Zero Correlation Linear Cryptanalysis with Reduced Data Complexity

Andrey Bogdanov[1⋆] and Meiqin Wang[1,2⋆]

[1] KU Leuven, ESAT/COSIC and IBBT, Belgium
[2] Shandong University, Key Laboratory of Cryptologic Technology and Information Security, Ministry of Education, Shandong University, Jinan 250100,China

Abstract. Zero correlation linear cryptanalysis is a novel key recovery technique for block ciphers proposed in [5]. It is based on linear approximations with probability of exactly 1/2 (which corresponds to the zero correlation). Some block ciphers turn out to have multiple linear approximations with correlation zero for each key over a considerable number of rounds. Zero correlation linear cryptanalysis is the counterpart of impossible differential cryptanalysis in the domain of linear cryptanalysis, though having many technical distinctions and sometimes resulting in stronger attacks.

In this paper, we propose a statistical technique to significantly reduce the data complexity using the high number of zero correlation linear approximations available. We also identify zero correlation linear approximations for 14 and 15 rounds of TEA and XTEA. Those result in key-recovery attacks for 21-round TEA and 25-round XTEA, while requiring less data than the full code book. In the single secret key setting, these are structural attacks breaking the highest number of rounds for both ciphers.

The findings of this paper demonstrate that the prohibitive data complexity requirements are not inherent in the zero correlation linear cryptanalysis and can be overcome. Moreover, our results suggest that zero correlation linear cryptanalysis can actually break more rounds than the best known impossible differential cryptanalysis does for relevant block ciphers. This might make a security re-evaluation of some ciphers necessary in the view of the new attack.

Keywords: block ciphers, key recovery, linear cryptanalysis, zero correlation linear cryptanalysis, data complexity, TEA, XTEA

1 Introduction

1.1 Motivation

Differential and linear cryptanalyses [3, 30] are the two basic tools for evaluating the security of block ciphers such as the former U.S. encryption standard DES as well as its successor AES. While DES was developed at the time when differential and linear cryptanalyses were not publicly known, the design of AES provably addresses these attacks.

Design strategies have been proposed such as the wide-trail design strategy [13] or decorrelation theory [42] to make ciphers resistant to the basic flavours of differential and linear cryptanalysis. However, a proof of resistance according to these strategies does not necessarily imply resistance to the extensions of these techniques such as impossible differential cryptanalysis [1, 6] and the recently proposed zero correlation linear cryptanalysis [5].

Standard differential cryptanalysis uses differentials with probabilities significantly higher than those expected for a randomly drawn permutation. Similarly, basic linear cryptanalysis uses linear approximations whose probabilities detectably deviate from 1/2. At the same time, impossible differential cryptanalysis and zero correlation linear cryptanalysis are based on structural deviations of another kind: Differentials with zero probability are targeted in impossible differential cryptanalysis and linear approximations with probability of exactly 1/2 correlation are exploited in zero correlation linear cryptanalysis. Thus, zero correlation linear cryptanalysis

⋆ Both authors are corresponding authors.

can be seen as the counterpart of impossible differential cryptanalysis in the domain of linear cryptanalysis.

The name of the attack originated from the notion of *correlation* [11,34]: If $\frac{1+c}{2}$ is the probability for a linear approximation to hold, c is called the correlation of this linear approximation. Clearly, putting $c = 0$ yields an unbiased linear approximation of probability $1/2$, or a *zero correlation linear approximation.*

Impossible differential cryptanalysis has been known to the cryptographic community since over a decade now. It has turned out a highly useful tool of attacking block ciphers [2,15,27–29, 41]. In fact, among meet-in-the-middle [14] and multiset-type attacks [18], it is the impossible differential cryptanalysis [28] that breaks the highest numbers of rounds of AES-128 and AES-256 in the classical single-key attack model as to date, the recent biclique cryptanalysis [4] being the notable exception though.

Zero correlation linear cryptanalysis is a novel promising attack technique that bears some technical similarities to impossible differential cryptanalysis but has its theoretical foundation in a different mathematical theory. Despite its newness, it has already been demonstrated to successfully apply to round-reduced AES and CLEFIA even in its basic form [5], which is highly motivating for further studies.

In this paper, we show how to remove the data requirement of the full codebook which was the major limitation of basic zero correlation linear cryptanalysis [5]. As an application of zero correlation linear cryptanalysis and this data complexity reduction technique, we propose attacks against round-reduced TEA and XTEA. For both ciphers, we can cryptanalyze more rounds than it was previously possible using less than the full code book.

1.2 Contributions

The work at hand has two major contributions.

Data complexity reduction for zero correlation linear cryptanalysis. The data requirements of the full codebook have been a crucial limitation for the recent zero correlation linear cryptanalysis to become a major cryptanalytic technique, though the length of the fundamental property (the length of the zero correlation linear approximation) was demonstrated to be comparable to that of impossible differentials for several cipher structures [5]. Overcoming this annoying limitation, a statistical technique of data complexity reduction for zero correlation linear cryptanalysis is the first contribution of this paper.

The data complexity reduction technique is based on the fact that, like any exploitable impossible differential, a typical zero correlation linear approximation is *truncated*: That is, once a zero correlation linear approximation has been identified that holds for all keys, it will as a rule imply an entire class of similar zero correlation linear approximations to exist. Those can be typically obtained by just changing several bits of the input mask, output mask or both. In other words, in most practical cases, there will be *multiple* zero correlation linear approximations available to the adversary which has been ignored by the previous analysis.

However, unlike in impossible differential cryptanalysis, the actual value of the correlation has to be estimated in zero correlation linear cryptanalysis and it is not enough to just wait for the impossible event to occur. In fact, the idea we use for zero correlation linear cryptanalysis is more similar to that of multiple linear cryptanalysis: We estimate the correlation of each individual linear approximation using a limited number of texts. Then, for a group of zero correlation linear approximations (i.e. for the right key), we expect the cumulative deviation of those estimations from 0 to be lower than that for a group of randomly chosen linear approximations (i.e. for a wrong key). Given the statistical behaviour of correlation for a randomly drawn permutation [12, 35], this consideration results in a χ^2 statistic and allows for a theoretical analysis of the

complexity and error probabilities of a zero correlation linear attack that are confirmed by experiments.

Table 1. Summary of cryptanalytic results on round-reduced TEA* and XTEA in the single-key setting

attack	#rounds	data	comp. compl.	memory	Pr[success]	ref.
TEA						
impossible differential	11	$2^{52.5}$ CP	2^{84}	NA	NA	[32]
truncated differential	17	1920 CP	$2^{123.37}$	NA	NA	[20]
impossible differential	17	2^{57} CP	$2^{106.6}$	2^{49}	NA	[8]
zero correlation linear	**21**	$\mathbf{2^{62.62}}$ **KP**	$\mathbf{2^{121.52}}$	**negligible**	**0.846**	**this paper**
zero correlation linear	**23**	$\mathbf{2^{64}}$	$\mathbf{2^{119.64}}$	**negligible**	**1**	**this paper**
XTEA						
impossible differential	14	$2^{62.5}$ CP	2^{85}	NA	NA	[32]
truncated differential	23	$2^{20.55}$ CP	$2^{120.65}$	NA	0.969	[20]
meet-in-the-middle	23	18 KP	2^{117}		$1-2^{-1025}$	[37]
impossible differential	23	$2^{62.3}$ CP	$2^{114.9}$	$2^{94.3}$	NA	[8]
impossible differential	23	2^{63}	2^{101} MA $+2^{105.6}$	2^{103}	NA	[8]
zero correlation linear	**25**	$\mathbf{2^{62.62}}$ **KP**	$\mathbf{2^{124.53}}$	$\mathbf{2^{30}}$	**0.846**	**this paper**
zero correlation linear	**27**	$\mathbf{2^{64}}$	$\mathbf{2^{120.71}}$	**negligible**	**1**	**this paper**

CP: Chosen Plaintexts, KP: Known Plaintexts.
Memory: the number of 32-bit words.
*The effective key length for TEA is 126 bit

Zero correlation linear cryptanalysis of round-reduced TEA and XTEA. TEA (Tiny Encryption Algorithm) is one of the first lightweight block ciphers. It is a 64-bit block cipher based on a balanced Feistel-type network with a simple ARX round function. TEA has 64 rounds and accepts a key of 128 bits. It favours both efficient hardware [22] and software implementations. TEA was designed by Wheeler and Needham and proposed at FSE'94 [43]. It was used in Microsoft's Xbox gaming console for checking software authenticity until its weakness as a hash function was used [40] to compromise the chain of trust. The block cipher XTEA [33] is the fixed version of TEA eliminating this property (having the same number rounds, block size, and key size). TEA and XTEA being rather popular ciphers, both are implemented in the Linux kernel.

