Joint Data and Key Distribution of Simple, Multiple, and Multidimensional Linear Cryptanalysis Test Statistic and Its Impact to Data Complexity

Céline Blondeau and Kaisa Nyberg

Department of Computer Science, Aalto University School of Science, Finland
celine.blondeau@aalto.fi, kaisa.nyberg@aalto.fi

Abstract. The power of a statistical attack is inversely proportional to the number of plaintexts needed to recover information on the encryption key. By analyzing the distribution of the random variables involved in the attack, cryptographers aim to provide a good estimate of the data complexity of the attack. In this paper, we analyze the hypotheses made in simple, multiple, and multidimensional linear attacks that use either non-zero or zero correlations, and provide more accurate estimates of the data complexity of these attacks. This is achieved by taking, for the first time, into consideration the key variance of the statistic for both the right and wrong keys. For the family of linear attacks considered in this paper, we differentiate between the attacks which are performed in the known-plaintext and those in the distinct-known-plaintext model.

Keywords: iterated block cipher, key-recovery attack, simple linear attack, multidimensional linear attack, zero-correlation linear attack, key-difference-invariant-bias attack, known plaintext, distinct known plaintext, chosen plaintext, key variance, statistical model.
MSC 2010 codes: 94A60, 11T71, 68P25

1 Introduction

1.1 Background and Previous Work

After being introduced a quarter of a century ago, the classical linear [25] and differential [5] cryptanalysis methods have been extended to various more evolved statistical attacks. A generalization of differential attacks, known as truncated differential attacks [21], take advantage of simultaneously multiple differential approximations. Since the invention of linear cryptanalysis, several authors have considered taking advantage of multiple linear approximations, but the first general statistical model was not presented until in [6]. The theoretically restrictive assumption of independence of linear approximations was removed in the model developed in [19] on the cost of taking into account a family of linear approximations which covers a linear space excluding zero. More recently zero-correlation linear attacks [9,11,13] were introduced. While the linear attacks usually exploit

linear approximations that are, for a random encryption key, expected to have correlation of large absolute value, zero-correlation attacks make use of linear approximations which are unbiased, that is, have correlation equal to zero, for all encryption keys.

The aim of a statistical key-recovery attack is to search for the correct value for some bits of the encryption key based on a known statistical property of the cipher. This property is expected to be detected only for the correct key candidate, while wrong key candidates which are far from satisfying the property can be discarded. To estimate the data complexity of a statistical attack, the probability distributions of the involved random variables for the right and wrong keys are analyzed. These distributions depend on both the data sample used to compute it as well as the encryption key and the key candidate.

Distinct-known-plaintext attacks. The work done in this paper is motivated by the zero-correlation linear attacks, where two different statistical models had been in use. The model for multidimensional linear zero-correlation attacks assumed distinct known plaintext [11], while the attacks using multiple independent linear approximations assumed just known plaintext not excluding repetitions [9, 13]. We observe that the differentiating factor of the statistical models is not whether the attack is multidimensional or multiple, but instead, whether distinct plaintext is assumed or not. In this paper, we develop on distinct-known-plaintext attacks. In particular, we show that avoiding repetition in the plaintexts when the data complexity is close to the full codebook could present some interest not only for multidimensional zero-correlation attacks but also for multiple zero-correlation attacks and more generally for all known-plaintext attacks. In particular using distinct-known-plaintext we improve the data complexity of some multiple zero-correlation linear attacks and key-invariant attacks [8].

Right- and wrong-key hypothesis in key-recovery attacks. Previously, most statistical models used in linear attacks determine and exploit distributions of the attack statistic with fixed keys and taking only the data as random variable. Then it is assumed that for all cipher keys, the distributions for wrong key candidates are identical, and similarly, that the distributions obtained with correct key are identical. This practice may be due to the fact that for most ciphers, we are only able to estimate the expected value of a linear correlation, but estimating the variance is difficult. Previously, in [17,18] the authors provide experiments to show that also significant variances occur. In particular they present an estimate of the variance of correlation in the wrong-key case. In [12], this influence of the wrong-key variance for a simple linear attack was taken into consideration and a better estimate of the data complexity of a linear attack was obtained and demonstrated in experiments. In [20], the distribution of the capacity for the right encryption key was established and was used to determine weak-key quantiles, that is, lower bounds of capacity that are satisfied by a given proportion, say one half, or 30% of the keys. Such approach was previously taken in [23] in the case of single linear hull. Zero-correlation multiple or multidimensional linear attack is a special case, where for all correct keys the data is drawn from an

identical probability distribution. For the wrong keys, however, the distributions are not identical, a fact that has been ignored in the literature until recently.

Contributions of this paper. In this paper, we present the first complete treatment of the statistical distributions of linear attack test statistics, that is, the empirical correlations and capacities, by considering both the data and the key as random variables. We analyze and combine the different models previously presented and go beyond by studying the joint probability distribution of the test statistic both in the wrong-key and right-key case and present formulas for success probability and data complexity. In addition, the new statistical model takes also into account whether the data sample is obtained by the usual known plaintext sampling or the recently considered distinct plaintext sampling first introduced in the context of multidimensional zero-correlation attacks [11].

Outline. The outline of this paper is as follows. We start in Section 2 by presenting all the required background about linear key-recovery attacks including statistical tools and properties of correlations. Section 3 is dedicated to the classical linear context. We present separately the case of linear approximation with single dominant characteristic and the case of several characteristic. In both contexts we take into consideration the key deviation of the correlation for both the wrong and right keys. Section 4 is dedicated to the presentation of the multiple and multidimensional linear attacks and a more accurate statistical model these attacks is presented. Based on this model, in Section 5, we provide new estimates of the data complexity of a multiple/multidimensional linear attacks and present in Section 6, as an application, the case of zero-correlation linear cryptanalysis. Section 7 concludes this paper.

2 Correlation and Statistical Key-Recovery Attack

2.1 Correlation and Key Search

Iterated block cipher. Matsui's Algorithm 2 [25] is a statistical cryptanalysis method for finding a part of the last round key for an iterated block cipher. An iterated block cipher with block size n bits processes plaintext $x \in \{0,1\}^n$ and expanded key $K' = (k_1, k_2, \ldots, k_r)$ by iterating a key-dependent round function g_k with $k = k_i$, $i = 1, 2, \ldots, r$, to obtain ciphertext y, see Figure 1.

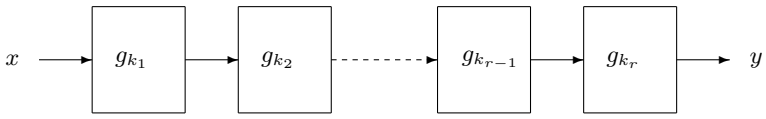


Fig. 1: Iterated block cipher of r rounds with round function g and expanded encryption key $(k_1, k_2, \ldots, k_r)$.

Last-round key-recovery attack. Key-recovery attacks on iterated block ciphers allow to recover information on the key-bits involved in one or more of the first rounds or last rounds, or both [4, 25]. To keep the notation simple, we restrict the description in this paper to key recovery attacks over the last round.

Let $K' = (K, k_r)$ denote the extended key, that is, K is the concatenation of the keys $k_1, k_2, \ldots, k_{r-1}$. The first prerequisite for the last-round key recovery attack is the so-called last round trick, which means that there is a part y' of ciphertext y that can be computed from a part of the input to the last round and a part of the last round key k_r using a bijective function. Let us denote this part of k_r by k'_r and the bijective function by $G_{k'_r}$. Then it holds (see Figure 2) that

$$y' = G_{k'_r}(E'_K(x)),$$

where E'_K is the encryption function over $r-1$ rounds with its range restricted to the domain of $G_{k'_r}$.

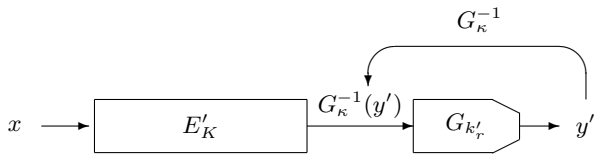


Fig. 2: Last round key recovery attack.

Then the attacker tries all possible key candidates κ of k'_r. For the right key candidate $\kappa = k'_r$ it holds that $G_\kappa^{-1}(y') = E'_K(x)$, that is, the resulting data is the same as obtained by encrypting plaintext x over $r-1$ rounds of the cipher.

The last round trick is not specific to linear cryptanalysis, but is often used also in the context of other statistical cryptanalysis, e.g., differential attacks. The important prerequisite for this type of attack is that the cipher has a statistical property that can be observed from the data obtained from $r-1$ rounds of the cipher.

Correlation. The classical linear cryptanalysis exploits a biased linear combination of input and output bits over $r-1$ rounds of encryption E'_K. Given a vector u in the plaintext space and a vector v in the output space of E'_K the Boolean function $u \cdot x \oplus v \cdot E'_K(x)$ is called the linear approximation over E'_K with input mask u and output mask v, where "$\cdot$" denotes the inner product. For example, $u \cdot x$ is the *modulo* 2 sum of the coordinate-wise products of u and x. The strength of this linear approximation, also denoted as (u, v), is measured by its correlation defined as

$$\mathrm{cor}(u,v)(K) =$$
$$2^{-n}\Big[\#\left\{x \in \{0,1\}^n | \, u \cdot x + v \cdot E'_K(x) = 0\right\} - \#\left\{x \in \mathbb{F}_2^n | u \cdot x + v \cdot E'_K(x) = 1\right\}\Big].$$

Description of the attack. In the offline analysis of the cipher, the attacker selects a linear approximation (u, v) such that $\mathrm{cor}(u, v)(K)$ is large in absolute value, for all K. To launch an online attack, the cryptanalyst obtains a data sample from the cipher. We denote the data sample by D and the number of data items in D by N. In the case of classical linear cryptanalysis D is a set of plaintext-ciphertext pairs (x, y). From the ciphertext y only the part y' is used in the attack.

Then the correct value k'_r is searched by trying all candidates κ and seeing if the cipher property, in this case large correlation, is observable from the data. To this end, the attacker obtains pairs $(x, G_\kappa^{-1}(y'))$, for all $(x, y') \in D$, and determines the empirical correlation

$$\hat{c}(D, K, k_r, \kappa) = \frac{2}{N}\#\{(x, y') \in D \mid u \cdot x + v \cdot G_\kappa^{-1}(y') = 0\} - 1.$$

After examining all candidates κ of k'_r, the cryptanalyst selects a set of key-candidates κ that achieve the top largest values $|\hat{c}(D, K, k_r, \kappa)|$.

Success probability and advantage of the attack. It has become customary to denote by 2^{-a} the proportion of keys that are discarded in this screening process and call the exponent a the advantage of the attack [30]. On the other hand, the cryptanalyst must take care that the correct key k'_r is among the survived keys. Let us denote by P_S the probability that the correct key k'_r survives. Then P_S is called the success probability of the attack. The value a can also interpreted as the number of key bits correctly determined by the screening process and therefore a is usually taken as a positive integer. Then the remaining key bits are determined by exhaustive search and the correct solution is found with probability P_S. Clearly, there is a trade-off between a and P_S. But more importantly, the larger $|\mathrm{cor}(u, v)(K)|$ is, the larger values a and P_S can be achieved by increasing the sample size N. The relationship between these quantities is established using a statistical model. In this section we provide the detail in the linear attack context.

(Distinct)-known plaintext attack. The statistical model depends also on the way the cryptanalysts obtain the data sample. The pairs (x, y') can be obtained by receiving them randomly, in which case, the attack is called known-plaintext (KP) attack. This can be modeled as picking x randomly and obtaining y' for it. In statistical terms, we say that the sampling of (x, y') is done with replacement. As the sample size grows, sorting out repetitions becomes infeasible due to the memory and time requirements. Therefore, KP sampling is the most commonly used model for linear cryptanalysis. In this paper, we will also study sampling of distinct-known plaintext (DKP). In practice, it means that the cryptanalyst generates a set of non-repeating but otherwise random plaintext x and obtains the corresponding ciphertexts. In statistical terms, the pairs (x, y') are sampled randomly without replacement.

2.2 Statistical Distributions

To estimate the data complexity of a statistical attack, we study the distribution of the random variables involved in the attack.

Let us denote by $Z = Z(D, K, k_r, \kappa)$ the random variable corresponding to the number of solutions of the equation $u \cdot x + v \cdot G_\kappa^{-1}(y') = 0$, where $(x, y') \in D$. The n-bit block cipher with a fixed key K determines a probability p for a randomly selected x to satisfy this equation. It is well known [18,25] that in case of KP sampling, the variable Z follows a binomial distribution with expected value Np and variance $Np(1-p)$. In case of DKP sampling, by definition of the hypergeometric distribution, the variable Z follows a hypergeometric distribution with expected value Np but with variance

$$Np(1-p)\frac{2^n - N}{2^n - 1},$$

which goes to zero as the sample size grows.

