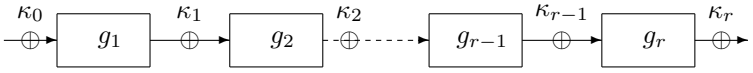
4.1 Linear trails

Let E_κ be the encryption function of an n-bit block cipher with n-bit plaintext $x = (x_1, \ldots, x_n) \in \{0,1\}^n$. A linear approximation $\langle a, x\rangle \oplus \langle b, E_\kappa(x)\rangle$ of E_κ is a single bit computed as $a_1x_1 \oplus \cdots \oplus a_nx_n \oplus b_1y_1 \oplus \cdots \oplus b_ny_n$, where $y = (y_1, \ldots, y_n) = E_\kappa(x)$. The n-bit vectors $a = (a_1 \ldots a_n)$ and $b = (b_1 \ldots b_n)$ are called input and output masks, respectively. The *correlation* of the linear approximation is defined as

$$\mathrm{cor}_{E_\kappa}(a,b) = 2^{-n}\sum_{x\in\{0,1\}^n}(-1)^{\langle a,x\rangle\oplus\langle b,E_\kappa(x)\rangle}. \tag{4}$$

It takes values between 1 and -1 included, and is a function of the key variable $\kappa \in \mathbb{F}_2^k$, where k is the key length in bits.

An iterative key-alternating block cipher with block size n processes plaintexts $x \in \mathbb{F}_2^n$ and round keys $\kappa_0, \kappa_1, \ldots, \kappa_r$ by iterating key-independent round functions g_i, $i = 1, \ldots, r$, round by round to obtain a ciphertext. Let us denote the sequence of round keys $\kappa_0, \ldots, \kappa_r$ by $\kappa \in \mathbb{F}_2^{(r+1)n}$ and denote the encryption function of the block cipher by E_κ. The round keys are added to the data between rounds using XOR addition:



Then the correlation of a linear approximation over E_κ can be expressed as

$$\mathrm{cor}_{E_\kappa}(a,b) = \sum_{\substack{\tau\\ \tau_0=a,\,\tau_r=b}}(-1)^{\langle\tau,\kappa\rangle}\prod_{i=1}^{r}\mathrm{cor}_{g_i}\left(\tau_{i-1},\tau_i\right), \tag{5}$$

where the sum is taken over all $(r+1)$-tuples $\tau = (\tau_0, \tau_1, \ldots, \tau_{r'})$ such that $\tau_0 = a$ and $\tau_r = b$ [DGV94]. The sequence τ is called a *linear trail* of the linear approximation $\langle a, x\rangle \oplus \langle b, E_\kappa\rangle$. The quantity $\prod_{i=1}^{r}\mathrm{cor}_{g_i}\left(\tau_{i-1},\tau_i\right)$ is called the *trail correlation* of trail τ and is independent of the key. The term $\langle\tau,\kappa\rangle = \langle\tau_0,\kappa_0\rangle \oplus \langle\tau_1,\kappa_1\rangle \oplus \cdots \oplus \langle\tau_r,\kappa_r\rangle$ depends solely on the key schedule.

The existing works on related-key linear cryptanalysis [RN13] and [BBR+13] consider the difference of correlations $\mathrm{cor}_{E_K}(a,b)$ and $\mathrm{cor}_{E_{K\oplus\Delta}}(a,b)$ for two keys (sequences of

round keys) with a fixed difference Δ. For an iterative block cipher, this difference can be written as follows

$$\begin{aligned}&\mathrm{cor}_{E_K}(a,b) - \mathrm{cor}_{E_{K\oplus\Delta}}(a,b)\\ &= \sum_{\substack{\tau \\ \tau_0=a,\, \tau_r=b}} \left(1-(-1)^{\langle\tau,\Delta\rangle}\right)(-1)^{\langle\tau,K\rangle}\prod_{i=1}^{r}\mathrm{cor}_{g_i}\left(\tau_{i-1},\tau_i\right).\end{aligned}$$

All terms with $\langle\tau,\Delta\rangle = 0$ cancel out, meaning that the number of the possible values of the correlation difference is smaller than the number of the values of the correlation. If moreover,

$$\prod_{i=1}^{r}\mathrm{cor}_{g_i}\left(\tau_{i-1},\tau_i\right) = 0$$

for all τ with $\langle\tau,\Delta\rangle = 1$, then the difference of the related-key correlations is equal to zero. This property, named as key difference invariant bias (KDIB), was identified in the block ciphers LBlock and TWINE in [BBR+13].