Similarly to the complementation property of DES, TEA has an equivalent key property and its effective key size is 126 bits (compared to 128 bits suggested by the nominal key input size) [23]. Kelsey, Scheier and Wagner [24] proposed a practical related-key attack on the full TEA. Using complementation cryptanalysis [7], up to 36 rounds of XTEA can be attacked with related keys for all keys. The work [7] also contains related-key attacks for up to 50 rounds of XTEA working for a weak key class.

In the classical single-key setting, however, by far not all rounds of TEA are broken by structural attacks (whereas the effective key size is 126 bits for the full cipher). The truncated differential result on 17 rounds remains the best cryptanalysis of TEA [20]. Impossible differential cryptanalysis [8] has yielded a faster attack against 17 rounds of TEA. Similarly, 23 rounds of XTEA have been cryptanalyzed so far using truncated differential [20], impossible differential [8] and well as meet-in-the-middle attacks [37]. That is, for both TEA and XTEA, there has been no progress in terms of the number of attacked rounds since 2003.

In this paper, using zero correlation linear cryptanalysis, we cryptanalyze 21 rounds of TEA and 25 rounds of XTEA with $2^{62.62}$ *known* plaintexts (in contrast to *chosen* texts required in impossible differential cryptanalysis). Certainly, zero correlation linear cryptanalysis for lower number of rounds yields a lower data complexity for both TEA and XTEA. Moreover, unlike most impossible differential attacks including those on TEA and XTEA [8], zero correlation linear cryptanalysis is able to profit from the full code available. If all 2^{64} texts are available to the

adversary, we propose zero correlation linear cryptanalysis for 23 rounds of TEA and 27 rounds of XTEA. Our cryptanalytic results are summarized and compared to previous cryptanalysis in Table 1.

As opposed to the initial intuition expressed in [5], both major contributions of this work — the data complexity reduction and the new attacks on more rounds of TEA and XTEA — demonstrate that zero correlation linear cryptanalysis can actually perform better than impossible differential cryptanalysis. Moreover, we expect the security of more ciphers to be reevaluated under the consideration of zero correlation linear cryptanalysis.

1.3 Outline

We start with a review of the basic zero correlation linear cryptanalysis for block ciphers in Section 2. In Section 3, we introduce a χ^2 statistical technique for reducing the data requirements of zero correlation linear cryptanalysis and thoroughly investigate its complexity. In Section 4, the 14- and 15-round zero correlation linear approximations are demonstrated for block ciphers TEA and XTEA. Section 5 gives several zero correlation key recoveries for round-reduced TEA and XTEA. We conclude in Section 6. Appendices contain proofs of some technical statements as well as further zero correlation linear attacks on round-reduced TEA and XTEA.

2 Basic zero correlation linear cryptanalysis

Zero correlation linear cryptanalysis has been introduced in [5]. Below we briefly review its basic ideas and methods.

2.1 Linear approximations with correlation zero

Consider an n-bit block cipher f_K with key K. Let P denote a plaintext which is mapped to ciphertext C under key K, $C = f_K(P)$. If Γ_P and Γ_C are nonzero plaintext and ciphertext linear masks of n bit each, we denote by $\Gamma_P \to \Gamma_C$ the linear approximation

$$\Gamma_P^T P \oplus \Gamma_C^T C = 0.$$

Here, $\Gamma_A^T A$ denotes the multiplication of the transposed bit vector Γ_A (linear mask for A) by a column bit vector A over $\mathbb{F}_2$. The linear approximation $\Gamma_P \to \Gamma_C$ has probability

$$p_{\Gamma_P,\Gamma_C} = \Pr_{P \in \mathbb{F}_2^n}\{\Gamma_P^T P \oplus \Gamma_C^T C = 0\}. \tag{1}$$

The value

$$c_{\Gamma_P,\Gamma_C} = 2p_{\Gamma_P,\Gamma_C} - 1 \tag{2}$$

is called the *correlation (or bias)* of linear approximation $\Gamma_P \to \Gamma_C$. Note that $p_{\Gamma_P,\Gamma_C} = 1/2$ is equivalent to *zero correlation* $c_{\Gamma_P,\Gamma_C} = 0$:

$$p_{\Gamma_P,\Gamma_C} = \Pr_{P \in \mathbb{F}_2^n}\{\Gamma_P^T P \oplus \Gamma_C^T C = 0\} = 1/2. \tag{3}$$

In fact, for a randomly drawn permutation of sufficiently large bit size n, zero is the most frequent single value of correlation for a nontrivial linear approximation. Correlation goes to small values for increasing n, the probability to get exactly zero decreases as a function of n though. More precisely, the probability of the linear approximation $\Gamma_P \to \Gamma_C$ with $\Gamma_P, \Gamma_C \neq 0$ to have zero correlation has been shown [5, Proposition 2] to be approximated by

$$\frac{1}{\sqrt{2\pi}} 2^{\frac{4-n}{2}}. \tag{4}$$

2.2 Two examples

Given a randomly chosen permutation, however, it is hard to tell a priori which of its nontrivial linear approximations in particular has zero correlation. At the same time, it is often possible to identify groups of zero correlation linear approximations for a block cipher f_K once it has compact description with a distinct structure. Moreover, in many interesting cases, these linear approximations will have zero correlation *for any key* K. Here are two examples provided in [5]:

- **AES:** The data transform of AES has a set of zero correlation linear approximations over 4 rounds (3 full rounds appended by 1 incomplete rounds with MixColumns omitted). If Γ and Γ' are 4-byte column linear masks with exactly one nonzero byte, then each of the linear approximations $(\Gamma, 0, 0, 0) \rightarrow (\Gamma', 0, 0, 0)$ over 4 AES rounds has zero correlation [5, Theorem 2].
- **CLEFIA-type GFNs:** CLEFIA-type generalized Feistel networks [39] (also known as type-2 GFNs with 4 lines [44]) have zero correlation linear approximations over 9 rounds, if the underlying F-functions of the Feistel construction are invertible. For $a \neq 0$, the linear approximations $(a, 0, 0, 0) \rightarrow (0, 0, 0, a)$ and $(0, 0, a, 0) \rightarrow (0, a, 0, 0)$ over 9 rounds have zero correlation [5, Theorem 1].

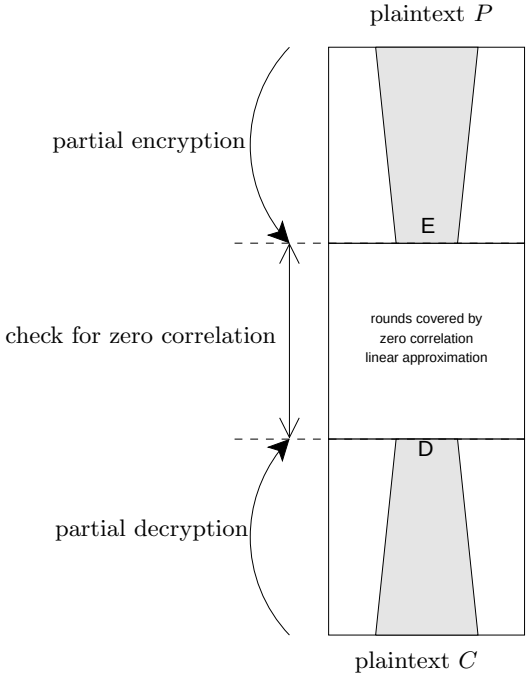

Fig. 1. High-level view of key recovery in zero correlation linear cryptanalysis

2.3 Key recovery with zero correlation linear approximations

Based on linear approximations of correlation zero, a technique similar to Matsui's Algorithm 2 [30] can be used for key recovery. Let the adversary have N known plaintext-ciphertexts and ℓ zero correlation linear approximations $\{\Gamma_E \rightarrow \Gamma_D\}$ for a part of the cipher, with $\ell = |\{\Gamma_E \rightarrow \Gamma_D\}|$. The linear approximations $\{\Gamma_E \rightarrow \Gamma_D\}$ are placed in the middle of the attacked cipher. Let E and D be the partial intermediate states of the data transform at the boundaries of the linear approximations.