In this paper, we often consider KP and DKP alternatives within the same model. To this end, we introduce the following constant B which is defined by

$$B = \begin{cases} 1, & \text{for KP,} \\ \dfrac{2^n - N}{2^n - 1}, & \text{for DKP.} \end{cases} \tag{1}$$

Both the binomial distribution and hypergeometric distribution allow tight approximation using the normal distribution [33]. It means that a discrete random variable Z with binomial or hypergeometric distribution, whose all values are integers, is related to a continuous normal deviate X such that

$$\Pr(Z = \zeta) \approx \Pr(\zeta - \frac{1}{2} \leq X < \zeta + \frac{1}{2}) \tag{2}$$

In particular, the expected values and variances of Z and X are equal. In this paper, when we say that a discrete random variable follows a normal (or some other continuous) distribution, it is in the sense given in Equation (2). We then denote by $Z \sim \mathcal{N}(\mu, \sigma^2)$ a random variable Z which follows a normal distribution with mean μ and variance σ^2.

In the rest of this paper, the empirical correlation is interpreted as a discrete random variable. Due to the connection

$$\hat{c}(D, K, k_r, \kappa) = \frac{2}{N} Z(D, K, k_r, \kappa) - 1$$

and the normal approximation of the distribution of $Z(D, K, k_r, \kappa)$, we say that $\hat{c}(D, K, k_r, \kappa)$ follows a normal distribution with expected value

$$\mathrm{Exp}_D(\hat{c}(D, K, k_r, \kappa)) = 2^{n-1} \#\{x \in \mathbb{F}_2^n \mid u \cdot x + v \cdot G_\kappa^{-1}(y') = 0\} - 1 = 2p - 1$$

and variance

$$\begin{aligned} \mathrm{Var}_D(\hat{c}(D, K, k_r, \kappa)) &= \mathrm{Exp}_D(\hat{c}(D, K, k_r, \kappa)^2) - (\mathrm{Exp}_D(\hat{c}^2(D, K, k_r, \kappa)))^2 \\ &= \frac{1}{N} 4p(1-p)B, \end{aligned} \tag{3}$$

where B is defined as in Equation (1).

It should be noted that the approximation of the binomial distribution by a normal distribution is commonly accepted and has been verified experimentally in [1, 12, 18]. For more detailed considerations about distributions of discrete random variables arising from differential and linear cryptanalysis we refer to [18].

Later in this paper, we will also consider discrete random variables that are formed as a sum of squares of independent binomial variables. Since a sum of squares of independent standard normal deviates follows χ^2 distribution, we will identify such a discrete random variable with a continuous χ^2 distributed variable, and say that this discrete variable follows χ^2 distribution.

The probability p depends on K, k_r and κ. The crucial distinction is made between the cases $\kappa = k'_r$ and $\kappa \neq k'_r$. Previously also the dependency of p on different values of κ for $\kappa \neq k'_r$ has been studied [12]. This paper is the first to present a statistical model of the behavior of p in the case $\kappa = k'_r$ as K varies in the classical linear setting.

2.3 Statistical Hypothesis Testing

Given two random variables T_W and T_R with respective cumulative distribution functions F_W and F_R consider a situation that we have a value Θ, called a threshold value, such that $F_W(\Theta) > F_R(\Theta)$. Having observed a value T we decide that T is drawn from the distribution of T_R if $T > \Theta$. If $T \leq \Theta$ we decide that T is drawn from the distribution of T_W. Then error probabilities, the false alarm ε_0 and the non-detection ε_1, are defined as

$$\varepsilon_0 = 1 - F_W(\Theta) \quad \text{and} \quad \varepsilon_1 = F_R(\Theta).$$

The condition $F_W(\Theta) > F_R(\Theta)$ guarantees that the probability that we reject T when it is drawn from T_R is less than the probability that we accept T when it is drawn from T_W. Then $\Theta = F_W^{-1}(1 - \varepsilon_0) = F_R^{-1}(\varepsilon_1)$ and

$$\varepsilon_1 = F_R(F_W^{-1}(1 - \varepsilon_0)). \tag{4}$$

In the context of statistical cryptanalysis the cumulative distribution functions depend on the size N of the sample in such a way that, as N grows, Equation (4) is satisfied with smaller error probabilities $1 - \varepsilon_0$ and ε_1. On the other hand, given one error probability, the cryptanalyst can find an appropriate sample size and the other error probability to satisfy Equation (4), and then can compute a threshold value for the test.

In this paper, we will compute examples of this principle for normally distributed test variables. Let us assume now that T_W and T_R are normal deviates with different means μ_W and μ_R. We consider w.l.o.g. the case $\mu_W < \mu_R$. Let us denote the standard deviations σ_W and σ_R, respectively. Then we then have $\varepsilon_0 = 1 - F_W(\Theta) = 1 - \Phi(\frac{\Theta - \mu_W}{\sigma_W})$ and $\varepsilon_1 = F_R(\Theta) = \Phi(\frac{\Theta - \mu_R}{\sigma_R})$, where Φ denotes the cumulative distribution function of the standard normal distribution. From the symmetry of central normal distribution we get that $\varepsilon_1 = 1 - \Phi(\frac{\mu_R - \Theta}{\sigma_R})$,

Let us denote by ζ_0 and ζ_1 the quantiles of the standard normal distribution corresponding to the probabilities $1-\varepsilon_0$ and $1-\varepsilon_1$. It means that $\Phi(\zeta_0) = 1-\varepsilon_0$ and $\Phi(\zeta_1) = 1-\varepsilon_1$, where Φ denote the cumulative distribution function of the standard normal distribution. Then we compute the threshold value $\Theta = \mu_R - \zeta_1\sigma_R = \mu_W + \zeta_0\sigma_W$ and by Equation (4) obtain

$$1-\varepsilon_1 = \Phi(\zeta_1) = \Phi\left(\frac{\mu_R - \mu_W - \sigma_W\Phi^{-1}(1-\epsilon_0)}{\sigma_R}\right), \tag{5}$$

which gives the success probability of T_R, that is, the probability that decision is correct when T is drawn from the distribution of T_R. Such a threshold can be found as soon as the standard deviations σ_W and σ_R are sufficiently small to satisfy

$$\mu_W + \sigma_W\zeta_0 \leq \mu_R - \sigma_R\zeta_1. \tag{6}$$

The data complexity N is determined as the least sample size to obtain this inequality.

3 Classical Linear Key-Recovery Attack

3.1 Matsui's Algorithm 2

A linear approximation (u, v) over E'_K consists of linear characteristics that are given as sequences $\tau = (\tau_0 = u, \tau_1, \ldots, \tau_{r-2}, \tau_{r-1} = v)$ and its correlation can be computed as a product of round-by-round correlation matrices [16]

$$c(u,v)(K) = \sum_{\tau}\prod_{i=1}^{r-1}\mathrm{cor}(\tau_{i-1}\cdot z + \tau_i\cdot g_{k_i}(z)), \tag{7}$$

where the sum is taken over all characteristics τ of the linear approximation (u, v).

The classical case of Matsui's Algorithm 2 relies on the assumption that there exists a single τ such that

$$c(u,v)(K) \approx \prod_{i=1}^{r-1}\mathrm{cor}(\tau_{i-1}\cdot z + \tau_i\cdot g_{k_i}(z))$$

for all keys K. Moreover, the original attack assumes that the block cipher is key-alternating, that is, the round function is of the form $g_k(z) = g(x\oplus k)$. Then a characteristic can be presented as follows

$$\prod_{i=1}^{r-1}\mathrm{cor}(\tau_{i-1}\cdot z + \tau_i\cdot g_{k_i}(z)) = (-1)^{\tau\cdot K}\rho_\tau,$$

where ρ_τ is independent of the key. Let us denote $|\rho_\tau| = c$.

Now we can formulate the assumptions about the statistical distributions of the empirical correlation for the wrong key and the right key. We restrict to the KP case for direct comparison with the previous treatments.

Let us denote by $K_W = (K, k_r, \kappa)$ the key parameters that are used in computing the empirical correlation $\hat{c}(D, K, K_r, \kappa)$ for the wrong key candidate $\kappa \neq k'_r$, and denote in this case the counter by $Z(D, K_W) = Z(D, K, K_r, \kappa)$ the empirical correlation by $\hat{c}(D, K_W) = \hat{c}(D, K, K_r, \kappa)$. Since the data is not obtained from the cipher, it is not expected to exhibit the bias of the linear approximation. Specifically, it is assumed that for a given key K_W, the counter $Z(D, K_W)$ is binomially distributed with $p = 1/2$, which leads to the following assumption about the continuous approximation of this probability distribution.

Wrong-key randomization hypothesis: For a wrong key candidate, $\hat{c}(D, K_W)$ follows normal distribution with parameters

$$\begin{aligned} \mathrm{Exp}_D\left(\hat{c}(D, K_W)\right) &= 0 \\ \mathrm{Var}_D\left(\hat{c}(D, K_W)\right) &= \frac{1}{N}. \end{aligned}$$

In a similar way, let us denote by $K_R = (K, k_r, \kappa)$ when $\kappa = k'_r$, and denote by $Z(D, K_R) = Z(D, K, K_r, \kappa)$ the counter and by $\hat{c}(D, K_R)$ the empirical correlation in this case. Note that following the notation of Section 2.1 we have

$$\hat{c}(D, K_R) = \frac{2}{N}\#\{x \in D \,|\, u \cdot x + v \cdot E'_K = 0\} - 1$$

which is independent of the value of k_r and $\kappa = k'_r$. The expected value of $\hat{c}(D, K_R)$ taken over data D is then the correlation $c(u, v)(K)$ of the linear approximation, which for each key K is assumed to be equal to c or $-c$ in this classical case of a single dominant characteristic. This leads to the following assumption.

Hypothesis of right-key equivalence: For all correct key $\kappa = k'_r$ the empirical correlation $\hat{c}(D, K_R)$ follows normal distribution with parameters

$$\begin{aligned} &\mathrm{Exp}_D\left(\hat{c}(D, K_R)\right) = \pm c \\ &\mathrm{Var}_D\left(\hat{c}(D, K_R)\right) = \mathrm{Exp}_D\left(\left(\hat{c}(D, K_R) - \mathrm{Exp}_D\left(\hat{c}(D, K_R)\right)\right)^2\right) = \frac{1}{N}(1 - c^2). \end{aligned}$$

Here it is usually estimated $1 - c^2 \approx 1$. Under these assumptions there are three normal distributions as depicted in Figure 3. In any practical instance with a fixed encryption key (K, k_r), only two of the distributions are present, the middle distribution and exactly one of the other two, but the cryptanalyst does not know which of the two. It implies that wrong keys will be accepted on both sides.

When testing the key candidates, the cryptanalyst is facing with the task of statistical hypothesis testing: given the value $|\hat{c}(D, K, k_r, \kappa)|$ computed from the data D, determine if the data is obtained from the cipher with the correct

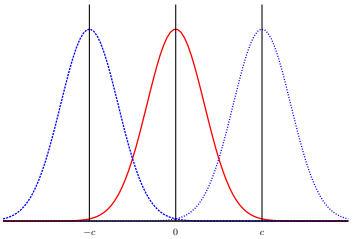


Fig. 3: Three normal distributions related to classical linear key-recovery attack:
middle curve = wrong key K_W with $\mathrm{Exp}_D\left(\hat{c}(D, K_W)\right) = 0$
left curve = right key K_R with $\mathrm{Exp}_D\left(\hat{c}(D, K_R)\right) = -c$
right curve = right key K_R with $\mathrm{Exp}_D\left(\hat{c}(D, K_R)\right) = c$

last-round key $\kappa = k'_r$, or if it is not from the cipher, and the key candidate κ is rejected.

Let us now apply the hypothesis testing paradigm explained in Section 2.3 to the key recovery attack. Let T_W be the observed correlation computed with the wrong key candidate and T_R the observed correlation with the right key. Then $\mu_W = 0$ and $\mu_R = c$, and $\sigma_W^2 = 1/N$ and $\sigma_R^2 = 1/N(1 - c^2)$.