4.2 Approximate continuous distributions

In the context of linear cryptanalysis the Boolean functions f considered in Section 2 are linear approximations of a keyed block cipher E_κ defined as follows

$$f_\kappa(x) = \langle a, E_\kappa(x)\rangle \oplus \langle b, x\rangle,\ x \in \mathbb{F}_2^n,$$

where $a, b \in F_2^n$ are some fixed linear masks and n is the block size of the cipher. It is practical to use continuous approximations of the discrete binomial and hypergeometric distributions for statistical analysis of the attacks. Similarly, the multinomial distribution and the multivariate hypergeometric distributions are approximated using the multivariate normal distribution. Further, it is well known that any linear (or affine) transformation of a multivariate normal deviate is again a multivariate normal deviate. This means that the distribution of the difference $X_{0,1} - X_{1,0}$ considered in Theorem 1 can be approximated using a normal distribution. From this we get the following corollary.

Corollary 1. *The difference $\widehat{c}$ of the sampled correlations of a linear approximation $\langle a, E_\kappa(x)\rangle \oplus \langle b, x\rangle$ of an n-bit block cipher E_κ computed for two different keys $\kappa = K$ and $\kappa = K'$ is approximately normally distributed with mean $c = \mathrm{cor}(f_K) - \mathrm{cor}(f_{K'})$ and variance equal to*

$$\frac{4B}{N}\left(q - \frac{c^2}{4}\right), \tag{6}$$

where N is the size of the sample of data triplets $(x, E_K(x), E_{K'}(x)$ and

$$q = 2^{-n}\left|\{x \in \mathbb{F}_2^n \mid \langle a, E_K(x)\rangle \neq \langle a, E_{K'}(x)\rangle\}\right|. \tag{7}$$

The constant B is defined by (1).

4.3 Cryptanalysis of Röck and Nyberg

By Equation [5], the correlation $\mathrm{cor}_{E_K}(a,b)$ is the sum of signed trail correlations, where the signs depend on the key, so the correlation takes different values as the key varies. Matsui's Algorithm 1 type attacks are possible for ciphers for which, for a significant proportion of the keys, the number of different values of $\mathrm{cor}_{E'_K}(a,b)$ is small and they are sufficiently apart from each other.

In Matsui's original attack of the DES algorithm, the linear approximation is composed of only one trail meaning that the correlation has only two possible values $\pm\rho$, where ρ is the trail correlation [Mat93]. To thwart this attack, modern ciphers are designed so that a single trail cannot determine the correlation of a linear approximation. In an attempt to reduce the number of trails, Röck and Nyberg presented a generalisation of Matsui's Algorithm 1 to the related key setting [RN13]. They divided the keys K into key classes $\mathcal{K}(c)$ as follows: a key K belongs to $\mathcal{K}(c)$ if $\mathrm{cor}(f_K) - \mathrm{cor}(f_{K\oplus\Delta}) = c$, where c is a possible value of the correlation difference.

To recover the key the cryptanalyst computes the sampled correlation difference. To determine the success and error probabilities of the solution the cryptanalyst needs to know the probability distribution of the sampled correlation difference. In [RN13], it was assumed that the sampled correlations of the two related keys are independent, by arguing that this is at least the case if for each key the sample is drawn separately and independently. Since the sampled correlation for a fixed key is normally distributed with variance $1/N$ by the approximation of the binomial distribution (assuming sampling with replacement), the difference of two such sampled correlations is then normally distributed with the mean $\mathrm{cor}(f_K) - \mathrm{cor}(f_{K\oplus\Delta})$ and variance $2/N$.

Using Corollary 1 we can remove the assumption about independence of the sampled correlations and get a more detailed understanding of the statistical behaviour of $\widehat{c}$ as given by the following result.