Then the key can be recovered using the following approach (see also Figure 1):

1. Guess the bits of the key needed to compute E and D. For each guess:

(a) Partially encrypt the plaintexts and partially decrypt the ciphertexts up to the boundaries of the zero correlation linear approximation $\Gamma_E \to \Gamma_D$.
(b) Estimate the correlations $\{\hat{c}_{\Gamma_E,\Gamma_D}\}$ of all linear approximations in $\{\Gamma_E \to \Gamma_D\}$ for the key guess using the partially encrypted and decrypted values E and D by counting how many times $\Gamma_E^T E \oplus \Gamma_D^T D$ is zero over N input/output pairs, see (1) and (2).
(c) Perform a test on the estimated correlations $\{\hat{c}_{\Gamma_E,\Gamma_D}\}$ for $\{\Gamma_E \to \Gamma_D\}$ to tell of the estimated values of $\{\hat{c}_{\Gamma_E,\Gamma_D}\}$ are compatible with the hypothesis that all of the actual values of $\{c_{\Gamma_E,\Gamma_D}\}$ are zero.

2. Test the surviving key candidates against a necessary number of plaintext-ciphertext pairs according to the unicity distance for the attacked cipher.

Step 1(c) of the technique above relies on an efficient test distinguishing between the hypothesis that $\{c_{\Gamma_E,\Gamma_D}\}$ are all zero and the alternative hypothesis. The work [5] requires the exact evaluation of the correlation value (defined by the probability of a linear approximation) and the data complexity is restricted to $N = 2^n$ in [5]. Thus, a small number ℓ of linear approximations is usually enough in [5] and $\hat{c}_{\Gamma_E,\Gamma_D} = c_{\Gamma_E,\Gamma_D}$, though the data complexity of the full codebook is too restrictive.

For most ciphers (including the examples of Subsection 2.2), however, a large number ℓ of zero correlation linear approximations is available. This freedom is not used in [5]. At the same time, it has been shown in the experimental work [9] that any value of correlation can be used for key recovery in a linear attack with reduced data complexity, once enough linear approximations are available. Despite its convincing experimental evidence, [9] gives no theoretical data complexity estimations and does not provide any ways of constructing linear approximations with certain properties.

In the next section of this paper, we provide a framework for reducing the data complexity N if many zero correlation linear approximations are known.

3 Reduction of data complexity with many approximations

3.1 Distinguishing between two normal distributions

Consider two normal distributions: $\mathcal{N}(\mu_0, \sigma_0)$ with mean μ_0 and standard deviation σ_0, and $\mathcal{N}(\mu_1, \sigma_1)$ with mean μ_1 and standard deviation σ_1. A sample s is drawn from either $\mathcal{N}(\mu_0, \sigma_0)$ or $\mathcal{N}(\mu_1, \sigma_1)$. It has to be decided if this sample is from $\mathcal{N}(\mu_0, \sigma_0)$ or from $\mathcal{N}(\mu_1, \sigma_1)$. The test is performed by comparing the value s to some threshold value t. Without loss of generality, assume that $\mu_0 < \mu_1$. If $s \le t$, the test returns "$s \in \mathcal{N}(\mu_0, \sigma_0)$". Otherwise, if $s > t$, the test returns "$s \in \mathcal{N}(\mu_1, \sigma_1)$". There will be error probabilities of two types:

$$\beta_0 = \Pr\{\text{"}s \in \mathcal{N}(\mu_1, \sigma_1)\text{"} | s \in \mathcal{N}(\mu_0, \sigma_0)\},$$
$$\beta_1 = \Pr\{\text{"}s \in \mathcal{N}(\mu_0, \sigma_0)\text{"} | s \in \mathcal{N}(\mu_1, \sigma_1)\}.$$

Here a condition is given on μ_0, μ_1, σ_0, and σ_1 such that the error probabilities are β_0 and β_1. The proof immediately follows from the basics of probability theory (see e.g. [17, 19]) and is given in Appendix A for completeness.

Proposition 1. *For the test to have error probabilities of at most β_0 and β_1, the parameters of the normal distributions $\mathcal{N}(\mu_0, \sigma_0)$ and $\mathcal{N}(\mu_1, \sigma_1)$ with $\mu_0 \neq \mu_1$ have to be such that*

$$\frac{z_{1-\beta_1}\sigma_1 + z_{1-\beta_0}\sigma_0}{|\mu_1 - \mu_0|} = 1,$$

where $z_{1-\beta_1}$ and $z_{1-\beta_0}$ are the quantiles of the standard normal distribution.

3.2 A known plaintext distinguisher with many zero correlation linear approximations

Let the adversary be given N known plaintext-ciphertext pairs and ℓ zero correlation linear approximations for an n-bit block cipher. The adversary aims to distinguish between this cipher and a randomly drawn permutation.

The procedure is as follows. For each of the ℓ given linear approximations, the adversary computes the number T_i of times the linear approximations are fulfilled on N plaintexts, $i \in \{1, \ldots, \ell\}$. Each T_i suggests an empirical correlation value $\hat{c}_i = 2\frac{T_i}{N} - 1$. Then, the adversary evaluates the statistic:

$$\sum_{i=1}^{\ell} \hat{c}_i^2 = \sum_{i=1}^{\ell} \left(2\frac{T_i}{N} - 1\right)^2. \tag{5}$$

It is expected that for the cipher with ℓ known zero correlation linear approximations, the value of statistic (5) will be lower than that for ℓ linear approximations of a randomly drawn permutation. In a key-recovery setting, the right key will result in statistic (5) being among the lowest values for all candidate keys if ℓ is high enough. In the sequel, we treat this more formally.

3.3 Correlation under right and wrong keys

Consider the key recovery procedure outlined in Subsection 2.3 given N known plaintext-ciphertext pairs. There will be two cases:

- *Right key guess:* Each of the values $\hat{c}_i$ in (5) approximately follows the normal distribution with zero mean and standard deviation $1/\sqrt{N}$ with good precision (c.f. e.g. [21, 38]) for sufficiently large N:
$$\hat{c}_i \sim \mathcal{N}(0, 1/\sqrt{N}).$$
- *Wrong key guess:* Each of the values $\hat{c}_i$ in (5) approximately follows the normal distribution with mean c_i and standard deviation $1/\sqrt{N}$ for sufficiently large N:
$$\hat{c}_i \sim \mathcal{N}(c_i, 1/\sqrt{N}) \text{ with } c_i \sim \mathcal{N}(0, 2^{-n/2}),$$
where c_i is the exact value of the correlation which is itself distributed as $\mathcal{N}(0, 2^{-n/2})$ over random permutations with $n \geq 5$ — a result due to [12,35]. Thus, our wrong key hypothesis is that for each wrong key, the adversary obtains a permutation with linear properties close to those of a randomly chosen permutation.

3.4 Distribution of the statistic

Based on these distributions of $\hat{c}_i$, we now derive the distributions of statistic (5) in these two cases.