Let us denote the error probabilities

$$\alpha_0 = 2^{-a} \text{ and } \alpha_1 = 1 - P_S,$$

where α_0 is the probability that a wrong key candidate is accepted and α_1 is the probability that the correct key is rejected. Since wrong keys can be accepted on both sides, the error probabilities for the test are

$$\varepsilon_0 = \frac{1}{2}\alpha_0 = 2^{-(a+1)} \text{ and } \varepsilon_1 = \alpha_1 = 1 - P_S.$$

By substituting these values to Equation (5) we can solve for data complexity bound N and threshold Θ such that $\Theta = \sqrt{1/N}\zeta_0 = c - \sqrt{(1-c^2)/N}\zeta_1$ and

$$P_S = \Phi(\zeta_1) = \Phi\left(\frac{c - \varphi_{a+1}\sqrt{1/N}}{\sqrt{1-c^2}}\right), \tag{8}$$

where we have denoted the quantiles as $\zeta_0 = \Phi^{-1}(1 - 2^{-(a+1)}) = \varphi_{a+1}$ and $\zeta_1 = \Phi^{-1}(P_S) = \varphi_{P_S}$.

The case $\mathrm{Exp}_D(\hat{c}(D, K_R)) < 0$ is a mirror image of the case explained above and $-\Theta$ can be taken as a threshold will be $-\Theta$. Then the key candidate κ is accepted if $\hat{c}(D, K, k_r, \kappa) > \Theta$ or $\hat{c}(D, K, k_r, \kappa) < -\Theta$ and rejected otherwise.

Let us summarize the derivations in the following theorem, which is a refinement of the original result of Matsui [25]. The same formula (assuming $1-c^2 \approx 1$) was derived in [30], Corollary 1, using an order statistic approach and a folded normal distribution for the right key.

Theorem 1. *Assume that an r-round block cipher has a linear approximation with a single dominant characteristic over $r-1$ rounds and correlation with absolute value about equal to c. Assume that the hypotheses of right-key equivalence and wrong-key randomization hold. Then the key-recovery attack presented in this section will succeed with probability P_S and advantage a if the size N of the available data sample satisfies*

$$N \geq \frac{(\varphi_{a+1} + \sqrt{1-c^2}\varphi_{P_S})^2}{c^2}.$$

3.2 Integrating Key Variable in the Model

Only recently, the hypothesis of right-key equivalence and the wrong-key randomization hypothesis have been questioned, as it has been observed in practical experiments that the statistical distributions may vary significantly as the key varies. The same holds for the wrong key case. For each wrong key candidate the statistical distribution of $\hat{c}(D, K_W)$ is different. Strong evidence was brought up that it is not accurate to model wrong keys to draw test statistic from the uniform distribution [12, 26].

The wrong key case. In [12] the distribution of the empirical correlation $\hat{c}(D, K_W)$ was examined in the case of a wrong key K_W. Specifically, it was noted that the empirical correlation depends on two mutually independent random variables K_W and D. Let $\kappa \neq k'_r$, and denote

$$\begin{aligned}\tilde{c}(K_W) = \tilde{c}(K, k_r, \kappa) &= 2^{n-1}\#\{x \in \mathbb{F}_2^n \mid u \cdot x + v \cdot G_\kappa^{-1}(y') = 0\} - 1 \\ &= \mathrm{Exp}_D(\hat{c}(D, K_W)).\end{aligned}$$

In other words, $\tilde{c}(K_W)$ is the correlation of the linear approximation (u, v) computed over the function $G_\kappa^{-1} \circ G_{k'_r} \circ E'_K$. The original wrong-key randomization hypothesis assumed that these correlations are equal for all wrong keys $K_W = (K, k_r, \kappa)$. Based on the remark after Corollary 4.3 of [18], in [12] they suggested to revise the wrong-key randomization hypothesis as follows.

Hypothesis 1 *Revised wrong-key randomization hypothesis: For each wrong key $K_W = (K, k_r, \kappa)$, the function $G_\kappa^{-1} \circ G_{k'_r} \circ E'_K$ is a random vectorial Boolean function and the correlation of its linear approximation has the following distribution*

$$\tilde{c}(K_W) \sim \mathcal{N}(0, 2^{-n}).$$

Based on this hypothesis the following result was stated in [12] but the proof was omitted. We give also the proof here.

Theorem 2. *Suppose that the revised wrong-key randomization hypothesis holds for an r-round block cipher. Then the empirical correlation $\hat{c}(D, K_W)$ is approximately normally distributed with parameters*

$$\mathrm{Exp}_{D,K_W}(\hat{c}(D, K_W)) = 0 \quad \textit{and} \quad \mathrm{Var}_{D,K_W}(\hat{c}(D, K_W)) = \frac{1}{N} + 2^{-n}. \tag{9}$$

Proof. The aim is to determine the distribution of $\hat{c}(D, K_W)$ when both D and K are simultaneously taken into consideration as random variables. We write $\hat{c}(D, K_W)$ as a sum of two random variables

$$(\hat{c}(D, K_W) - \tilde{c}(K_W)) + \tilde{c}(K_W), \tag{10}$$

where, for each fixed K_W,

$$\hat{c}(D, K_W) - \tilde{c}(K_W) \sim \mathcal{N}(0, \frac{1}{N}(1 - \tilde{c}(K_W)^2)).$$

We observe that $2^n \tilde{c}(K_W)^2 \sim \chi^2$ with one degree of freedom and has the mean 2^{-n} and variance 2^{1-2n}. Hence it is negligible and often omitted in similar derivations, see e.g. [30], by replacing the true variance of the first variable by a key-independent upper-bound $\frac{1}{N}$. Then we apply the revised wrong-key randomization hypothesis to the second part of Equation (10), which is independent of D, to obtain that $\hat{c}(D, K_W)$ can be expressed as a sum of two independent normally distributed variables, the first one $\hat{c}(D, K_W) - \tilde{c}(K_W) \sim \mathcal{N}(0, \frac{1}{N})$ depending on D only, and the second one $\tilde{c}(K_W) \sim \mathcal{N}(0, 2^{-n})$ depending on K_W. □

The right key case. Next we complete the statistical model of the classical linear attack by making the corresponding adjustment to the variance of the empirical correlation in the right key case. Let us start by recalling the Linear hull theorem for iterated block ciphers. The proof of this classical result was given in [28]. The special case of key-alternating cipher was considered and proven in [17]. It is interesting to note that the Linear hull theorem, stated for a general Boolean function, has found applications also in coding theory [15] and in the theory of Boolean complexity [24].

Theorem 3. *Let (u, v) be a linear approximation and denote $c(u, v)(K) = \mathrm{cor}(u \cdot x + v \cdot E'_K(x))$ and $c(u, \tau, v) = \mathrm{cor}(u \cdot x + \tau \cdot K + c \cdot E'_K(x))$. Then the average of $c(u, v)(K)^2$ taken over K is equal to the sum of $c(u, \tau, v)^2$ taken over τ.*

Note that $c(u, v)(K)$ is computed as the correlation over the space of the plaintext x with a fixed key K, while correlation $c(u, \tau, v)$ is computed over the plaintext x and the key K. The latter is also called the correlation of the characteristics τ. Let $|K|$ denote the length of K in bits. The quantity

$$2^{-|K|} \sum_K c(u, v)(K)^2 = \sum_\tau c(u, \tau, v)^2$$

where the sum on right side is taken over all characteristics τ of the linear approximation (u, v) is called the expected linear potential of (u, v) and denoted as $ELP(u, v)$ or just ELP if the linear approximation is clear from the context.

Matsui's Algorithm 2 assumes a single characteristic τ with a dominating correlation, which takes the form $(-1)^{\tau \cdot K} \rho_\tau$. Let us denote by $\mathcal{K}_0$ the keys for

which $\tau \cdot K = 0$ and by $\mathcal{K}_1$ the keys for which $\tau \cdot K = 1$. Then the assumption about dominating characteristic can be formalized as follows

$$\mathrm{Exp}_{K \in \mathcal{K}_0}(c(u,v)(K)) = \rho_\tau \text{ and } \mathrm{Exp}_{K \in \mathcal{K}_1} c(u,v)(K) = -\rho_\tau.$$

So the expected values of the correlations taken over the the two disjoint halfs of the keyspace are $\pm c$. Moreover, it is natural to assume that the variances of $c(u,v)(K)$ over the two disjoint halfs of the keyspace are equal, that is,

$$\mathrm{Exp}_{K \in \mathcal{K}_0}(c(u,v)(K)^2) - c^2 = \mathrm{Exp}_{K \in \mathcal{K}_1}(c(u,v)(K)^2) - c^2,$$

in which case

$$\begin{aligned}\mathrm{Exp}_{K \in \mathcal{K}_0}(c(u,v)(K)^2) &= \mathrm{Exp}_{K \in \mathcal{K}_1}(c(u,v)(K)^2) \\ &= \mathrm{Exp}_K(c(u,v)(K)^2) = ELP(u,v).\end{aligned}$$

It follows that

$$\mathrm{Var}_{K \in \mathcal{K}_0}(c(u,v)(K)) = \mathrm{Var}_{K \in \mathcal{K}_1}(c(u,v)(K)) = \sum_{t \neq \tau} c(u,t,v)^2 = ELP - c^2.$$

Let us state this property, which was derived under specific assumptions, as a following hypothesis.

Hypothesis 2 *Revised hypothesis of right-key equivalence: When the key K is taken as a random variable over the half space $\mathcal{K}_0$ or $\mathcal{K}_1$, the correlations are distributed as follows*

$$c(u,v)(K) \sim \mathcal{N}(\pm c, ELP - c^2),$$

where the mean is positive for one half space and negative for the other.

Theorem 4. *In the context of a linear key-recovery attack of an iterated block cipher described in this section, the empirical correlation $\hat{c}(D, K_R)$ approximately follows exactly one of the two normal distributions with parameters*

$$\mathrm{Exp}_{D,K}(\hat{c}(D,K_R)) = \pm c \quad \textit{and} \quad \mathrm{Var}_{D,K}(\hat{c}(D,K_R)) = \frac{1}{N} + ELP - c^2. \tag{11}$$

The choice between these distribution happens with probability equal to one half.

Proof. The proof is similar to the one in the wrong key case. Assuming that these two probability distributions can be approximated by a normal distribution, we can write $\hat{c}(D, K_R)$ as a sum of two normal deviates

$$(\hat{c}(D,K_R) - c(u,v)(K)) + c(u,v)(K)$$

For each fixed key K_R, the first term is a normal deviate of the random variable D with mean 0 and variance $\frac{1}{N}(1 - c(u,v)(K)^2)$. Replacing $\frac{1}{N}(1 - c(u,v)^2)$ by its close upper bound $\frac{1}{N}$, we obtain that $(\hat{c}(D,K_R) - c(u,v)(K)) \sim \mathcal{N}(0, \frac{1}{N})$ and is independent of K. The distribution of the second term can be approximated with exactly one of two normal distributions $\mathcal{N}(c, ELP - c^2)$ or $\mathcal{N}(-c, ELP - c^2)$ depending if the distribution is taken over $K \in \mathcal{K}_0$ or $K \in \mathcal{K}_1$. We obtain the result since we have a sum of two independent random variables. □

By substituting the adjusted variances to the formula of the success probability given in Equation (8) we get

$$P_S \approx \Phi\left(\frac{c\sqrt{N} - \varphi_{a+1}\sqrt{1+N2^{-n}}\varphi_{a+1}}{\sqrt{1+N(ELP-c^2)}}\right).$$

If $ELP = c^2$ for all encryption keys as assumed by the Hypothesis of right-key equivalence, then this formula is identical to Equation (6) in [12].

In reality, it would be more accurate to assume the value

$$ELP - c^2 = \sum_{t \neq \tau} c(u,\tau,v)^2$$

to be bounded from below by the variance of random noise which is equal to 2^{-n}. If the equality can be assumed to hold, that is, the linear approximation is composed of one linear characteristic and pure noise, then by taking $ELP = c^2 + 2^{-n}$ we obtain the following complexity bound for the KP sampling in Matsui's Algorithm 2

$$N \geq \frac{(\varphi_{a+1} + \varphi_{P_S})^2}{c^2 - 2^{-n}(\varphi_{a+1} + \varphi_{P_S})^2}.$$

This complexity bound is larger than the one given in Theorem 1 due to the variance over the key, and gives meaningful values for $c > 2^{-n/2}(\varphi_{a+1} + \varphi_{P_S})$.