Theorem 4. *In the setting of [RN13], let us denote by $Q(c)$ the average of q, defined by Equation (7) taken over all keys in $\mathcal{K}(c)$. Then the distribution of $\widehat{c}$ over a random sample of size N and a random key in $\mathcal{K}(c)$ is approximately normal with the mean c and variance equal to*

$$\frac{4B}{N}\left(Q(c) - \frac{c^2}{4}\right).$$

Proof. Since the expected value of $\widehat{c}$ is constant for all key pairs in $\mathcal{K}(c)$, the variance is the mean of the variance (6). □

If $Q(c) = \frac{1}{2}$ and $c^2 \ll \frac{1}{2}$ and $B = 1$ we obtain the same distribution parameters as in [RN13] but now without the assumption about independence of the sampled correlations. For ciphers with $Q(c) < \frac{1}{2}$, if any, the variance could be smaller than estimated, which may lead to improvements of the attack.

4.4 Key difference invariant bias

The KDIB cryptanalysis proposed by Bogdanov *et al.* [BBR+13] is a key-recovery attack similar to Matsui's Algorithm 2 [Mat93]. The distinguisher is based on the difference between the statistical distributions of some test statistic computed for the wrong key and the right key. While in Matsui's linear cryptanalysis this test statistic is the sampled correlation, Bogdanov *et al.* used the (squared) difference of correlations computed for two keys with a fixed difference in their corresponding sequences of round keys. In statistical cryptanalysis, the wrong-key behaviour of the statistic is modelled according to the behaviour of the corresponding statistic computed for a random permutation. The right-key behaviour, which should be different, is based on some non-randomness property, which in this case, is the KDIB property.

By the KDIB property there is a linear approximation with input mask a and output mask b and key difference Δ of the sequences of round keys such that the difference of correlations $\mathrm{cor}_{E_K}(a,b) - \mathrm{cor}_{E_{K\oplus\Delta}}(a,b)$ is equal to zero for all keys [BBR+13]. Hence the difference of sampled correlations computed for a data sample obtained from the cipher

with a KDIB property can be expected to have a smaller variance compared to the one in the random case.

Using Corollary 1 we get the following result.

Theorem 5. *Suppose a key-alternating block cipher has the KDIB property with input mask a and output mask b and a key difference Δ. Then the probability distribution of the difference of the sampled correlations*

$$\widehat{c} = \widehat{\mathrm{cor}}_{E_K}(a,b) - \widehat{\mathrm{cor}}_{E_{K\oplus\Delta}}(a,b)$$

taken over a random sample of N plaintexts and over a random key K is approximately normal with the mean equal to zero and the variance equal to

$$\frac{4BQ}{N},$$

where Q is the average of the probability $\Pr(E_K(a,b) \neq E_{K\oplus\Delta}(a,b))$ taken over a random key K.

Proof. By the KDIB property we set $c = 0$, for all keys, in Corollary 1. It follows that the variance of c taken over a random key is equal to zero. Hence the variance of $\widehat{c}$ is equal to the mean of the variance given in (6) over a random key. □

By setting $Q = \frac{1}{2}$ and $B = 1$ for sampling with replacement, we obtain that $\frac{N}{2}\widehat{c}^2$ is a χ^2 deviate with one degree of freedom. If we take λ such variables, where λ is high enough, assume their independence, and use their sum as a statistic as done in [BBR+13], we get the result of their Proposition 2 after approximating the χ^2 distribution with a normal distribution. Proposition 2 of [BBR+13] makes an additional assumption that, for each linear approximation, the two sampled correlations (or the counters) computed for the related key pair are statistically independent.

In general, the related-key linear cryptanalysis of a block cipher is based on some non-randomness property of the difference of the expected values of the correlations, or equivalently, of the data distributions, which holds for all related key pairs. The success of the attack depends on how well the related-key behaviour of the cipher can be distinguished from the random behavior of the difference of correlations and data distributions.

For an example how to set up a linear related-key distinguisher for key recovery and establish a connection between the error probabilities and the data requirement we refer to [BBR+13]. In the next subsection, we will determine the wrong-key distribution of the correlation difference.

4.5 Wrong-key distribution for related-key linear distinguisher

The distribution of the correlation of a linear approximation taken over a random permutation can be approximated by the distribution of the correlation of a random Boolean function [DR07]. This property was later established also for the sampled correlation [AKN21]. Key-recovery attacks on block ciphers (key-dependent permutations), which exploit a statistical distinguisher between the right-key and wrong-key behaviours of the cipher, typically model the wrong-key behaviour according to the random case.