Right key guess. In this case, we deal with ℓ zero correlation linear approximations:

$$\sum_{i=1}^{\ell} \hat{c}_i^2 \sim \sum_{i=1}^{\ell} \mathcal{N}^2\left(0, 1/\sqrt{N}\right) = \frac{1}{N}\sum_{i=1}^{\ell} \mathcal{N}^2(0,1) = \frac{1}{N}\chi_\ell^2,$$

where χ_ℓ^2 is the χ^2-distribution with ℓ degrees of freedom which has mean ℓ and standard deviation $\sqrt{2\ell}$, assuming the independency of underlying distributions. For sufficiently large ℓ, χ_ℓ^2 converges to the normal distribution. That is, χ_ℓ^2 approximately follows $\mathcal{N}(\ell, \sqrt{2\ell})$, and:

$$\sum_{i=1}^{\ell} \hat{c}_i^2 \sim \frac{1}{N}\chi_\ell^2 \approx \frac{1}{N}\mathcal{N}\left(\ell, \sqrt{2\ell}\right) = \mathcal{N}\left(\frac{\ell}{N}, \frac{\sqrt{2\ell}}{N}\right). \tag{6}$$

Proposition 2. *Consider ℓ nontrivial zero correlation linear approximations for a block cipher with a fixed key. If N is the number of known plaintext-ciphertext pairs, T_i is the number of times such a linear approximation is fulfilled for $i \in \{1, \ldots, \ell\}$, and ℓ is high enough, then, assuming the counters T_i are independent, the following approximate distribution holds for sufficiently large N and n:*

$$\sum_{i=1}^{\ell} \left(2\frac{T_i}{N} - 1\right)^2 \sim \mathcal{N}\left(\frac{\ell}{N}, \frac{\sqrt{2\ell}}{N}\right).$$

Wrong key guess. The wrong key hypothesis is that we deal with pick a permutation at random for each wrong key. Therefore, the ℓ given linear approximations will have randomly drawn correlations, under this hypothesis. Thus, as mentioned above:

$$\sum_{i=1}^{\ell} \hat{c}_i^2 \sim \sum_{i=1}^{\ell} \mathcal{N}^2\left(c_i, 1/\sqrt{N}\right), \text{ where } c_i \sim \mathcal{N}\left(0, 2^{-n/2}\right).$$

First, we show that the underlying distribution of $\hat{c}_i$ is actually normal with mean 0. Then we show that the sum approximately follows χ^2-distribution assuming the independency of underlying distributions, and can be approximated by another normal distribution.

Since

$$\begin{aligned}\mathcal{N}\left(c_i, 1/\sqrt{N}\right) &= c_i + \mathcal{N}\left(0, 1/\sqrt{N}\right) \\ &= \mathcal{N}\left(0, 1/\sqrt{2^n}\right) + \mathcal{N}\left(0, 1/\sqrt{N}\right) \\ &= \mathcal{N}\left(0, \sqrt{1/N + 1/2^n}\right),\end{aligned}$$

the distribution above is a χ^2-distribution with ℓ degrees of freedom up to a factor, under the independency assumption:

$$\begin{aligned}\textstyle\sum_{i=1}^{\ell} \mathcal{N}^2\left(c_i, 1/\sqrt{N}\right) &= \textstyle\sum_{i=1}^{\ell} \mathcal{N}^2\left(0, \sqrt{\frac{1}{N} + \frac{1}{2^n}}\right) \\ &= \left(\tfrac{1}{N} + \tfrac{1}{2^n}\right) \textstyle\sum_{i=1}^{\ell} \mathcal{N}^2(0,1) \\ &= \left(\tfrac{1}{N} + \tfrac{1}{2^n}\right) \chi_\ell^2.\end{aligned}$$

As for the right keys, for sufficiently large ℓ, χ_ℓ^2 can be approximated by the normal distribution with mean ℓ and standard deviation $\sqrt{2\ell}$. Thus:

$$\begin{aligned}\textstyle\sum_{i=1}^{\ell} \hat{c}_i^2 &\sim \left(\tfrac{1}{N} + \tfrac{1}{2^n}\right) \chi_\ell^2 \approx \left(\tfrac{1}{N} + \tfrac{1}{2^n}\right) \mathcal{N}\left(\ell, \sqrt{2\ell}\right) \\ &= \mathcal{N}\left(\tfrac{\ell}{N} + \tfrac{\ell}{2^n}, \tfrac{\sqrt{2\ell}}{N} + \tfrac{\sqrt{2\ell}}{2^n}\right).\end{aligned}$$

Proposition 3. *Consider ℓ nontrivial linear approximations for a randomly drawn permutation. If N is the number of known plaintext-ciphertext pairs, T_i is the number of times a linear approximation is fulfilled for $i \in \{1, \ldots, \ell\}$, and ℓ is high enough, then, assuming the independency of T_i, the following approximate distribution holds for sufficiently large N and n:*

$$\sum_{i=1}^{\ell} \left(2\frac{T_i}{N} - 1\right)^2 \sim \mathcal{N}\left(\frac{\ell}{N} + \frac{\ell}{2^n}, \frac{\sqrt{2\ell}}{N} + \frac{\sqrt{2\ell}}{2^n}\right).$$

3.5 Data complexity of the distinguisher

Combining Propositions 2 and 3 with Proposition 1, one obtains the condition:

$$\frac{z_{1-\beta_1}\left(\frac{\sqrt{2\ell}}{N} + \frac{\sqrt{2\ell}}{2^n}\right) + z_{1-\beta_0}\frac{\sqrt{2\ell}}{N}}{\left(\frac{\ell}{N} + \frac{\ell}{2^n}\right) - \frac{\ell}{N}} = 1.$$

The left part of this equation can be simplified to

$$\frac{2^{n+0.5}}{N\sqrt{\ell}}\left(z_{1-\beta_0}+z_{1-\beta_1}\right)+\frac{z_{1-\beta_1}\sqrt{2}}{\sqrt{\ell}},$$

which yields

Theorem 1. *With the assumptions of Propositions 1 to 3, using ℓ nontrivial zero correlation linear approximations, to distinguish between a wrong key and a right key with probability β_1 of false positives and probability β_0 of false negatives, a number N of known plaintext-ciphertext pairs is sufficient if the following condition is fulfilled:*

$$\frac{2^{n+0.5}}{N\sqrt{\ell}}\left(z_{1-\beta_0}+z_{1-\beta_1}\right)+\frac{z_{1-\beta_1}\sqrt{2}}{\sqrt{\ell}}=1.$$

The success probability of an attack is defined by the probability β_0 of false negatives. The probability β_1 of false positives determines the number of surviving key candidates and, thus, influences the computational complexity of the key recovery.

4 Linear approximations with correlation zero for TEA and XTEA

In [5], a sufficient condition is given for a linear approximation to have a correlation of zero. Namely, if for a linear approximation there exist no linear characteristics with non-zero correlation contributions, then the correlation of the linear approximation is exactly zero.

4.1 The block ciphers TEA and XTEA

TEA is a 64-round iterated block cipher with 64-bit block size and 128-bit key which consist of four 32-bit words $K[0], K[1], K[2]$ and $K[3]$. TEA does not have any iterative key schedule algorithm. Instead, the key words are used directly in round functions. The round constant is derived from the constant $\delta = 9e3779b9_x$ and the round number. We denote the input and the output of the r-th round for $1 \le r \le 64$ by (L_r, R_r) and (L_{r+1}, R_{r+1}), respectively. $L_{r+1} = R_r$ and R_{r+1} is computed as follows:

$$R_{r+1} = \begin{cases} L_r + (((R_r \ll 4) + K[0]) \oplus (R_r + i\cdot\delta) \oplus (R_r \gg 5 + K[1])) & r = 2i-1, 1 \le i \le 32, \\ L_r + (((R_r \ll 4) + K[2]) \oplus (R_r + i\cdot\delta) \oplus (R_r \gg 5 + K[3])) & r = 2i, 1 \le i \le 32. \end{cases}$$

Like TEA, XTEA is also a 64-round Feistel cipher with 64-bit block and 128-bit key. Its 128-bit secret key K is represented by four 32-bit words $K[0], K[1], K[2]$ and $K[3]$ as well. The derivation of the subkey word number is slightly more complex though. The input of the r-th round is (L_r, R_r) and the output is (L_{r+1}, R_{r+1}). Again, $L_{r+1} = R_r$ and R_{r+1} is derived as:

$$R_{r+1} = \begin{cases} L_r + (((R_r \ll 4 \oplus R_r \gg 5) + R_r) \oplus ((i-1)\cdot\delta + K[((i-1)\cdot\delta \ll 11)\&3])) & r = 2i-1, 1 \le i \le 32, \\ L_r + (((R_r \ll 4 \oplus R_r \gg 5) + R_r) \oplus (i\cdot\delta + K[(i\cdot\delta \ll 11)\&3])) & r = 2i, 1 \le i \le 32. \end{cases}$$

These round functions of TEA and XTEA are illustrated in Figure 2.