3.3 Several Dominant Characteristics

If the number of dominant characteristics is small, the approach of single dominant characteristic discussed in the previous section can be applied and the key space will be divided into 2^d parts according to the expected value of the correlation. This approach was taken in [29] where it was shown that it is possible to distinguish between different values of correlations up to seven rounds of PRESENT. As the number of dominant characteristics grows, the correlation as expressed in Equation (7) will take several different values with non-negligible absolute value and it becomes unfeasible to distinguish between them. Moreover, this often means that for an increasing number of encryption keys the correlation will be equal, or close, to zero. Most modern ciphers have been designed like this to avoid linear cryptanalysis. Then the key space cannot be partitioned as in the classical Matsui's Algorithm 2, but instead, it is considered as a whole. Consequently, the expected value of the empirical correlation, now taken over all keys, is typically close to zero. Indeed, it is very likely that the average correlation is equal to zero. This situation is interpreted by Daemen and Rijmen [18] as follows: "The average correlation of a hull gives no indication about the complexity of a linear attack. Therefore, we only talk about the ELP of a hull." Next we elaborate what this means in practice.

Let us consider a linear approximation (u, v) and denote $\mathrm{Exp}_K(c(u,v)(K))$ by c. Then by definition of the ELP we have $\mathrm{Var}_K(c(u,v)(K)) = ELP - c^2$ that

we denote by σ^2. Moreover, we assume normal distribution as state it as the following.

Hypothesis 3 *Right-key randomization hypothesis: The correlations of a block cipher are random variables of key K and follow normal distribution*

$$c(u,v)(K) \sim \mathcal{N}(c, ELP - c^2).$$

We now state the following theorem. This theorem is the basic building block of the new statistical model of multiple linear cryptanalysis. Therefore we include both KP and DKP cases of data sampling in this theorem.

Theorem 5. *Assume that the Right-key randomization hypothesis holds. Then the empirical correlation $\hat{c}(D, K_R)$ of (u, v) is a normally distributed random variable of D and K and*

$$\mathrm{Exp}_{D,K}(\hat{c}(D,K_R)) = c \quad \textit{and} \quad \mathrm{Var}_{D,K}(\hat{c}(D,K_R)) = \frac{B}{N} + ELP - c^2, \tag{12}$$

where B is defined depending of the sampling in Equation (1).

Proof. The claim follows by splitting the random variable $\hat{c}(D, K_R)$ to two parts

$$(\hat{c}(D,K_R) - c(u,v)(K)) + c(u,v)(K).$$

Similarly, as in the case of a single dominant characteristic, the probability distribution of the first part can be approximated by a normal deviate which is independent of the key variable K. It has expected value equal to zero and variance equal to $\frac{B}{N}$. By assumption, the second part is a normal deviate of the random variable K. Then the probability distribution of $\hat{c}(D, K_R)$ can be approximated by a sum of two independent normal deviates which has the claimed parameters. □

The form of the probability distribution of $c(u, v)(K)$ has been seen normal for many practical ciphers [1]. One example is the cipher PRESENT [10], for which, in addition, the expected values c of correlations are practically equal to zero. In such a case the means of the empirical correlations in the wrong and right key cases are the same, that is, equal to zero. But the variances are different which makes distinguishing possible also in this case. By Theorem 2 the variance in the wrong key case is equal to 2^{-n} which is less than ELP. The approach is to use the square $\hat{c}(D, K, k_r, \kappa)^2$ as a test statistic, which in both cases is a constant multiple of a χ^2-distributed random variable where the constant is equal to $\frac{B}{N} + ELP$ in the right key case and $\frac{B}{N} + 2^{-n}$ in the wrong key case. Hence the test statistic has different means $\frac{B}{N} + ELP$ or $\frac{B}{N} + 2^{-n}$ and variances $2(\frac{B}{N} + ELP)^2$ or $2(\frac{B}{N} + 2^{-n})^2$ in the right key and wrong key case, respectively. Now the hypothesis testing approach of Section 2.3 can be applied analogically to the classical case by replacing the normal distribution by gamma distribution.

The larger ELP, the better is the distinguisher. If multiple linear approximations with large ELP are taken into account simultaneously, the difference

in variances can be amplified even further and the distinguisher improved. This approach was first modeled in [6] by assuming that the involved linear approximations are independent. This model covered the statistics as a random variable over the data sample and assumed that the behavior is practically the same for all keys. Similarly, the extension of the statistical model to multidimensional linear cryptanalysis presented in [19] did not take into account the effect of the key. The goal of the next section is to complete these models and include the key variable to them.

4 Multiple and Multidimensional Linear Attacks

4.1 Capacity

In [6] the statistical model of taking advantage of multiple independent linear approximations in key-recovery attacks was presented. In the more recent multidimensional linear attacks introduced in [19], the attacker exploits of all linear approximations with linear masks $(u, v) \neq 0$ in a linear space. The main benefit of the latter approach is that it does not require the assumption of independence of the linear approximations.

The main motivation and challenges of the work presented in this section originate from the multidimensional linear attack, but due to the generic link between linear and differential types of attacks [7], the results can also be applied to truncated differential attacks.

To collect information of the correlations of all the linear approximations over E'_K used in the attack, the notion of capacity was introduced in [6] and generalized in [19]. Given a set of input and output linear mask pairs (u_j, v_j), $j = 1, \ldots, \ell$, where $(u_j, v_j) \neq 0$, their capacity is defined as the sum of the squared correlations:

$$C(K) = \sum_{j=1}^{\ell} c(u_j, v_j)(K)^2. \tag{13}$$

In case the linear approximations (u_j, v_j) form the set of non-zero elements of a linear space $U \times V$ of dimension s, that is, $\ell + 1 = 2^s$, then the capacity can also be computed as

$$C(K) = 2^s \sum_{\eta=0}^{\ell} \left(p_\eta(K) - 2^{-s}\right)^2, \tag{14}$$

where $p_\eta(K)$ is the probability that a data value $(x, E'_K(x))$ restricted to $U \times V$ takes the value $\eta \in U \times V$. In other words, $C(K)$ is the squared Euclidean imbalance [3] of the probability distribution $p_\eta(K)$.

While multiple and multidimensional linear attacks take advantage of a set of linear approximations with large capacity, multiple and multidimensional zero-correlation linear attacks [9, 11, 13] exploit linear approximations with correlation equal to zero. These attacks have been proven efficient on word-oriented

structures such as Feistel-type ciphers. When multiple approximations with zero-correlation are used, the capacity $C(K)$ of the set of linear approximations is equal to zero for all keys K.

To capture the statistical behavior of capacities we develop a model that takes also the key variance into account. The case of wrong keys is straightforward as all empirical correlations are independently and identically distributed as D and K_W varies. The right key case requires some additional assumptions. In [20], the cipher correlations were assumed independent and certain weak key quantiles where determined and tested for multiple linear key recovery attack on the PRESENT cipher.

4.2 Key-Recovery Attack

We consider the same last-round key-recovery setting as in Section 2. In (zero-correlation) multiple/multidimensional linear key-recovery attacks the test statistic is the empirical capacity of a set of linear approximations. We denote it by $T(D, K, k_r, \kappa)$ and it is computed as follows

$$T = T(D, K, k_r, \kappa) = N \sum_{j=1}^{\ell} \hat{c}_j(D, K, k_r, \kappa)^2, \tag{15}$$

where $\hat{c}_j(D, K, k_r, \kappa)$ is the empirical correlation of the j-th linear approximation $u_j \cdot x + v_j \cdot G_\kappa^{-1}(E_{K'}(x))$.

In multidimensional linear key-recovery attacks, the online test statistic is computed over all non-zero linear approximations, in which case, instead of the individual empirical correlations, cryptanalyst may compute the test statistic over the observed data $(x, G_\kappa^{-1}(y'))$, $x \in U$ and $G_\kappa^{-1}(y') \in V$, also as follows

$$T = T(D, K, k_r, \kappa) = \sum_{\eta=0}^{\ell} \frac{(V[\eta] - N2^{-s})^2}{N2^{-s}}, \tag{16}$$

where $V[\eta]$ corresponds to the number of occurrences of the value η of the observed data distribution. In the offline analysis, only a subset of all linear approximations in a linear space $U \times V$ are taken into consideration when the capacity estimate is computed.

In the following, we study the statistical distributions of multiple and multidimensional linear cryptanalysis as the encryption key and the wrong key candidates vary and provide new key-dependent distribution parameters. With these new developments we aim at improving the estimate of the data complexity. One of the main obstacles that still remains is how to estimate the variance of $T(D, K, k_r, \kappa)$ for $\kappa = k'_r$. We will also present some approaches how to tackle this problem.

4.3 Statistical Model for the Wrong Keys

The wrong key behavior is modeled according to the statistics of random cipher as in [17] under the assumption that the revised wrong-key hypothesis stated

in Section 3.2 holds. As shown in Theorem 2 and the remark after it, the empirical correlation $\hat{c}(D, K_W)$, $K_W = (K, k_r, \kappa)$, where $\kappa \neq k'_r$, is approximately normally distributed with parameters

$$\mathrm{Exp}_{D,K_W}(\hat{c}(D, K_W)) = 0 \quad \text{and} \quad \mathrm{Var}_{D,K_W}(\hat{c}(D, K_W)) = \frac{1}{N}B + 2^{-n},$$

where the constant B is determined as in Equation (1) according to KP or DKP sampling.

Let us denote the number of exploited linear approximations by ℓ. Then we can write the test statistic computed for wrong key K_W in multiple or multidimensional linear key recovery attack as follows

$$T = T(D, K_W) = N \sum_{j=1}^{\ell} \hat{c}_j(D, K_W)^2 = (B + 2^{-n}N) \sum_{j=1}^{\ell} \frac{\hat{c}_j(D, K_W)^2}{\frac{1}{N}B + 2^{-n}}.$$

Note that $\dfrac{\hat{c}_j(D, K_W)}{\sqrt{\frac{1}{N}B + 2^{-n}}} \sim \mathcal{N}(0, 1)$. As for a random function, the linear approximations are independent, it follows that

$$\frac{T}{B + 2^{-n}N} \sim \chi^2_\ell \tag{17}$$

and the parameters of the probability distribution of T are as given by the following theorem.

Theorem 6. *Assuming that the revised wrong-key hypothesis holds for ℓ linear approximations involved in a multiple or multidimensional linear attack, then the statistic $T(D, K_W)$ computed as in Equation (15) (or Equation (16)) is a constant multiple of a χ^2-distributed variable with ℓ degrees of freedom and has the following mean and variance*

$$\begin{aligned} \mathrm{Exp}_{D,K_W}(T(D, K_W)) &= B\ell + N2^{-n}\ell \text{ and} \\ \mathrm{Var}_{D,K_W}(T(D, K_W)) &= \frac{2}{\ell}\left(B\ell + N2^{-n}\ell\right)^2, \end{aligned} \tag{18}$$

where B is defined as in Equation (1).

Proof. From Equation (17) and by definition of the χ^2 distribution, we have

$$\mathrm{Exp}_{D,K_W}\left(\frac{T}{B + 2^{-n}N}\right) = \ell$$

and

$$\mathrm{Var}_{D,K_W}\left(\frac{T}{B + 2^{-n}N}\right) = 2\ell.$$

□

We highlight the following special case and state it as a separate corollary.

Corollary 1. *In the context of Theorem 6 suppose that sampling is done using distinct known plaintexts. Then the statistic $T(D, K_W)$ has the following mean and variance*

$$\begin{aligned}\mathrm{Exp}_{D,K_W}\left(T(D,K_W)\right) &\approx \ell\\ \mathrm{Var}_{D,K_W}\left(T(D,K_W)\right) &\approx 2\ell.\end{aligned}$$

Proof. By substituting $B = (2^n - N)/(2^n - 1)$ to Equation (18) we get the result. □

Interestingly, these are exactly the parameters that have been used in previous works to model the wrong-key distribution for DKP sampling in multidimensional zero-correlation attacks in [9, 11]. However, no justification of these parameter values can be found in the previous literature. As KP sampling from a uniform distribution will yield the same distribution parameters for $T(D, K_W)$, it is possible that those parameters have been reused in DKP case in the lack of anything better. Fortunately, the parameter values were correct and the existing zero-correlation attacks that use DKP remain correct.

The situation is not that fortunate for general multidimensional linear attacks that use KP. The data complexity estimate as given in Equation (21) has been derived under the hypothesis that the wrong-key data is drawn from the uniform distribution with $\mathrm{Exp}_{D,K_W}\left(T(D,K_W)\right) = \ell$ and $\mathrm{Var}_{D,K_W}\left(T(D,K_W)\right) = 2\ell$, see Equation (17) of [19]. This is certainly too optimistic for the attacker, since the data distributions in multidimensional linear approximation in random case are unlikely to become completely uniform. The more realistic values of parameters in the KP case are $\mathrm{Exp}_{D,K_W}\left(T(D,K_W)\right) = \ell(1 + N2^{-n})$ and $\mathrm{Var}_{D,K_W}\left(T(D,K_W)\right) = 2\ell(1 + N2^{-n})^2$ as given by our analysis in Theorem 6.