Analogically, a practical model of the wrong-key behaviour of the (sampled) difference of related-key correlations of a linear approximation is obtained by imitating the behaviour of the difference of the (sampled) correlations of two random Boolean functions. By Theorem 3 we get the following result.

Corollary 2. *The difference of sampled related-key correlations*

$$\widehat{c} = \widehat{\mathrm{cor}}\left(\langle a, E_K\rangle \oplus \langle b, x\rangle\right) - \widehat{\mathrm{cor}}\left(\langle a, E_{K'}\rangle \oplus \langle b, x\rangle\right)$$

taken over a random wrong related-key pair (K, K') and over a random sample of size N is approximately normally distributed with mean equal to zero. The variance is equal to

$$\frac{2}{N}\left(1+\frac{N-1}{2^n}\right)$$

if sampling is with replacement. If sampling is without replacement, the variance is equal to

$$\frac{2}{N}.$$

Proposition 3 [BBR+13] gives an approximate probability distribution of the sum of squares of correlation differences for λ linear approximations under the following assumptions

1. λ is high enough,
2. all 2λ sampled correlations are statistically independent, and
3. the sample of N known plaintexts may contain repetitions.

Using Corollary 2 we can remove assumption 1 and give the result without normal approximation of χ^2 distribution (which required λ to be high enough). In addition to the case defined by assumption 3 we also consider sampling without replacement. We can relax assumption 2 and allow the two counters for each linear approximation to be dependent. Yet we still need statistical dependence of the linear approximations. For the formulation of this result, see Corollary 3 in the next section.

5 Multiple and Multidimensional Linear Approximations

5.1 Definition of the statistic

Let

$$\langle a_\alpha, E_\kappa(x)\rangle \oplus \langle b_\alpha, x\rangle,\ \alpha = 1,\ldots,M, \tag{8}$$

be a set of M nonzero linear approximations of an n-bit block cipher E_κ. Let (K, K') be a pair of related keys and f_α and f'_α denote these linear approximations applied to E_K and $E_{K'}$, respectively. Further, we denote

$$\begin{aligned} c_\alpha &= \mathrm{cor}(f_\alpha) - \mathrm{cor}(f'_\alpha) \text{ and} \\ \widehat{c}_\alpha &= \widehat{\mathrm{cor}}(f_\alpha) - \widehat{\mathrm{cor}}(f'_\alpha), \end{aligned}$$

where the sampled correlations $\widehat{\mathrm{cor}}(f_\alpha)$, and $\widehat{\mathrm{cor}}(f'_\alpha)$, $\alpha = 1,\ldots,M$, are computed for a random set of N plaintexts drawn either with or without replacement. Further, we denote by q_α the probability that $f_\alpha(x) \neq f'_\alpha(x)$ taken over a random plaintext $x \in \mathbb{F}_2^n$. By Corollary 2 we have that

$$\frac{N}{2\left(1+(N-1)2^{-n}\right)}\widehat{c} \quad \text{or} \quad \frac{N}{2}\widehat{c},$$

if sampling is with or without replacement, respectively, follows the χ^2 distribution with one degree of freedom. In statistical cryptanalysis we can assume that N and 2^n are large and make the following approximations $N \approx N-1$ and $2^n \approx 2^n - 1$.

In the analysis of the sampled related-key correlation difference $\widehat{c}_\alpha$ we propose to use the following statistic

$$T = \frac{N}{2\left(B + N2^{-n}\right)} \sum_{\alpha=1}^{M} \widehat{c}_\alpha^2, \tag{9}$$

where B is the constant defined in (1).

5.2 Wrong-key distribution of the statistic

Corollary 2 gives the distribution of a sampled related-key correlation difference over wrong key pairs. Using it we obtain the following information about the distribution of T for a family of M linear approximations.

Corollary 3. *The mean of the statistic*

$$T = \frac{N}{2\,(B + N2^{-n})} \sum_{\alpha=1}^{M} \widehat{c}_\alpha^2,$$

taken over a random sample of size N and a random wrong related-key pair is equal to M. Suppose, moreover, that the linear approximations are independent, in the sense that $\widehat{c}_\alpha$, $\alpha = 1, \ldots, M$, are statistically independent over a random sample of size N and over a random pair of permutations. Then the statistic T follows the χ^2 distribution with M degrees of freedom.