4.2 Notations

To demonstrate zero correlation linear approximations for TEA and XTEA, we will need the following notations (the least significant bit of a word has number 0):

- $e_{i,\sim}$ is a 32-bit word that has zeros in bits $(i+1)$ to 31, one in bit i and undefined values in bits 0 to $(i-1)$,

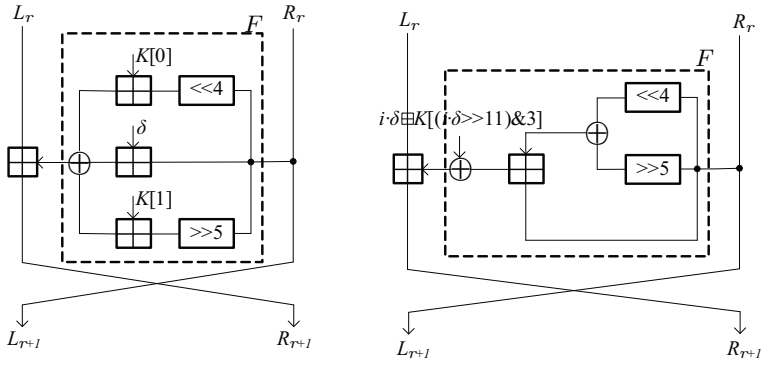

Fig. 2. Round function for TEA(left) and XTEA(right)

- $e_{i\sim j}$ is a 32-bit word that has zeros in bits $(i+1)$ to 31 and bits 0 to $(j-1)$, a one in bit i and undefined values in bits j to $(i-1)$ for $j < i$,
- $\bar{e}_{i,\sim}$ is a 32-bit word that has zeros in bits $(i+1)$ to 31, undefined values in bits 0 to i,
- ? is an undefined value,
- $X^{i\sim j}$ is bits from j to i of the value X, $j < i$, and
- X^i is the value of bit i of X.

4.3 Linear approximation of modular addition

Here, we first demonstrate the properties of linear approximations with non-zero correlation over the modular addition, which is the only nonlinear part of the TEA and XTEA transformation (summarized as Property 1). Then we use it to show a condition for linear approximation with non-zero correlation for one round of TEA and XTEA (stated as Property 2).

For the modular addition of two n-bit inputs x and y, the output z can be computed as:

$$z = (x + y) \mod 2^n.$$

We denote the mask values for x, y and z as $\Gamma x, \Gamma y$ and Γz, respectively ($x, y, z, \Gamma x, \Gamma y$, and $\Gamma z \in \mathbb{F}_2^n$). The linear approximation for the modular addition is then $\Gamma x^T \cdot x \oplus \Gamma y^T \cdot y = \Gamma z^T \cdot z$ and is referred to as

$$+ : (\Gamma x | \Gamma y) \to \Gamma z.$$

Property 1 (Modular addition). In any linear approximation $(\Gamma x|\Gamma y) \to \Gamma z$ of the modular addition with a non-zero correlation, the most significant non-zero mask bit for $\Gamma x, \Gamma y$ and Γz is the same.

Property 1 is proven in Appendix B.

4.4 Linear approximation of one TEA/XTEA round

Using Property 1 for modular addition, as all other operations in TEA and XTEA are linear, we can derive conditions on a special class of approximations with non-zero correlation for the round function of TEA and XTEA. See Figures 4 and 3 for an illustration.

As in Subsection 4.1, the input and output of round r in TEA and XTEA are $(L_r|R_r)$ and $(L_{r+1}|R_{r+1})$, respectively. Correspondingly, $(\Gamma_r^L|\Gamma_r^R)$ and $(\Gamma_{r+1}^L|\Gamma_{r+1}^R)$ are input and output linear masks of the round. So the linear approximation over the round is

$$\text{(X)TEA round } r : (\Gamma_r^L|\Gamma_r^R) \to (\Gamma_{r+1}^L|\Gamma_{r+1}^R)$$

and has the following

Property 2 (One round). If $\Gamma_r^L = e_{i,\sim}$ *and* $\Gamma_r^R = e_{j,\sim}$, $(j < i)$, *then one needs* $\Gamma_{r+1}^R = e_{i,\sim}$ *and* $\Gamma_{r+1}^L = e_{i,\sim} \oplus e_{i+5\sim 5}$ *for the approximation to have a non-zero correlation. Similarly, for the decryption round function of TEA, if the input mask and the output mask for round* r *are* $(\Gamma_r^L|\Gamma_r^R)$ *and* $(\Gamma_{r+1}^L|\Gamma_{r+1}^R)$, *respectively. If* $\Gamma_r^R = e_{i,\sim}$ *and* $\Gamma_r^L = e_{j,\sim}$, $(j < i)$, *then we have* $\Gamma_{r+1}^L = e_{i,\sim}$ *and* $\Gamma_{r+1}^R = e_{i,\sim} \oplus e_{i+5\sim 5}$.

4.5 Zero correlation approximations for 14 and 15 rounds of TEA/XTEA

With the one-round property of linear approximation in TEA and XTEA derived in the previous subsection, we can identify zero correlation approximations over 14 and 15 rounds of both TEA and XTEA.

Proposition 4. *Over 15 rounds of TEA and XTEA, any linear approximation with input mask* $(\Gamma_1^R|\Gamma_1^L) = (1|0)$ *and output mask* $(\Gamma_{15}^R|\Gamma_{15}^L) = (0|e_{1,\sim})$ *has a correlation of exactly zero. Moreover, over 14 rounds of TEA and XTEA, any linear approximation with input mask* $(\Gamma_1^R|\Gamma_1^L) = (1|0)$ *and output mask* $(\Gamma_{14}^R|\Gamma_{14}^L) = (e_{1,\sim}|\bar{e}_{5,\sim})$ *has zero correlation.*

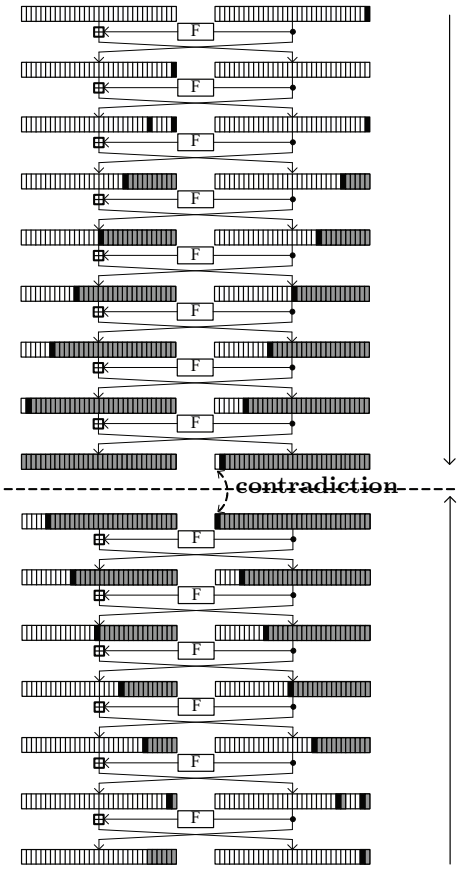

Fig. 3. Zero correlation linear approximation for 14-round TEA and XTEA (grey – undefined bits, black – bits set to 1)

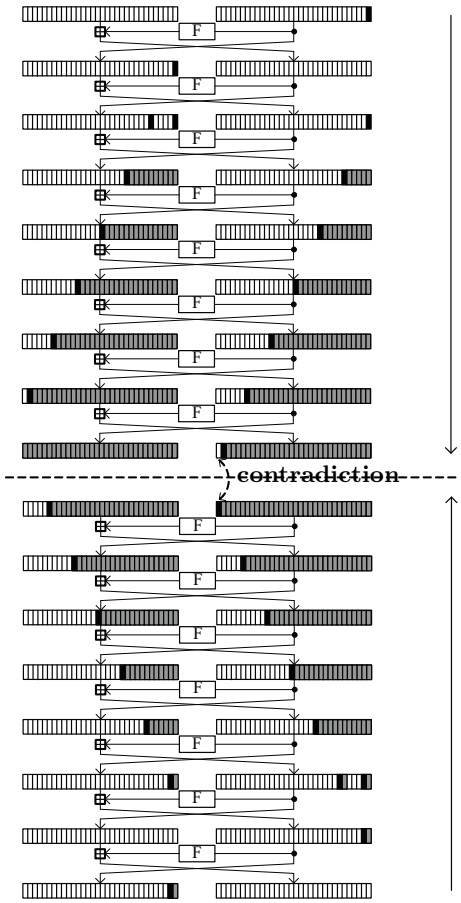

Fig. 4. Zero correlation linear approximation for 15-round TEA and XTEA (grey – undefined bits, black – bits set to 1)

Proof. First, we follow the linear approximation in the forward direction. From $\Gamma_1^L = 0$ and $\Gamma_1^R = 1$, it is obtained that $\Gamma_2^L = 0$ and $\Gamma_2^R = 1$, then we get $\Gamma_3^L = 1 \oplus (1 << 5)$ and $\Gamma_3^R = 1$. From Property 2, $\Gamma_3^L = 1 \oplus (1 << 5)$ and $\Gamma_3^R = 1$, then we have $\Gamma_4^R = e_{5,\sim}$ and $\Gamma_4^L =$

$e_{5,\sim} \oplus e_{5+5\sim 5} \oplus 1 = e_{10,\sim}$. Similarly, we get $(\Gamma_5^R|\Gamma_5^L) = (e_{10,\sim}|e_{15,\sim})$, $(\Gamma_6^R|\Gamma_6^L) = (e_{15,\sim}|e_{20,\sim})$, $(\Gamma_7^R|\Gamma_7^L) = (e_{20,\sim}|e_{25,\sim})$, $(\Gamma_8^R|\Gamma_8^L) = (e_{25,\sim}|e_{30,\sim})$ and $(\Gamma_9^R|\Gamma_9^L) = (e_{30,\sim}|?)$.