Remark 1. We denote by $C(K_W) = \sum_{j=1}^{\ell} \tilde{c}_j(K_W)^2$ with $\tilde{c}_j(K_W) = \tilde{c}(K, k_r, \kappa) = 2^{n-1}\#\{x \in \mathbb{F}_2^n \mid u_j \cdot x + v_j \cdot G_\kappa^{-1}(y') = 0\} - 1$ the capacity for a given wrong key. Assuming the Revised wrong-key randomization hypothesis, we have $c_j(K_W) \sim \mathcal{N}(0, 2^{-n})$ and by definition, a constant multiplier of $C(K_W)$ follows a χ^2 distribution with ℓ degrees of freedom. It means that $C(K_W)$ follows a gamma distribution with $\mathrm{Exp}_{K_W}(C(K_W)) = \ell/2^n$ and $\mathrm{Var}_{K_W}(C(K_W)) = 2^{1-2n}\ell$.

4.4 Modeling the Right Key Behavior

In Section 3.3, we gave the distribution of correlation of a single linear approximation (see Theorem 5). In this section we model the behavior of the statistic T involved in a multiple/multidimensional linear attack for the right key using the same straightforward approach as in the wrong-key case presented above. The result is obtained by combining the squares of normally distributed correlations into a (constant multiple of) a χ^2 distributed statistic. This is basically how the distributions of $T(D, K_R)$, for the right key $K_R = (K, k_r, \kappa)$, where $\kappa = k'_r$, were determined in [20]. It takes the following assumption to do this:

Hypothesis 4 *(Key-variance hypothesis – multiple:) The empirical correlations* $\hat{c}_j(D, K_R)$, $j = 1, 2, \cdots, \ell$ *of the multiple linear approximations involved in Equation (15) are statistically independent and their expected values* $\mathrm{cor}(u_j, v_j)(K)$ *taken over the data are normal deviates of* K *and have equal variances.*

For the sake of clarity, let us introduce the following notations before presenting the result. Given $1 \leq j \leq \ell$, the correlations of the ℓ linear approximations are, for a fixed key K, denoted by $c_j(K) = \mathrm{cor}(u_j, v_j)(K) = \mathrm{Exp}_D(\hat{c}_j(D, K_R))$ and $c_j = \mathrm{Exp}_K(c_j(K))$. Given the expected linear potentials $ELP_j = ELP(u_j, v_j)$ as defined in Section 3.2, the capacity is

$$C = \mathrm{Exp}_K(C(K)) = \sum_{j=1}^{\ell} \mathrm{Exp}_K(c_j(K)^2) = \sum_{j=1}^{\ell} ELP_j.$$

Since in general $\mathrm{Exp}_K(c_j(K)^2) \neq \mathrm{Exp}_K(c_j(K))^2$, we also introduce $C_0 = \sum_{j=1}^{\ell} c_j^2$ which can be interpreted as the capacity of the expected correlations. We then state the following result.

Theorem 7. *Suppose that the linear approximations involved in the computation of the attack statistic as in Equation (15)*

$$T = T(D, K_R) = \sum_{j=1}^{\ell} \hat{c}_j(D, K_R)^2$$

satisfy Hypotheses 2 and 4. We have

$$Q = \frac{T}{B + \frac{N}{\ell}(C - C_0)} \sim \chi^2_\ell(\delta),$$

where the non-centrality parameter of the χ^2 *distribution is*

$$\delta = \frac{NC_0}{B + \frac{N}{\ell}(C - C_0)},$$

and B *is defined as in Equation (1).*

Proof. By Hypotheses 2 and 4 we can apply Theorem 5 for each $j = 1, 2, \ldots, \ell$ and get that each $\hat{c}_j(D, K_R)$ follows normal distribution with parameters

$$\mathrm{Exp}_{D,K}\, \hat{c}_j(D, K_R) = c_j \text{ and } \mathrm{Var}_{D,K}\, \hat{c}_j(D, K_R) = \frac{B}{N} + ELP_j - c_j^2,$$

and moreover, they are all independent.

With the notations previously defined, the summation over j gives

$$\sum_{j=1}^{\ell} \mathrm{Var}_{D,K}(\hat{c}_j(D, K_R)) = \frac{B}{N}\ell + C - C_0.$$

It then follows by Hypothesis 4 that

$$\mathrm{Var}_{D,K}(\hat{c}_j(D, K_R)) = \frac{B}{N} + \frac{C - C_0}{\ell}.$$

Moreover, by Hypothesis 4 the empirical correlations are normal deviates, meaning that for all j,

$$\frac{\hat{c}_j(D, K_R)}{\sqrt{\frac{B}{N} + \frac{C-C_0}{\ell}}} \sim \mathcal{N}\left(\frac{c_j}{\sqrt{\frac{B}{N} + \frac{C-C_0}{\ell}}}, 1\right).$$

By definition of the non-central χ^2 distribution we obtain that

$$Q = \frac{T}{B + \frac{N}{\ell}(C - C_0)} = \sum_{j=1}^{\ell} \left(\frac{\hat{c}_j(D, K_R)}{\sqrt{\frac{B}{N} + \frac{C-C_0}{\ell}}}\right)^2$$

follows a non-central χ^2 distribution with ℓ degrees of freedom and non-central parameter

$$\delta = \sum_{j=1}^{\ell} \frac{c_j}{\sqrt{\frac{B}{N} + \frac{C-C_0}{\ell}}}.$$

□

Let us recall that in the multidimensional case, when the key K is fixed and the statistic T is computed by Equation (16) over 2^s data values η, the empirical frequencies $V[\eta]$ are independent random variables of D for all but one η, for which $V[\eta]$ is determined by the other values. Therefore, we only need an assumption about their expected values $p_\eta(K)$ to prove Theorem 7 in this case. This assumption is formulated as follows, where we recall that $\sum_{\eta=0}^{\ell} p_\eta = 1$ and therefore, for each K, only ℓ of the values $p_\eta(K)$ can be taken as free variables.

Hypothesis 5 *(Key-variance hypothesis – multidimensional:) For each fixed data value $\eta = 0, \ldots, 2^s - 1$, the probabilities $p_\eta(K)$ taken as random variables of K, are normal deviates with equal variances. Moreover, each subset of ℓ values $p_\eta(K)$ are independent, and they determine the remaining value uniquely.*

Under the previous hypothesis we obtain the following result which is similar to the one of Theorem 7.

Theorem 8. *Suppose that the expected values $p_\eta(K) = \mathrm{Exp}_D(V[\eta]/N)$ involved in the computation of the attack statistic as in Equation (16)*

$$T = T(D, K, k_r, \kappa) = \sum_{\eta=0}^{\ell} \frac{(V[\eta] - N2^{-s})^2}{N2^{-s}},$$

satisfy Hypothesis 5. Then

$$Q = \frac{T}{B + \frac{N}{\ell}(C - C_0)} \sim \chi^2_\ell(\delta),$$

where the non-centrality parameter of the χ^2 distribution is

$$\delta = \frac{NC_0}{B + \frac{N}{\ell}(C - C_0)},$$

and B is defined as in Equation (1).

Proof. Let us start by computing the sum of the variances of $p_\eta(K)$ over $\eta = 0, 1, 2, \ldots, \ell = 2^s - 1$. We obtain

$$\begin{aligned}\sum_{\eta=0}^{\ell} \mathrm{Var}_K\, p_\eta(K) &= \sum_{\eta=0}^{\ell} \left(\mathrm{Exp}_K(p_\eta(K)^2) - p_\eta^2\right) \\ &= 2^{-s}(C+1) - 2^{-s}(C_0+1) = \frac{C - C_0}{\ell + 1}.\end{aligned}$$

Without loss of generality, we denote by $p_0(K)$ the probability that is determined by the other probabilities. $p_\eta(K)$, $\eta \neq 0$. Then by Hypothesis 5

$$p_\eta(K) \sim \mathcal{N}\left(p_\eta, \frac{C - C_0}{\ell(\ell+1)}\right), \eta \neq 0.$$

It follows that

$$Np_\eta(K) \sim \mathcal{N}\left(Np_\eta, N^2\frac{C - C_0}{\ell(\ell+1)}\right), \eta \neq 0.$$

On the other hand, for each fixed K, the frequency $V[\eta]$, $\eta \neq 0$ follows a binomial distribution with probability $p_\eta(K)$. We estimate $p_\eta(K)(1 - p_\eta(K)) \approx 2^{-s}$ as usual. Then

$$V[\eta] - Np_\eta(K) \sim \mathcal{N}(0, BN2^{-s}),\ \eta \neq 0,$$

and this distribution is independent of K, and hence, independent of $Np_\eta(K)$. It follows that

$$V[\eta] \sim \mathcal{N}\left(Np_\eta, \frac{N}{\ell+1}(B + \frac{N}{\ell}(C - C_0)\right).$$

The rest of the proof is analogical to the proof of Theorem 7. □

The distribution of $T = T(D, K_R)$. Summarizing the results of Theorems 7 and 8, the test statistic T has the following parameters

$$\begin{aligned}\mathrm{Exp}_{D,K}(T(D, K_R)) &= (\ell + \delta)(B + \frac{N}{\ell}(C - C_0) = B\ell + NC \\ \mathrm{Var}_{D,K}(T(D, K_R)) &= 2(\ell + 2\delta)(B + \frac{N}{\ell}(C - C_0)^2 \\ &= \frac{2}{\ell}((\ell+\delta)^2 - \delta^2))(B + \frac{N}{\ell}(C - C_0)^2 \\ &= \frac{2}{\ell}((B\ell + NC)^2 - (NC_0)^2).\end{aligned} \tag{19}$$

The form of the distribution of T can be determined in two cases:

1. $\ell > 50$, in which case normal approximation can be used, or
2. $C_0 = 0$, in which case T follows a gamma distribution

$$T(D, K_R) \sim \Gamma\left(\frac{\ell}{2}, 2(B + N\frac{C}{\ell})\right).$$

Both of these cases are important for applications.

In the next section, we present results of our experiments when $\ell > 50$ and $C_0 = 0$. In particular we compare the theoretical estimates of the mean and the variance from Equation (19) with the experimental ones.

In the experiments we use versions of the SMALLPRESENT cipher where we can compute the correlations and capacities for the full codebook. In practical applications, however, the problem of how to obtain accurate estimates of the expected values and variances of the test statistics which are needed to estimate the data complexity of the key recovery attack. This question, as well as testing the accuracy of Hypotheses 4 and 5 require cipher specific information. We have investigated the case of the key-alternating block cipher and will present the results of this work in a forthcoming paper.

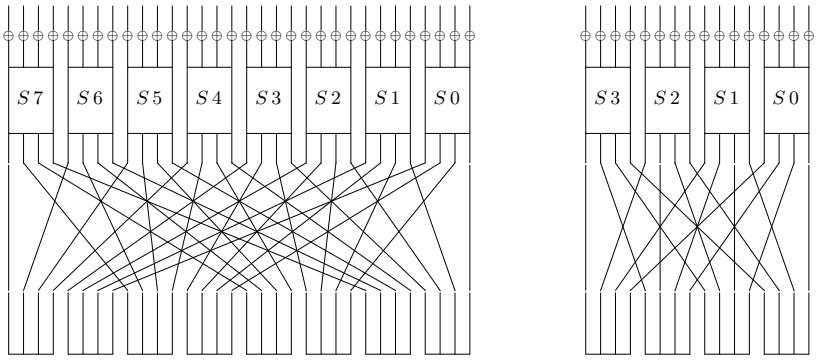


Fig. 4: The round function of SMALLPRESENT-[8] (left) and SMALLPRESENT-[4] (right).

4.5 Experiments on SMALLPRESENT

In this section, we verify developed theory by comparing the experimental and theoretical mean $\mathrm{Exp}_{D,K}(T(D, K_R))$ of the test statistic T in the cases of DKP and KP sampling.

The experiments have been conducted on two scale versions [22] of the block cipher PRESENT [10]. SMALLPRESENT-[8] is a 32-bit cipher designed with the 80-bit original key-schedule of PRESENT. SMALLPRESENT-[4] is a 16-bit cipher. The round functions of both ciphers are depicted in Figure 4. For the experiments on SMALLPRESENT-[4], a 20-bit key-schedule has been defined. The multidimensional distributions are respectively involving $\ell = 255$ and $\ell = 63$ linear approximations.