The shape of the distribution of T must be considered separately for different kinds of families of linear approximations. It can be argued that if the linear approximations $\langle a_\alpha, E_\kappa(x)\rangle \oplus \langle b_\alpha, x\rangle$, $\alpha = 1, \ldots, M$, are linearly independent, that is, the mask pairs (a_α, b_α) are linearly independent, then they are also essentially statistically independent. In practice, χ^2 distribution may work well also for other kinds of sets of linear approximations even if the prerequisites of Pearson's χ^2 test are not fully satisfied [BTV18, FN20].

It might be possible, although elaborate, to use the properties of the multinomial distribution in the similar way it was done in [AKN21] for the single-key setting to compute the variance of the capacity of a multidimensional linear approximation. In this way, one could obtain the variance of T, while the form of the distribution still would remain an open problem.

5.3 The statistic without independence of linear approximations

We start by examining the statistic T for a fixed pair of permutations, either random or cipher, with related keys identified by a key pair (K, K'). We determine the expected value of T over a random sample of N plaintexts.

Theorem 6. *For any given pair of keys (K, K') with $c_\alpha = \mathrm{cor}(f_\alpha) - \mathrm{cor}(f'_\alpha)$ the statistic T defined by (9) has the following mean over a random sample of size N*

$$\mathrm{Exp}\,(T) = \frac{2B}{B + N2^{-n}} \sum_{\alpha=1}^{M} q_\alpha + \frac{N}{2\,(B + N2^{-n})} \sum_{\alpha=1}^{M} c_\alpha^2, \tag{10}$$

where B is the constant defined in (1).

If the set of linear approximations satisfies

$$\frac{1}{2^t - 1} \sum_{\alpha=1}^{2^t-1} q_\alpha = Q$$

then the mean of T taken over a random data sample of size N is equal to

$$\frac{2(2^t - 1)B}{B + N2^{-n}} Q + \frac{N}{2\,(B + N2^{-n})} \sum_{\alpha=1}^{M} c_\alpha^2.$$

Proof. For each $\alpha = 1, \ldots, M$, the expected value of $\widehat{c}_\alpha$ is equal to c_α. By applying the expression (6) of the variance we get

$$\mathrm{Exp}\left(\widehat{c}_\alpha^2\right) = \frac{4B}{N}\left(q_\alpha - \frac{c_\alpha^2}{4}\right) + c_\alpha^2 = \frac{4B}{N} q_\alpha + \left(1 - \frac{B}{N}\right) c_\alpha^2.$$

By summing over $\alpha = 1, \ldots, M$ and using $N - B \approx N$ we get the claim. □

We leave it as an open task to investigate the distribution of T. In practical applications the χ^2 distribution may often give a sufficiently accurate approximation. For example, in the case where $Q = \frac{1}{2}$ the distribution of $(1 + (N/B)2^{-n})\,T$ could be close to the noncentral χ^2 distribution with M degrees of freedom and noncentrality parameter approximately equal to

$$\frac{N}{2B} \sum_{\alpha=1}^{M} c_\alpha^2.$$

5.4 The distribution view

In single-key multidimensional linear cryptanalysis, the χ^2 distribution of the statistic computed from the squared correlations of the linear approximations arises naturally from the related multinomial distribution using Pearson's χ^2 test [HCN19, AKN21]. In related-key multidimensional linear cryptanalysis this is not clear. Let us have a closer look.

A multidimensional linear approximation is a linear space where the nonzero elements are given by the mask pairs (a_α, b_α), $\alpha = 1, \ldots, M$. If the dimension is t, then $M = 2^t - 1$. Then the multidimensional linear approximation can also be given by a vectorial Boolean function. For example, when applied to the cipher E_κ with key κ, the nonzero components of this vectorial Boolean function are the Boolean functions defined by the expression (8). Moreover, we can assume that the indexing is such that $\alpha \in \mathbb{F}_2^t$ and the mapping $\alpha \mapsto (a_\alpha, b_\alpha)$ is a linear isomorphism.