Second, we follow the 7-round linear approximation in the backward direction. From $\Gamma_{16}^L = e_{1,\sim}$ and $\Gamma_{16}^R = 0$, we can derive that $(\Gamma_{15}^R|\Gamma_{15}^L) = (e_{1,\sim}|0)$, $(\Gamma_{14}^R|\Gamma_{14}^L) = (e_{1,\sim} \oplus e_{6\sim 5}|e_{1,\sim})$, $(\Gamma_{13}^R|\Gamma_{13}^L) = (e_{11,\sim}|e_{6,\sim})$, $(\Gamma_{12}^R|\Gamma_{12}^L) = (e_{16,\sim}|e_{11,\sim})$, $(\Gamma_{11}^R|\Gamma_{11}^L) = (e_{21,\sim}|e_{16,\sim})$, $(\Gamma_{10}^R|\Gamma_{10}^L) = (e_{26,\sim}|e_{21,\sim})$ and $(\Gamma_9^R|\Gamma_9^L) = (e_{31,\sim}|e_{26,\sim})$.

From the forward direction, the most significant bit of Γ_9^R has to be zero, and from the backward direction, the most significant bit of Γ_9^R has to be one. This yields a contradiction and shows that there are no characteristics for this linear approximation. By the sufficient condition of [5] for constructing zero correlation linear approximations, this is enough for the approximation to have correlation zero. So the linear approximation for 15-round TEA and XTEA with the input mask $(1|0)$ and the output mask $(0|e_{1,\sim})$ has zero correlation. By restricting this linear approximation to 14 rounds and adding several undefined bits to the output mask, one gets all the claims of the proposition. □

There are only 2 zero correlation linear approximations of this form over 15 rounds. We note however that there are 2^7 different zero correlation linear approximations over 14 rounds of both TEA and XTEA. They can be generated by setting the undefined bits (depicted in gray in Figure 3 and Figure 4) to different values.

5 Zero correlation linear cryptanalysis of round-reduced (X)TEA

5.1 Key recovery for 21 rounds of TEA

For the cryptanalysis of 21-round TEA, we use the 14-round zero correlation approximations of the type depicted in Figure 3 of Subsection 4.5. The availability of many such approximations allows us to use the data complexity reduction technique of Section 3.

We place the 14-round zero correlation linear approximations in the middle of the 21-round TEA. It covers rounds 5 to 18. Following the procedure outlined in Subsection 2.3, up to the boundaries of the linear approximations, we partially encrypt over the 4 first rounds 1 to 4 and partially decrypt over the 3 last rounds 19 to 21. The attack is illustrated in Figure 5.

The linear approximations involve 9 state bits: R_5^0, $R_{19}^{1\sim 0}$, and $L_{19}^{5\sim 0}$. In the corresponding 9 bits of the input and output masks, only 7 can take on 0 and 1 values: $\Gamma_{19}^{R\,0}$ and $\Gamma_{19}^{L\,5\sim 0}$. For the evaluation of the linear approximations from a plaintext-ciphertext pair, we guess 54 key bits $K_0^{15\sim 0}$, $K_1^{15\sim 0}$, $K_2^{10\sim 0}$, and $K_3^{10\sim 0}$. The attack flow is as follows given N known plaintext-ciphertext pairs.

For each possible guess of the 54-bit subkey $\kappa = (K_0^{15\sim 0}|K_1^{15\sim 0}|K_2^{10\sim 0}|K_3^{10\sim 0})$:

1. Allocate a 128-bit counter W and set it to zero. W will contain the χ^2 statistic for the subkey guess κ.
2. Allocate a 64-bit counter $V[x]$ for each of 2^9 possible values of
$$x = (R_5^0|R_{19}^{1\sim 0}|L_{19}^{5\sim 0})$$
and set it to 0. $V[x]$ will contain the number of times the partial state value x occurs for N texts.
3. For each of N plaintext-ciphertext pairs: partially encrypt 4 rounds and partially decrypt 3 rounds, obtain the 9-bit value for $x = (R_5^0|R_{19}^{1\sim 0}|L_{19}^{5\sim 0})$ and add one to the counter $V[x]$.
4. For each of 2^7 zero correlation linear approximations:
   (a) Set the 64-bit counter U to zero.

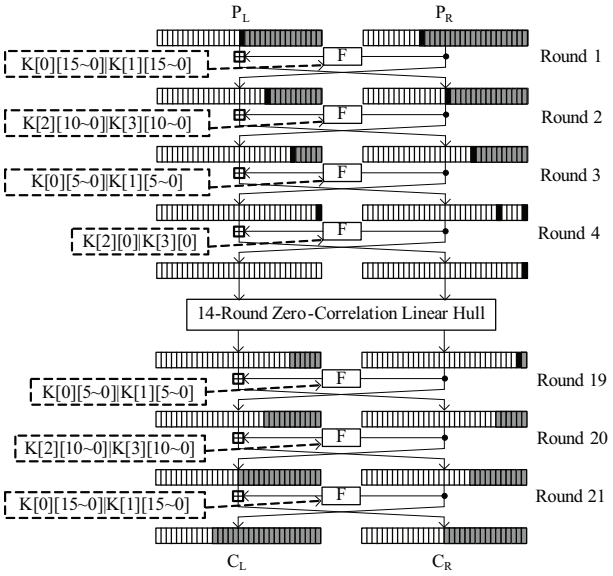

Fig. 5. Key recovery for 21 rounds of TEA. For the estimation of correlation, grey and black bits need to be computed and white bits are irrelevant. Uses the zero correlation approximation of Figure 3.

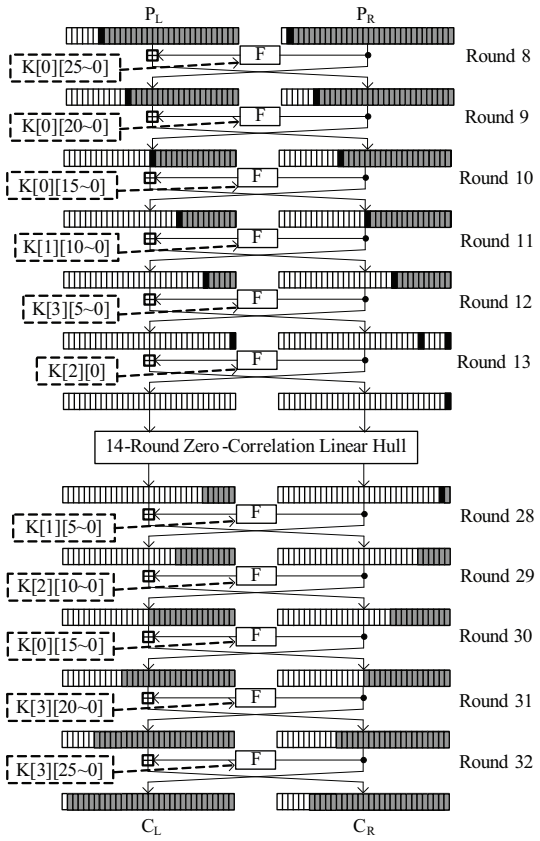

Fig. 6. Key recovery for 25 rounds of XTEA For the estimation of correlation, grey and black bits need to be computed and white bits are irrelevant. Uses the zero correlation approximation of Figure 3.