In all cases the capacity of the multidimensional approximation used in the theoretical models is the true value determined from the cipher.

In Figures 5 and 6, we compare the theoretical means given by Equation (19) of statistic T with the experimental ones for both distinct and non-distinct plaintext. For this cipher, the values seem to match very well.

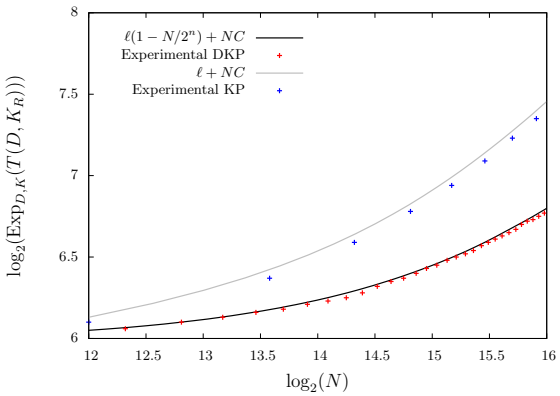


Fig. 5: The mean $\mathrm{Exp}_{D,K}(T(D, K_R))$ for a 6-bit multidimensional distribution ($\ell = 2^6 - 1$) over 4 rounds of SMALLPRESENT-[4] with capacity $C = 2^{-9.20}$.

In Figures 7 and 8, the corresponding variances are analyzed. We observe that the theoretical value is significantly improved when the key variance is taken into account. Still there is in all cases a clear gap between the theoretical value $\mathrm{Var}_{D,K}(T(D, K_R))$ given by Equation (19) and the experimental values of the variance.

In the computation of the theoretical value ℓ is taken equal to $2^s - 1$ where s is the dimension of the multidimensional linear approximation. It means that the model relies on the multidimensional Hypothesis 5, where moreover, we assume that $p_\eta = 0$, that is, $C_0 = 0$. We checked the validity of these assumptions, and

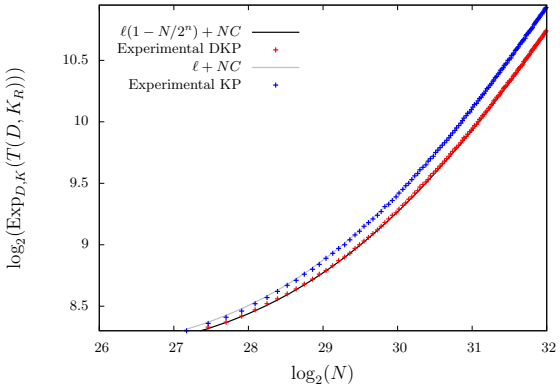


Fig. 6: The mean $\mathrm{Exp}_{D,K}(T(D,K_R))$ for a 8-bit multidimensional distribution ($\ell = 2^8 - 1$) over 9 rounds of SMALLPRESENT-[8] with capacity $C = 2^{-21.29}$.

only small deviation from it was observed in simulations on these SMALLPRESENT variants. On the other hand, it is known that due to the linear properties of the S-box, PRESENT ciphers allow accurate estimation of the capacity using single-bit linear characteristics that can be considered statistically independent. Therefore also the alternative approach of multiple independent linear approximations, that is, the use of Hypothesis 4 would be justified.

Let us examine these alternative approaches in the case of SMALLPRESENT-[8]. The observed multidimensional linear approximation consists of 4 bits of input to one S-box and 4 bits output of one S-box after 9 rounds. If Hypothesis 5 is applied, we take $\ell = 2^8 - 1$. By this approach we get an underestimate of the variance of $T(D, K_R)$ which is depicted in Figure 8. An alternative approach could be to include only the most dominant linear characteristics between these S-boxes, that is, by taking all single-bit characteristics leading from the three leftmost bits from the output of first S-box to the three leftmost bits of the input to the last S-box. This approach, however, seems in our experiments to give an overestimate of the variance of $T(D, K_R)$. The true value lies between these two extremes. Closing this gap and obtaining more accurate estimate of the variance is left for future work.

5 Data Complexity

5.1 Previous Model

The multidimensional linear cryptanalysis [19] traditionally assumes known plaintext and that the cryptanalyst does not have any means to check for repetitions in the plaintext. Then the statistic T given in Equation (16) computed from the multidimensional distribution for a fixed key (K, k_r, κ) follows, up to a constant multiplier, a χ^2 distribution both for $\kappa = k'_r$ and $\kappa \neq k'_r$. The parameters in the

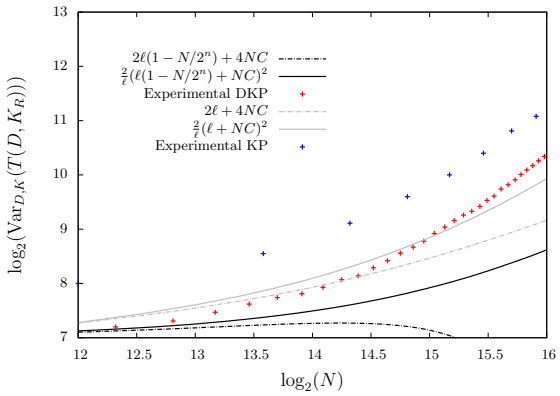


Fig. 7: The variance $\mathrm{Var}_{D,K}(T(D,K_R))$ for a 6-bit multidimensional distribution ($\ell = 2^6-1$) over 4 rounds of SMALLPRESENT-[4] with capacity $C = 2^{-9.20}$.

KP model given in [19] were the following:

$$\begin{array}{lll} \mathrm{Exp}(T(D,K_R)) \approx \ell + N\cdot C & \text{and} & \mathrm{Var}(T(D,K_R)) \approx 2(\ell + 2\cdot N\cdot C), \\ \mathrm{Exp}(T(D,K_W)) \approx \ell & \text{and} & \mathrm{Var}(T(D,K_W)) \approx 2\ell. \end{array} \tag{20}$$

The distributions for wrong and right key candidates have different means and variances and thus it is possible to distinguish between them by statistical inference analogical to the hypothesis testing method described in Section 3.1 but this time between two χ^2 distributed random variables. Then the error probabilities and the corresponding quantiles by replacing the cumulative density function Φ by the cumulative density function of the χ^2 distribution with ℓ degrees of freedom.

In practice, it is accurate to approximate χ^2 distributions using normal distribution as soon as $\ell = 2^s - 1 > 50$. Using the parameters given by Equation (20), and by denoting the advantage and success probability of the key-recovery attack respectively by a and P_S, the following estimate of the data complexity of a KP multidimensional linear attack was given in [19]

$$N \approx \frac{\sqrt{4a\ell} + 4\Phi^{-1}(2P_S-1)^2}{C}. \tag{21}$$

To derive this result in [19] it is assumed that the capacity $C(K)$ of the cipher data distribution, for any key K, is equal to its mean C, and the same value C is used for all encryption keys. Moreover, it is often difficult to obtain an accurate estimate of C. Usually only a rough lower bound can be obtained in the offline analysis of the cipher. Secondly, it is assumed that, for any wrong key candidate $\kappa \neq k'_r$, the data is drawn from the uniform distribution. In [12, 26] this simple approach has been criticized and shown to produce too optimistic (for the attacker) results in practice.

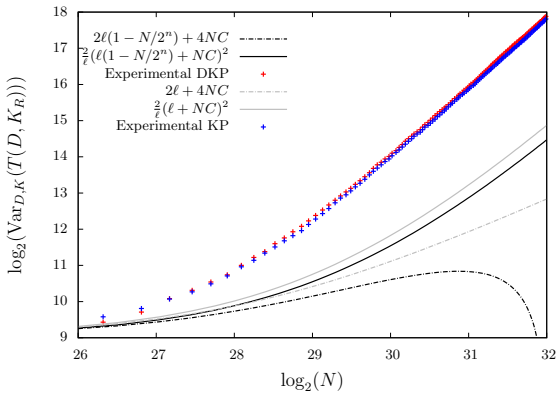


Fig. 8: The variance $\mathrm{Var}_{D,K}(T(D,K_R))$ for a 8-bit multidimensional distribution ($\ell = 2^8 - 1$) over 9 rounds of SMALLPRESENT-[8] with capacity $C = 2^{-21.29}$.

5.2 Data Complexity Estimates

In the hypothesis testing context described in Section 2.3, we can estimate the data complexity of a multiple/multidimensional linear attack. In this section, since we do not yet know how to estimate the value of C_0, we assume that $C_0 = 0$. As explained at the end of Section 4.4, this means that the random variable $T(D,K_R)$ follows a gamma distribution. Using Equation (4) we could therefore obtain an accurate formula of the success probability of the attack. However in order to compare with previous works we assume that $\ell > 50$, which is the case in most attacks, and use normal approximations of the gamma distributions of $T(D,K_W)$ and $T(D,K_R)$. We denote by $C_R = \mathrm{Exp}_K(C(K))$ and $C_W = \mathrm{Exp}_{K_W}(C(K_W))$ the expected values of the capacity for respectively the right and wrong keys.

Corollary 2. *Assuming that* $\mathrm{Var}_{D,K}(T(D,K_R)) = \frac{2}{\ell}(B\ell + NC_R)^2$ *and a normal approximation of the gamma distribution we obtain that the data complexity estimates* N^{KP} *and* N^{DKP} *in respectively the non-distinct and distinct context are given by the following formulas.*

$$N^{\mathrm{KP}} \approx \frac{\sqrt{2\ell}(\varphi_{P_S} + \varphi_a)}{|C_R - C_W| - \sqrt{2/\ell}(C_W\varphi_a + C_R\varphi_{P_S})}. \tag{22}$$

$$N^{\mathrm{DKP}} \approx \frac{\sqrt{2\ell}(\varphi_{P_S} + \varphi_a)}{|C_R - C_W| - \sqrt{2/\ell}(C_W\varphi_a + C_R\varphi_{P_S}) + 2^{-n}\sqrt{2\ell}(\varphi_{P_S} + \varphi_a)}. \tag{23}$$

Proof. According to Equation (5) and Theorems 6 and 7 (or 8), using the notations $\mu_R = \mathrm{Exp}_{D,K}(T(D,K_R))$, $\sigma_R^2 = \mathrm{Var}_{D,K}(T(D,K_R))$, we have

$$P_S \approx \Phi\left(\frac{N|C_R - C_W| - \sqrt{2/\ell}(B\ell + NC_W)\varphi_a}{\sqrt{2/\ell}(B\ell + NC_R)}\right). \tag{24}$$

We then deduce that

$$\sqrt{2/\ell}(B\ell + NC_R)\varphi_{P_S} \approx N|C_R - C_W| - \sqrt{2/\ell}(B\ell + NC_W)\varphi_a$$

and that

$$N\left(|C_R - C_W| - \sqrt{2/\ell}(C_W\varphi_a + C_R\varphi_{P_S})\right) \approx \sqrt{2\ell}B(\varphi_{P_S} + \varphi_a).$$

When the sampling is with replacement then $B = 1$ and we obtain the result. If we consider distinct plaintexts then $B \approx 1 - \dfrac{N}{2^n}$ and

$$N\left(|C_R - C_W| - \sqrt{2/\ell}(C_W\varphi_a + C_R\varphi_{P_S}) + 2^{-n}\sqrt{2\ell}(\varphi_{P_S} + \varphi_a)\right) \approx \sqrt{2\ell}(\varphi_{P_S} + \varphi_a).\square$$

In the following to compare Equations (22) and (23) we denote by $\lambda \geq 0$ the quantity defined by $C_R = \lambda \cdot C_W$, $\lambda \geq 0$. From Remark 1 we have $C_W = \ell/2^n$ and Equations (22) and (23) become

$$\begin{aligned} N^{\mathrm{KP}} &\approx \frac{2^n(\varphi_a + \varphi_{P_S})}{|\lambda - 1|\sqrt{\ell/2} - (\varphi_a + \lambda\varphi_{P_S})}, \\ N^{\mathrm{DKP}} &\approx \frac{2^n(\varphi_a + \varphi_{P_S})}{|\lambda - 1|\sqrt{\ell/2} - (\lambda - 1)\varphi_{P_S}}. \end{aligned} \tag{25}$$

In the zero-correlation context we have $\lambda = 0$ and we obtain the results recalled in Section 6 for respectively the non-distinct and distinct sampling methods.