If we denote by F this vectorial Boolean function for $\kappa = K$ then we can assume that

$$\langle \alpha, F(x) \rangle = \langle a_\alpha, E_K(x) \rangle \oplus \langle b_\alpha, x \rangle, \text{ for all } \alpha \in \mathbb{F}_2^t.$$

Similarly, we define the vectorial Boolean function F' to correspond this multidimensional linear approximation applied to $E_{K'}$. For each $\eta \in \mathbb{F}_2^t$ we define the probabilities

$$\begin{aligned} p_\eta &= \Pr\left(F(x) = \eta\right) = 2^{-n} |\{x \in \mathbb{F}_2^n \mid F(x) = \eta\}|, \text{ and} \\ p'_\eta &= \Pr\left(F'(x) = \eta\right) = 2^{-n} |\{x \in \mathbb{F}_2^n \mid F'(x) = \eta\}|. \end{aligned}$$

To observe the difference of the distibutions p_η and p'_η we draw a random sample S of N plaintexts x from $\mathbb{F}_2^n$. We denote

$$\begin{aligned} X(\eta) &= N^{-1} |\{x \in S \mid F(x) = \eta\}| \\ X'(\eta) &= N^{-1} |\{x \in S \mid F'(x) = \eta\}|. \end{aligned}$$

Then the values $F(x)$ and $F'(x)$ computed for a single x and a single key pair (K, K') typically increment counters for two different values of η, which means that the categories (labelled by η) are not sampled independently thus violating the prerequisites of Pearson's χ^2 test.

Based on the connections

$$\begin{aligned} p_\eta &= 2^{-t} \sum_{\alpha=0}^{2^t-1} (-1)^{\langle \alpha, \eta \rangle} c_\alpha, \text{ and} \\ X(\eta) &= N 2^{-t} \sum_{\alpha=0}^{2^t-1} (-1)^{\langle \alpha, \eta \rangle} \widehat{c}_\alpha, \; \eta \in \mathbb{F}_2^t, \end{aligned}$$

we get that the statistic T can also be expressed in the following form

$$T = \frac{2^t}{2N\left(B + N2^{-n}\right)} \sum_{\eta \in \mathbb{F}_2^t} \left(X(\eta) - X'(\eta)\right)^2 .$$

The form of T given above is applicable to the analysis of different types of distributions of cipher data obtained using linear projections. Such distributions occur for example in statistical saturation attack. The relation between the statistical saturation attack and multidimensional linear cryptanalysis (in the single-key setting) has been studied by Blondeau and Nyberg [BN14].

6 Conclusions

In this work, we studied the probability distribution of the difference of sampled correlations of two Boolean functions over a random sample of their inputs and showed that it is approximately normal and gave its parameters. Further, we established this distribution also over a random pair of Boolean functions. These results were then applied to related-key linear cryptanalysis. By modelling the wrong-key behaviour of the correlation of a linear approximation according to random behaviour, we obtain the wrong-key distribution. For the right key, the cryptanalyst exploits some non-random property of the cipher. We revisited the KDIB cryptanalysis and established the right-key distribution for a single linear approximation without any independence assumption about the sampled correlations computed for related keys. The variance of this distribution depends on the probability q that the linear approximations take different values when computed for the cipher with two related keys. This probability may not always equal to $\frac{1}{2}$. It would be interesting to determine the probability q and study its impact to the variance of the correlation difference for LBlock and TWINE. Another line of work, for related-key linear cryptanalysis more generally, would be its applications to tweaked block ciphers.

We also discuss the previously proposed solution to obtain independence of the sampled related-key correlations by using two independent samples to compute the correlations. While this would work for distributions taken over the data, it does not help when the distributions are taken over a random (right or wrong) key. In related-key cryptanalysis, in particular, it is not realistic to assume that the sampled related-key correlations are independent.

For related-key applications involving multiple linear approximations we proposed a χ^2-type statistic, which indeed has a χ^2 distribution under the additional assumption that the linear approximations are independent. When trying to remove this assumption by considering linear or affine spaces of linear approximations, as it is done in the single-key setting, we encountered problems, which were left for future work.

Acknowledgments

The author thanks Meiqin Wang for motivating discussions and finding a gap in an earlier version of the proof of Theorem 1. The constructive comments of the reviewers are also gratefully acknowledged.

References

[AKN21] Tomer Ashur, Mohsin Khan, and Kaisa Nyberg. Structural and statistical analysis of multidimensional linear approximations of random functions and permutations. *IEEE Trans. IT*, 2021. Accepted for publication, available: https://ieeexplore.ieee.org/document/9617455.