(b) For 2^9 values of x, verify if the linear approximation holds. If so, add $V[x]$ to U.
(c) $W = W + (2 \cdot U/N - 1)^2$.

5. If $W < t$, then κ is a possible subkey candidate and all cipher keys it is compatible with are tested exhaustively against a maximum of 3 plaintext-ciphertext pairs.

The correct 54-bit subkey κ is likely to be among the candidates with the χ^2 statistic W lower than the threshold $t = \sigma_0 \cdot z_{1-\beta_0} + \mu_0 = \frac{\sqrt{2l}}{N} \cdot z_{1-\beta_0} + \frac{l}{N} = \frac{\sqrt{2 \cdot 2^7}}{N} \cdot z_{1-\beta_0} + \frac{2^7}{N}$, see Subsection 3.1 with its Proposition 1 as well as Theorem 1.

In this attack, we set $\beta_0 = 2^{-2.7}$, $\beta_1 = 2^{-4.49}$ and get $z_{1-\beta_0} = 1$, $z_{1-\beta_1} = 1.7$. Note once again that $n = 64$ and $\ell = 2^7$. Theorem 1 suggests the data complexity of $N = 2^{62.62}$ known plaintext-ciphertexts with those parameters. The decision threshold is $t = 2^{-55.56}$.

The computational complexity is dominated by Steps 3 and 5. The computational complexity T_3 of Step 3 is 2^{54} times 7 half-round encryptions for each of N texts. This gives $T_3 = 2^{54} \cdot 2^{62.62} \cdot 7 \cdot 0.5/21 = 2^{114.03}$ 21-round TEA encryptions.

One in $1/\beta_1 = 2^{4.49}$ keys is expected to survive the test against zero correlation. The remaining key space is be covered by exhaustive search which is performed in Step 5. The computational complexity T_5 of Step 5 is about $T_5 = 2^{126-4.49} = 2^{121.51}$ 21-round encryptions using the equivalent key property. T_5 dominates the total computational complexity.

Summarizing the attack, its computational complexity is about $2^{121.51}$, data complexity is about $2^{62.62}$ known plaintext-ciphertext pairs, and the memory complexity is negligible. The success probability is about 0.846.

5.2 Key recovery for 25-round XTEA

Similarly to the attack on 21 rounds of TEA provided in the previous subsection, we use the 14-round zero correlation linear approximation depicted in Figure 3 to attack 25-round XTEA. Note that the attack covers rounds 8 to 32. It is illustrated in Figure 6. The linear approximations are placed in rounds 14 to 27. We partially encrypt 6 rounds (8 to 13) and partially decrypt 5 rounds (28 to 32) to evaluate the parity of approximations.

The linear approximations involve 9 bits and in the corresponding 9 bits of the input and output masks, again only 7 can take on 0 and 1 values: $\Gamma_{28}^{R\,0}$ and $\Gamma_{28}^{L\,5\sim0}$. For the evaluation of the linear approximations from a plaintext-ciphertext pair, we guess altogether 74 key bits $K_0^{25\sim0}$, $K_1^{10\sim0}$, $K_2^{10\sim0}$, and $K_3^{25\sim0}$. The attack itself is similar to that on 21-round TEA.

For each possible 63-bit value of $(K_0^{25\sim0}|K_1^{10\sim0}|K_3^{25\sim0})$:

1. Allocate and set to zero the 32-bit counter $V_1[x]$ for each of 2^{30} possible values of
$$x = (R_{13}^0|R_{13}^5|L_{13}^0|R_{30}^{10\sim0}|L_{30}^{15\sim0}).$$
2. For each of N plaintext-ciphertext pairs: partially encrypt 5 rounds and partially decrypt 3 rounds, obtain 30-bit $x = (R_{13}^0|R_{13}^5|L_{13}^0|R_{30}^{10\sim0}|L_{30}^{15\sim0})$, and add one to $V_1[x]$.
3. For each possible 11 bits value of $K_2^{10\sim0}$:
   (a) Allocate and set to zero a 128-bit counter W.
   (b) Allocate and set to zero a 64-bit counter $V_2[y]$ for each of 2^9 possible values of
   $$y = (R_{14}^0|L_{28}^{5\sim0}|R_{28}^{1\sim0}).$$
   (c) Encrypt one round and decrypt two rounds for 2^{30} values for x to get 9 bits of y and add $V_1[x]$ to $V_2[y]$.
   (d) For each of 2^7 zero correlation linear approximations:
      i. Set the 64-bit counter U to zero.
      ii. For 2^9 values of y, verify if the linear approximation holds. If so, add $V_2[y]$ to the counter U.
      iii. $W = W + (2 \cdot U/N - 1)^2$.
   (e) If $W < t$, then κ is a possible subkey candidate and all cipher keys it is compatible with are tested exhaustively against a maximum of 3 plaintext-ciphertext pairs.

The correct 74-bit subkey is likely to be among the candidates with the χ^2 statistic W lower than the threshold t. As we again set $\beta_0 = 2^{-2.7}$ and $\beta_1 = 2^{-4.49}$, we obtain $N = 2^{62.62}$ and $t = 2^{-55.56}$.

The computational complexity is dominated by Step 2 and checking for false positives in Step 3(e). T_2 of Step 2 is constituted by $2^{63}N$ computations of 5 rounds of 25-round XTEA and by $2^{63}N$ increments in the memory of 2^{30} 32-bit counters. Assuming that one increment of a memory cell costs one XTEA round, we obtain $T_2 = 2^{63} \cdot 2^{62.62} \cdot (5/25 + 1/25) = 2^{123.56}$. In Step 3(e), the remaining $T_{3(e)} = 2^{128-4.49} = 2^{123.51}$ keys can checked exhaustively by the same number of 25-round XTEA encryptions. Thus, the overall computational complexity is about $T_2 + T_{3(e)} = 2^{123.56} + 2^{123.51} = 2^{124.53}$ 25-round XTEA encryptions. The memory complexity is 2^{30} 32-bit words. Again, the data complexity is about $2^{62.62}$ known plaintext-ciphertext pairs, and the success probability is about 0.846.

5.3 Attacking more rounds with the full codebook

The attacks in the previous subsections use 14-round zero correlation linear approximations to enable data complexity reduction. As we only identified 2 15-round approximations, we cannot use this longer property to attack more rounds and still get a non-negligible decrease in the number of texts required. By taking advantage of the full codebook, we are however able to perform key recovery for up to 23 rounds of TEA and up to 27 rounds of XTEA, see Appendix D.

6 Conclusions

In this paper, we have demonstrated a technique for data complexity reduction for the promising recent zero correlation linear cryptanalysis which is based on linear approximations holding with a probability of exactly 1/2. This attack vector can be seen as the counterpart of the successful impossible differential cryptanalysis in the domain of linear cryptanalysis. Using ℓ linear approximations, we are able to reduce the data complexity to $\mathcal{O}(2^n/\sqrt{\ell})$, where n is the block size of the cipher.

As an application, we show 14- and 15-round linear approximations with correlation zero for round-reduced TEA and XTEA. Based on those, we propose key recovery attacks on 21-round TEA and 25-round XTEA with data complexity $2^{62.62}$ as well as on 23-round TEA and 27-round XTEA by taking advantage of all 2^{64} texts. All four attacks are the best key recoveries for both TEA and XTEA published to date in the single secret key setting. For these ciphers, our zero correlation linear attacks outperform their differential counterpart (impossible differential attacks), among other techniques.

These two contributions make the zero correlation linear cryptanalysis one of the major cryptanalytic techniques available today for attacking and evaluating symmetric-key ciphers.

Acknowledgements. We would like to thank Vincent Rijmen and Gregor Leander for insightful discussions. Andrey Bogdanov is postdoctoral fellow of the Fund for Scientific Research - Flanders (FWO). This work has been supported in part by the IAP Programme P6/26 BCRYPT of the Belgian State, by the European Commission under contract number ICT-2007-216676 ECRYPT NoE phase II, by KU Leuven-BOF (OT/08/027), by the Research Council KU Leuven (GOA TENSE), by NSFC Projects (No.61133013, No.61070244 and No.60931160442) as well as Outstanding Young Scientists Foundation Grant of Shandong Province (No.BS2009DX030).