Remark 2. For practical attacks we have $P_S \geq 0.5$ and $a \geq 1$ meaning that φ_{P_S} and φ_a are positive values, and that $(\lambda - 1)\varphi_{P_S} \leq \varphi_a + \lambda\varphi_{P_S}$. Therefore we can verify that $N^{\mathrm{DKP}} \leq N^{\mathrm{KP}}$ as expected.

5.3 Experiments on SMALLPRESENT-[4]

To perform a meaningful key-recovery attack, we simulated an attack on the 16-bit reduced version of PRESENT. For this attack we selected a multidimensional linear approximation of size $\ell = 2^6 - 1$ over 4 rounds. The key-recovery attack was on 6 rounds meaning that 2 rounds were partially inverted. In Figure 9, we give the results of the experiments.

When repetition of plaintexts is allowed, our model provides an underestimate of the success probability at least up to the data complexity corresponding to the full codebook. In that case, it seems that it is also possible to have a data complexity larger than the full codebook. Through these graphics we also illustrated that for a same advantage and data complexity the success probability is larger when using distinct plaintexts.

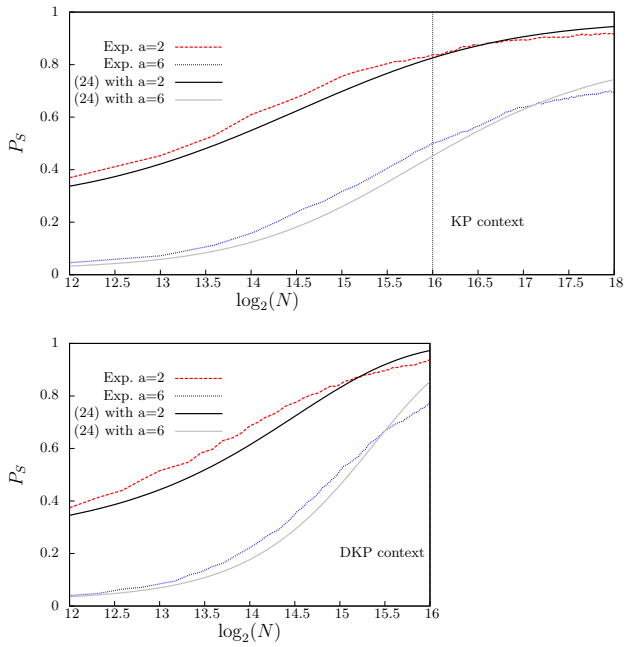


Fig. 9: Success probability of a key-recovery attack. The theoretical success probability is computed from Equation (24) using the normal distribution . The experimental results (Exp.) are represented with dotted lines. The parameters are $n = 16$, $\ell = 2^6 - 1$, $C_R = 2^{-9.20}$, $C_W = 2^{-10}$. Top: Using non-distinct plaintexts ($B = 1$). Bottom: Using distinct plaintexts ($B = 1 - N/2^{16}$).

6 Zero-Correlation Linear Cryptanalysis

6.1 Multiple and Multidimensional Zero-Correlation Linear Attacks

Zero-correlation linear cryptanalysis is a special case of multiple/multidimensional linear cryptanalysis with $C_R = 0$ and $C_W = \ell 2^{-n}$. Applying Corollary 2 we obtain the following estimate of the data complexity of a multiple/multidimensional zero-correlation linear attack.

Corollary 3. *The number N^{KP} of known plaintexts required in a multiple or multidimensional zero-correlation linear attack is:*

$$N^{\mathrm{KP}} \approx \frac{2^n(\varphi_{PS} + \varphi_a)}{\sqrt{\ell/2 - \varphi_a}}. \tag{26}$$

The number N^{DKP} of distinct-known plaintexts required in a multiple or multidimensional zero-correlation linear attack is:

$$N^{\mathrm{DKP}} \approx \frac{2^n(\varphi_{PS} + \varphi_a)}{\sqrt{\ell/2 + \varphi_{PS}}}. \tag{27}$$

Proof. The result is straightforward while putting $\lambda = 0$ in Equation (25). □

In [9, 11, 13], Equations (26) and (27) were given for respectively KP multiple zero-correlation attacks and DKP multidimensional zero-correlation linear attacks. The results of this paper allows now to consider the two other cases: DKP multiple zero-correlation attacks and KP multidimensional zero-correlation attacks is backed up by experiments in the next section.

Since for most attacks $0.5 \leq P_S \leq 0.99$, meaning that $0 \leq \varphi_{P_S} \leq 2.4$, the difference between Equation (27) and Equation (26) is particularly noticeable when $\sqrt{\ell/2}$ and φ_a are in the same order of magnitude. From Equation (27) and Equation (26) we deduce that the success probability of a known-plaintext zero-correlation linear attack is:

$$P_S \approx \Phi\left(\frac{N^{\mathrm{KP}}}{2^n}\sqrt{\ell/2} - \varphi_a\left(\frac{N^{\mathrm{KP}} + 2^n}{2^n}\right)\right), \tag{28}$$

and the one of a distinct-known-plaintext zero-correlation linear attack is:

$$P_S \approx \Phi\left(\frac{N^{\mathrm{DKP}}\sqrt{\ell/2}}{2^n - N^{\mathrm{DKP}}} - \varphi_a\frac{2^n}{2^n - N^{\mathrm{DKP}}}\right). \tag{29}$$

6.2 Experimental Results

We have implemented experiments on a Feistel-type cipher which is depicted in Figure 10 and could correspond to scaled versions of CLEFIA [31] (a 16-bit type-II GFN with 4 branches) .

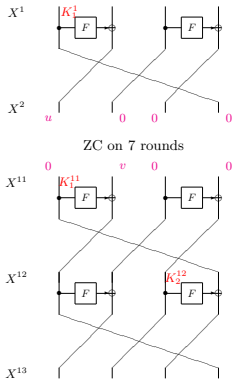
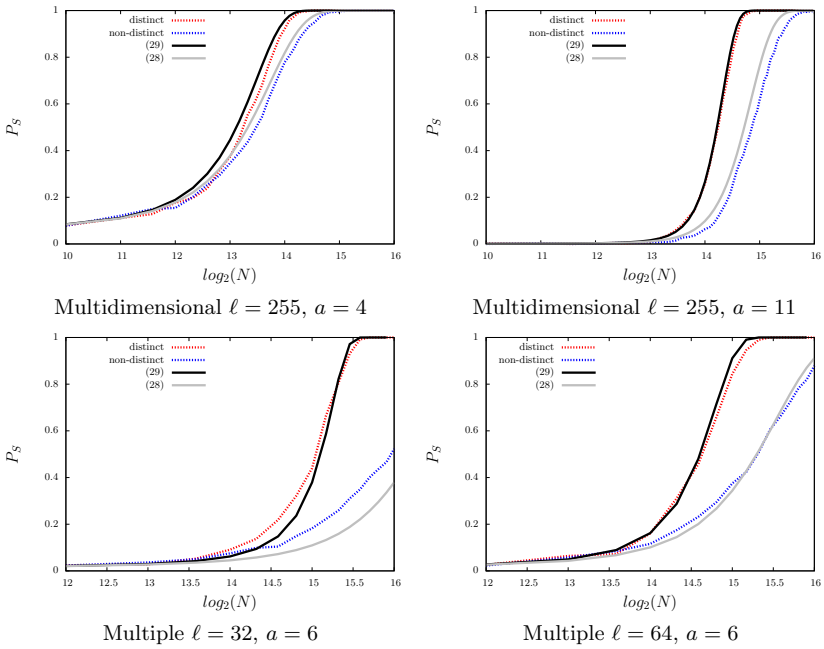


Fig. 10: Description of the key-recovery attack done on a Type-II GFN.

While in [32] experiments showing the distribution of $\mathrm{Exp}_{D,K}(T(D, K_R))$ and $\mathrm{Exp}_{D,K_W}(T(D, K_W))$ have been presented, there is, to the best of our knowledge, no previous mentioning of experimental zero-correlation linear attacks in the literature.

The results of our experimental attacks averaged over 1000 keys are provided in Figure 11. In these graphics we compare the success probability of multidimensional and multiple zero-correlation linear attacks with the theoretical ones given for KP by Equation (28) and for DKP by Equation (29). These experiments support the theory given in Section 6.1 showing that the same formula can be used to compute the complexity of multiple zero-correlation and multidimensional zero-correlation linear attacks. The difference lies only in the way of sampling, whether distinct or non-distinct known plaintexts are used in the attack. The bottom left graphic of Figure 11 corresponds to a case where only 32 approximations are taken into consideration. The gap between the theoretical and experimental success probability observed in these experiments is due to the approximation of the gamma distributions by normal distributions which is not accurate since $\ell = 32$. Using Equation 4 with the corresponding gamma distribution we obtain a more accurate estimate of the success probability of the attack.



Multidimensional $\ell = 255$, $a = 4$ | Multidimensional $\ell = 255$, $a = 11$

Multiple $\ell = 32$, $a = 6$ | Multiple $\ell = 64$, $a = 6$

Fig. 11: Attacks on a type-II-GFN cipher. Top: multidimensional zero-correlation linear attacks, bottom: multiple zero-correlation linear attacks.

6.3 Applications

Multiple zero-correlation linear attacks. As explained in detail later in this paper, by considering distinct-known plaintexts we can use Equation (27) to compute the data complexity of a multiple zero-correlation linear attack. As the data complexity of multidimensional linear attacks has already been computed under this setting, and because other comparable (in number of attacked rounds) attacks have been performed in the chosen-plaintext model, this should give us a better comparison factor. The result of our computation and a comparison with the best attacks on the block cipher Camellia [2] are provided in Table 1. The attack is from [9]. The data complexity has been computed using Equation (27) instead of using Equation (26) with the parameters of the attack chosen as $P_S = 0.85$ and $a = 96$ or $a = 160$. The time complexity has been computed according to the description given in [9]. We use the abbreviations KP, DKP and CP for known plaintext, distinct-known plaintext and chosen plaintext, respectively.

Version	#R	Type	ℓ	a	P_S	N	Time	Mem.	Ref.
128	11	ID	-	-	-	$2^{118.4}$ CP	$2^{118.43}$	$2^{96.4}$	[14]
128	11	ZC	2^{14}	96	85%	$2^{125.3}$ KP	$2^{125.8}$	2^{112}	[9]
128	11	ZC	2^{14}	96	85%	$\mathbf{2^{125.1}}$ **DKP**	$2^{125.8}$	2^{112}	Equation (27)
192	12	ID	-	-	-	$2^{119.7}$ CP	$2^{161.06}$	$2^{147.7}$	[14]
192	12	ZC	2^{14}	160	85%	$2^{125.7}$ KP	$2^{125.8}$	2^{112}	[9]
192	12	ZC	2^{14}	160	85%	$\mathbf{2^{125.46}}$ **DKP**	$2^{125.8}$	2^{112}	Equation (27)

Table 1: Best key-recovery attacks on Camellia-128 and Camellia-192 (attacks starting from the first round). The memory is expressed in number of bytes. #R denotes the number of attacked rounds. ID stands for impossible differential, ZC for zero-correlation.

Similarly we can improve the data complexity of the multiple zero-correlation linear attack on CAST-128 [34]. The parameters of the attack being $n = 128$, $\ell = 64770$, $a = 50$ and $P_S = 0.85$, the data complexity of the attack using known plaintextsfootnoteWith these parameters, the data complexity can not be equal to $2^{123.2}$ as given in [34]. is $N = 2^{123.73}$ and the data complexity of the attack using distinct-known plaintexts is $N = 2^{123.67}$.

Key-difference-invariant-bias attacks. Key-difference-invariant-bias attack is a related-key linear attack introduced in [8]. In this attack the attacker is taking advantage of linear approximations with same bias for different related keys. For these attacks, the statistical analysis is similar to the one done for zero-correlation linear attacks. For these attacks we can estimate the data complexity using the formulas provided in Corollary 3. In Table 2 we summarize the complexity of the best related key-attacks on LBlock [36]. These ones were obtained assuming a KP sampling. Assuming a DKP sampling, and the formula of the

data complexity provided by Equation (27), we show that the data complexity and a-fortiori the time complexity of the attack can be improved. Similar improvement can be obtained for the related-key attack on TWINE presented in [8]. The two-letter abbreviation RK refers to related-key attack throughout the table.