[BBR+13] Andrey Bogdanov, Christina Boura, Vincent Rijmen, Meiqin Wang, Long Wen, and Jingyuan Zhao. Key difference invariant bias in block ciphers. In Kazue Sako and Palash Sarkar, editors, *Advances in Cryptology - ASIACRYPT 2013 - 19th International Conference on the Theory and Application of Cryptology and Information Security, Bengaluru, India, December 1-5, 2013, Proceedings, Part I*, volume 8269 of *Lecture Notes in Computer Science*, pages 357–376. Springer, 2013.

[BN14] Céline Blondeau and Kaisa Nyberg. Links between truncated differential and multidimensional linear properties of block ciphers and underlying attack complexities. In Phong Q. Nguyen and Elisabeth Oswald, editors, *Advances in Cryptology - EUROCRYPT 2014 - 33rd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Copenhagen, Denmark, May 11-15, 2014. Proceedings*, volume 8441 of *Lecture Notes in Computer Science*, pages 165–182. Springer, 2014.

[BN17] 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 Cryptogr.*, 82(1-2):319–349, 2017.

[BTV18] Andrey Bogdanov, Elmar Tischhauser, and Philip S. Vejre. Multivariate profiling of hulls for linear cryptanalysis. *IACR Trans. Symmetric Cryptol.*, 2018(1):101–125, 2018.

[DGV94] Joan Daemen, René Govaerts, and Joos Vandewalle. Correlation matrices. 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 275–285. Springer, 1994.

[DR07] Joan Daemen and Vincent Rijmen. Probability distributions of correlation and differentials in block ciphers. *J. Mathematical Cryptology*, 1(3):221–242, 2007.

[FN20] Antonio Flórez-Gutiérrez and María Naya-Plasencia. Improving key-recovery in linear attacks: Application to 28-round PRESENT. In Anne Canteaut and Yuval Ishai, editors, *Advances in Cryptology - EUROCRYPT 2020 - 39th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, May 10-14, 2020, Proceedings, Part I*, volume 12105 of *Lecture Notes in Computer Science*, pages 221–249. Springer, 2020.

[HCN19] Miia Hermelin, Joo Yeon Cho, and Kaisa Nyberg. Multidimensional linear cryptanalysis. *J. Cryptology*, 32(1):1–34, 2019.

[Mat93] Mitsuru Matsui. Linear cryptanalysis method for DES cipher. In Tor Helleseth, editor, *Advances in Cryptology - EUROCRYPT '93, Workshop on the Theory and Application of of Cryptographic Techniques, Lofthus, Norway, May 23-27, 1993, Proceedings*, volume 765 of *Lecture Notes in Computer Science*, pages 386–397. Springer, 1993.

[RN13] Andrea Röck and Kaisa Nyberg. Generalization of Matsui's Algorithm 1 to linear hull for key-alternating block ciphers. *Des. Codes Cryptography*, 66(1-3):175–193, 2013.

[RT89] J. N. K. Rao and D. R. Thomas. Chi-squared tests for contingency tables. In C. J. Skinner, D. Holt, and T. M. F. Smith, editors, *Analysis of Complex Surveys*, pages 89–104. John Wiley & Sons, Chichester, UK, 1989.

A Properties of the Multinomial Distribution

The set of k integers $(Z_1, \ldots, Z_k)$ is said to follow a multinomial distribution with parameters m and $p_1, \ldots, p_k$, where m is a positive integer and p_η, $\eta = 1, \ldots, k$ are positive and $p_1 + \cdots + p_k = 1$, if $Z_1 + \cdots + Z_k = m$ and

$$\Pr(Z_1 = z_1, \ldots, Z_k = z_k) = \frac{m!}{z_1! \cdots z_k!} p_1^{z_1} \cdots p_k^{z_k},$$

for any k-tuple of nonnegative integers $(z_1, \ldots, z_k)$ with $z_1 + \cdots + z_k = m$.

The following properites are needed in the proof of Theorem 1:

$$\begin{aligned} \mathrm{Exp}(Z_\eta) &= mp_\eta, \\ \mathrm{Exp}(Z_\eta Z_\zeta) &= m(m-1)p_\eta p_\zeta,\ \eta \neq \zeta, \\ \mathrm{Exp}(Z_\eta^2) &= mp_\eta + m(m-1)p_\eta^2. \end{aligned}$$