#R	Type	#Keys	ℓ	a	P_S	N	Time	Mem.	Ref.
23	RKID	4	-	-	-	$2^{61.4}$ RKCP	$2^{78.3}$	$2^{61.4}$	[35]
24	KIB	32	$2^{7.81}$	4.5	85%	$2^{62.29}$ RKKP	$2^{74.59}$	2^{61}	[8]
24	KIB	32	$2^{7.81}$	8.5	85%	$2^{62.95}$ RKKP	$2^{70.67}$	2^{61}	[8]
24	KIB	32	$2^{7.81}$	8.5	85%	$\mathbf{2^{62.38}}$**RKDKP**	$2^{70.67}$	2^{61}	Equation (27)
24	KIB	32	$2^{7.81}$	**16**	85%	$\mathbf{2^{62.84}}$ **RKDKP**	$\mathbf{2^{66.57}}$	2^{61}	Equation (27)*

Table 2: Best related-key attacks on LBlock. *: Computation of the time complexity according to the description given in Section 5.3 of [8]. RKID stands for related-key impossible differential, KIB for key-invariant bias.

7 Conclusion

In this paper, we presented enhancements to the statistical models of linear type attacks on iterated block ciphers. Our main result is a general statistical model that considers both data and key as random variables and covers multiple and multidimensional linear attacks, including zero-correlation attacks, which use random sampling of plaintext-ciphertext pairs, either with replacement or without replacement.

First, we elaborate in detail the regular key-recovery attack that exploits one linear approximation with a single dominant characteristic. When integrating the key as a random variable in the model, the data complexity of distinguishing between wrong and right key candidates can be expressed as a function of the ELP of the linear approximation. Before, the data complexity is determined from a statistical model assuming that for each fixed cipher key and key candidate the distribution of the test statistic has identical distribution. The new integrated statistical model gives the data complexity estimate for a random key. As a consequence, the issue raised in [27] is resolved. In particular, the fact that multiple strong characteristics cancel each other for many keys is not a problem for linear cryptanalysis in general. While it has been known by most researchers that ELP is the right quantity to consider in the context of linear attacks, no satisfactory presentation of how it determines the complexity of the attack for a random encryption key has not been given in the literature until now.

We then extend the statistical model with both data and key as random variables to multiple and multidimensional linear cryptanalysis. This model is built under the simplifying assumption that all linear approximations involved

in the attack are independent and have equal variance over data and key. In the case of multidimensional linear cryptanalysis it suffices to make an assumption on variance with the key. The assumptions are the same that were previously used by Huang et al. [20] and shown to produce accurate results in simulations of key-recovery attacks.

We also tested the validity of the new model in experiments on various multidimensional linear approximations over small versions of SMALLPRESENT which allows to compute the exact values of distribution parameters. While the new model seems to predict the expected value of the test statistic very accurately, the variance in the multidimensional linear attack is significantly underestimated. Resolving this issue by developing more accurate variance estimates is left for future work. Based on our preliminary investigations we believe that improvements can be obtained for certain type of ciphers such as key-alternating block ciphers.

Finally we apply the new model to zero-correlation linear attacks. Previous model for multidimensional zero-correlation attacks was given for DKP sampling, while the multiple zero-correlation linear attack has been modeled only for KP sampling. The new unified model contains these models as special cases. In addition, it now allows also the multidimensional attacks to use KP data, which is the standard setting in linear cryptanalysis. In the zero-correlation attacks, the distribution of the test statistic is key dependent only for the wrong key. One of the results given in this paper is a justification of the previously used ad hoc model for the wrong key behavior in the multidimensional zero-correlation cryptanalysis.

Acknowledgements

We wish to thank the anonymous reviewers for insightful comments that were very helpful in improving the presentation of this paper. In particular, we followed their advice to include more tutorial type material on linear key-recovery attacks and elaborate the case of a single linear approximation in detail.

References

1. Mohamed Ahmed Abdelraheem, Martin Ågren, Peter Beelen, and Gregor Leander. On the distribution of linear biases: Three instructive examples. In Reihaneh Safavi-Naini and Ran Canetti, editors, *Advances in Cryptology - CRYPTO 2012 - 32nd Annual Cryptology Conference, Santa Barbara, CA, USA, August 19-23, 2012. Proceedings*, volume 7417 of *Lecture Notes in Computer Science*, pages 50–67. Springer, 2012.
2. Kazumaro Aoki, Tetsuya Ichikawa, Masayuki Kanda, Mitsuru Matsui, Shiho Moriai, Junko Nakajima, and Toshio Tokita. Camellia: A 128-Bit Block Cipher Suitable for Multiple Platforms - Design and Analysis. In Douglas R. Stinson and Stafford E. Tavares, editors, *SAC 2000*, volume 2012 of *LNCS*. Springer, 2001.

3. Thomas Baignères, Pascal Junod, and Serge Vaudenay. How far can we go beyond linear cryptanalysis? In *Advances in Cryptology - ASIACRYPT 2004*, volume 3329 of *Lecture Notes in Computer Science*, pages 432–450. Springer-Verlag, 2004.
4. Eli Biham and Adi Shamir. Differential Cryptanalysis of DES-like Cryptosystems. In Alfred Menezes and Scott A. Vanstone, editors, *CRYPTO*, volume 537 of *LNCS*, pages 2–21. Springer, 1990.
5. Eli Biham and Adi Shamir. Differential cryptanalysis of des-like cryptosystems. In *Advances in Cryptology - CRYPTO '90, 10th Annual International Cryptology Conference, Santa Barbara, California, USA, August 11-15, 1990, Proceedings*, volume 537 of *Lecture Notes in Computer Science*, pages 2–21. Springer, 1991.
6. Alex Biryukov, Christophe De Cannière, and Michaël Quisquater. On Multiple Linear Approximations. In *CRYPTO 2004*, pages 1–22, 2004.
7. Céline Blondeau and Kaisa Nyberg. Links Between Truncated Differential and Multidimensional Linear Properties of Block Ciphers and Underlying Attack Complexities. In Elisabeth Oswald and Phong Q. Nguyen, editors, *Eurocrypt 2014*, volume 8441 of *LNCS*. Springer-Verlag, 2014.
8. 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, *ASIACRYPT 2013*, volume 8269 of *LNCS*, pages 357–376. Springer, 2013.
9. Andrey Bogdanov, Huizheng Geng, Meiqin Wang, Long Wen, and Baudoin Collard. Zero-Correlation Linear Cryptanalysis with FFT and Improved Attacks on ISO Standards Camellia and CLEFIA. In *SAC'13*, LNCS. Springer, 2014.
10. Andrey Bogdanov, Lars R. Knudsen, Gregor Leander, Christof Paar, Axel Poschmann, Matthew J. B. Robshaw, Yannick Seurin, and C. Vikkelsoe. PRESENT: An Ultra-Lightweight Block Cipher. In Pascal Paillier and Ingrid Verbauwhede, editors, *CHES*, volume 4727 of *LNCS*, pages 450–466. Springer, 2007.
11. Andrey Bogdanov, Gregor Leander, Kaisa Nyberg, and Meiqin Wang. Integral and Multidimensional Linear Distinguishers with Correlation Zero. In Xiaoyun Wang and Kazue Sako, editors, *ASIACRYPT*, volume 7658 of *LNCS*, pages 244–261. Springer, 2012.
12. Andrey Bogdanov and Elmar Tischhauser. On the wrong key randomisation and key equivalence hypotheses in Matsui's Algorithm 2. In Shiho Moriai, editor, *Fast Software Encryption - 20th International Workshop, FSE 2013, Singapore, March 11-13, 2013. Revised Selected Papers*, volume 8424 of *Lecture Notes in Computer Science*, pages 19–38. Springer, 2013.
13. Andrey Bogdanov and Meiqin Wang. Zero Correlation Linear Cryptanalysis with Reduced Data Complexity. In Anne Canteaut, editor, *FSE*, volume 7549 of *LNCS*, pages 29–48. Springer, 2012.
14. Christina Boura, María Naya-Plasencia, and Valentin Suder. Scrutinizing and Improving Impossible Differential Attacks: Applications to CLEFIA, Camellia, LBlock and Simon. In Palash Sarkar and Tetsu Iwata, editors, *ASIACRYPT 2014*, volume 8873 of *LNCS*, pages 179–199. Springer, 2014.
15. Anne Canteaut, Claude Carlet, Pascal Charpin, and Caroline Fontaine. On cryptographic properties of the cosets of r(1, m). *IEEE Trans. IT*, 47(4):1494–1513, 2001.
16. Joan Daemen, René Govaerts, and Joos Vandewalle. Correlation matrices. In *Fast Software Encryption - FSE 1994*, volume 1008 of *Lecture Notes in Computer Science*, pages 275–285. Springer-Verlag, 1995.
17. Joan Daemen and Vincent Rijmen. Probability distributions of correlation and differentials in block ciphers. *IACR Cryptology ePrint Archive*, 2005:212, 2006.

18. Joan Daemen and Vincent Rijmen. Probability distributions of correlation and differentials in block ciphers. *J. Mathematical Cryptology*, 1(3):221–242, 2007.
19. Miia Hermelin, Joo Yeon Cho, and Kaisa Nyberg. Multidimensional Extension of Matsui's Algorithm 2. In *FSE*, volume 5665 of *LNCS*, pages 209–227. Springer, 2009.
20. Jialin Huang, Serge Vaudenay, Xuejia Lai, and Kaisa Nyberg. Capacity and data complexity in multidimensional linear attack. In Rosario Gennaro and Matthew Robshaw, editors, *Advances in Cryptology - CRYPTO 2015 - 35th Annual Cryptology Conference, Santa Barbara, CA, USA, August 16-20, 2015, Proceedings, Part I*, volume 9215 of *Lecture Notes in Computer Science*, pages 141–160. Springer, 2015.
21. Lars R. Knudsen. Truncated and higher order differentials. 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 196–211. Springer, 1994.
22. Gregor Leander. Small scale variants of the block cipher PRESENT. *IACR Cryptology ePrint Archive*, 2010:143, 2010.
23. Gregor Leander. On Linear Hulls, Statistical Saturation Attacks, PRESENT and a Cryptanalysis of PUFFIN. In K. G. Paterson, editor, *EUROCRYPT*, volume 6632 of *LNCS*, pages 303–322. Springer, 2011.
24. N. Linial, Y. Mansour, and N. Nisan. Constant depth circuits, fourier transform, and learnability. *Journal of the ACM*, 40(3):607–620, 1993.
25. Mitsuru Matsui. Linear Cryptanalysis Method for DES Cipher. In Tor Helleseth, editor, *EUROCRYPT*, volume 765 of *LNCS*, pages 386–397. Springer, 1993.
26. James McLaughlin and John A. Clark. Filtered nonlinear cryptanalysis of reduced-round serpent, and the wrong-key randomization hypothesis. In *Cryptography and Coding - 14th IMA International Conference, IMACC 2013, Oxford, UK, December 17-19, 2013. Proceedings*, volume 8308 of *Lecture Notes in Computer Science*, pages 120–140. Springer, 2013.
27. Sean Murphy. The effectiveness of the linear hull effect. Technical report, Royal Holloway College London, 2009.
28. Kaisa Nyberg. Linear approximation of block ciphers. In *Advances in Cryptology - EUROCRYPT'94*, volume 950 of *Lecture Notes in Computer Science*, pages 439–444. Springer-Verlag, 1995.
29. 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.
30. Ali Aydin Selçuk. On Probability of Success in Linear and Differential Cryptanalysis. *J. Cryptology*, 21(1):131–147, 2008.
31. Taizo Shirai, Kyoji Shibutani, Toru Akishita, Shiho Moriai, and Tetsu Iwata. The 128-Bit Block cipher CLEFIA (Extended Abstract). In Alex Biryukov, editor, *FSE*, volume 4593 of *LNCS*, pages 181–195. Springer, 2007.
32. Hadi Soleimany and Kaisa Nyberg. Zero-correlation linear cryptanalysis of reduced-round LBlock. *Des. Codes Cryptography*, 73(2):683–698, 2014.
33. Eric Weisstein. Binomial distribution. *Wolfram MathWorld*, 2016.
34. Long Wen, Meiqin Wang, Andrey Bogdanov, and Huaifeng Chen. General Application of FFT in Cryptanalysis and Improved Attack on CAST-256. In Willi Meier and Debdeep Mukhopadhyay, editors, *INDOCRYPT 2014*, volume 8885 of *LNCS*, pages 161–176. Springer, 2014.
35. Long Wen, Meiqin Wang, and Jingyuan Zhao. Related-Key Impossible Differential Attack on Reduced-Round LBlock. *J. Comput. Sci. Technol.*, 29(1):165–176, 2014.

36. Wenling Wu and Lei Zhang. LBlock: A Lightweight Block Cipher. In Javier Lopez and Gene Tsudik, editors, *ACNS*, volume 6715 of *LNCS*, pages 327–344, 2011.