References

1. E. Biham, A. Biryukov, A. Shamir: Cryptanalysis of Skipjack Reduced to 31 Rounds Using Impossible Differentials. In: EUROCRYPT'99, LNCS, pp. 12–23, Springer-Verlag, 1999.
2. E. Biham, O. Dunkelman, N. Keller: Related-Key Impossible Differential Attacks on 8-Round AES-192. In: CT-RSA'06, LNCS, pp. 21–33, Springer-Verlag, 2006.
3. E. Biham, A. Shamir: Differential Cryptanalysis of DES-like Cryptosystems. In: CRYPTO'90, LNCS, pp. 2–21, Springer-Verlag, 1990.
4. A. Bogdanov, D. Khovratovich, C. Rechberger: Biclique Cryptanalysis of the Full AES. In: ASIACRYPT'11, LNCS, pp. 344–371, Springer-Verlag, 2011.
5. A. Bogdanov, V. Rijmen: Zero Correlation Linear Cryptanalysis of Block Ciphers, IACR Eprint Archive Report 2011/123, March 2011.
6. J. Borst, L. R. Knudsen, V. Rijmen: Two Attacks on Reduced IDEA. In: EUROCRYPT'97, LNCS, pp. 1–13, Springer-Verlag, 1997.
7. C. Bouillaguet, O. Dunkelman, G. Leurent, P.-A. Fouque: Another Look at Complementation Properties. In: FSE 2010, LNCS, vol. 6147, pp. 347–364, 2010.
8. J. Chen, M. Wang, B. Preneel: Impossible Differential Cryptanalysis of Lightweight Block Ciphers TEA, XTEA and HIGHT. IACR Eprint Archive Report 2011/616, 2011.
9. B. Collard and F.-X. Standaert: Experimenting Linear Cryptanalysis. In P. Junod, A. Canteaut (eds.) *Advanced Linear Cryptanalysis of Block and Stream Ciphers*, vol, 7 of *Cryptology and Information Security Series*. IOS Press, 2011.
10. B. Collard, F.-X. Standaert, J.-J. Quisquater: Improving the Time Complexity of Matsui's Linear Cryptanalysis. In: ICISC'07, LNCS, vol. 4817, pp 77–88, Springer-Verlag, 2007.
11. J. Daemen, R. Govaerts, J. Vandewalle: Correlation Matrices. In: FSE 1994, LNCS, vol. 1008, pp. 275–285, Springer-Verlag, 1995.
12. J. Daemen, V. Rijmen: Probability distributions of correlations and differentials in block ciphers. Journal on Mathematical Cryptology 1(3), pp. 221-242, 2007.
13. J. Daemen, V. Rijmen. *The Design of Rijndael: AES – The Advanced Encryption Standard*. Springer-Verlag, 2002.

14. H. Demirci and A.A. Selçuk: A Meet-in-the-Middle Attack on 8-Round AES. In: FSE'08, LNCS, vol. 5086, pp. 116–126, Springer-Verlag, 2008.
15. O. Dunkelman, N. Keller: An Improved Impossible Differential Attack on MISTY1. In: ASIACRYPT'08, LNCS, vol. 5350, pp. 441–454, Springer-Verlag, 2008.
16. J. Etrog, M. J. B. Robshaw: On Unbiased Linear Approximations. In ACISP'10, LNCS, vol. 6168, pp. 74–86. Springer-Verlag, 2010.
17. W. Feller: *An Introduction to Probability Theory and Its Applications*, vol. 1, Wiley & Sons, 1968.
18. O. Dunkelman, N. Keller, A. Shamir: Improved Single-Key Attacks on 8-Round AES-192 and AES-256. In: ASIACRYPT'10, LNCS, vol. 6477, pp. 158–176, Springer-Verlag, 2010.
19. P. Hoel, S. Port, C. Stone: *Introduction to Probability Theory*, Brooks Cole, 1972.
20. S. Hong, D. Hong, Y. Ko, D. Chang, W. Lee, S. Lee: Differential Cryptanalysis of TEA and XTEA. In: ICISC'03, LNCS, vol. 2971, pp. 402–417, Springer-Verlag, 2004.
21. P. Junod: On the Complexity of Matsui's Attack. In: SAC'01, LNCS, vol. 2259, pp. 199–211, Springer-Verlag, 2001.
22. J.-P. Kaps: Chai-Tea, Cryptographic Hardware Implementations of xTEA. In: INDOCRYPT 2008, LNCS, vol. 5365, pp. 363–375, Springer-Verlag, 2008.
23. J. Kelsey, B. Schneier, D. Wagner: Key-Schedule Cryptoanalysis of IDEA, G-DES, GOST, SAFER, and Triple-DES. In: CRYPTO 1996, LNCS, vol. 1109, pp. 237–251, Springer-Verlag, 1996.
24. J. Kelsey, B. Schneier, D. Wagner: Related-key Cryptanalysis of 3-WAY, Biham-DES,CAST, DES-X, NewDES, RC2, and TEA. In: ICICS'97, LNCS, vol. 1334, pp. 233–246, Springer-Verlag, 1997.
25. E. Lee, D. Hong, D. Chang, S. Hong, J. Lim: A Weak Key Class of XTEA for a Related-Key Rectangle Attack. In: VIETCRYPT 2006, LNCS, vol. 4341, pp. 286–297, Springer-Verlag, 2006.
26. J. Lu: Related-key rectangle attack on 36 rounds of the XTEA block cipher. International Journal of Information Security 8(1), 1-11 (2009)
27. J. Lu, J. Kim, N. Keller, O. Dunkelman: Improving the Efficiency of Impossible Differential Cryptanalysis of Reduced Camellia and MISTY1. In: CT-RSA'08, LNCS, vol. 4964, pp. 370–386, Springer-Verlag, 2008.
28. J. Lu, O. Dunkelman, N. Keller, J. Kim: New Impossible Differential Attacks on AES. In: INDOCRYPT'08, LNCS, vol. 5365, pp. 279–293, Springer-Verlag, 2008.
29. H. Mala, M. Dakhilalian, V. Rijmen, M. Modarres-Hashemi: Improved Impossible Differential Cryptanalysis of 7-Round AES-128. In: INDOCRYPT'10, LNCS, vol. 6498, pp. 282–291, Springer-Verlag, 2010.
30. M. Matsui: Linear cryptanalysis method for DES cipher. In: EUROCRYPT'93, LNCS, vol. 765, pp. 386–397, Springer-Verlag, 1993.
31. M. Matsui: The First Experimental Cryptanalysis of the Data Encryption Standard. In: CRYPTO'94, LNCS, vol. 839, pp. 1–11, Springer-Verlag, 1994.
32. D. Moon, K. Hwang, W. Lee, S. Lee, J. Lim,: Impossible Differential Cryptanalysis of Reduced Round XTEA and TEA. In: FSE 2002, LNCS, vol. 2365, pp. 49–60, Springer-Verlag, 2002.
33. R.M. Needham, D.J. Wheeler: Tea extensions. Technical report, Computer Laboratory, University of Cambridge, October 1997, http://www.cix.co.uk/∼ klockstone/xtea.pdf
34. K. Nyberg: Correlation theorems in cryptanalysis. Discrete Applied Mathematics, 111(1-2):177–188, 2001.
35. L. O'Connor: Properties of Linear Approximation Tables. In: FSE 1994, LNCS, vol. 1008, pp. 131–136, Springer-Verlag, 1995.
36. A. Röck, K. Nyberg: Exploiting Linear Hull in Matsui's Algorithm 1. WCC'11, 2011.
37. G. Sekar, N. Mouha, V. Velichkov, B. Preneel: Meet-in-the-Middle Attacks on Reduced-Round XTEA. In: CT-RSA 2011, LNCS, vol. 6558, pp. 250-267, Springer-Verlag, 2011.
38. A.A. Selçuk: On Probability of Success in Linear and Differential Cryptanalysis. Journal of Cryptology, Volume 21(1), pp. 131–147, Springer-Verlag, 2008.
39. T. Shirai, K. Shibutani, T. Akishita, S. Moriai, T. Iwata: The 128-Bit Blockcipher CLEFIA (Extended Abstract). In: FSE'07, LNCS, vol.4593, pp. 181–195. Springer-Verlag, 2007.
40. M. Steil: 17 Mistakes Microsoft Made in the Xbox Security System. Chaos Communication Congress 2005, 2005. http://events.ccc.de/congress/2005/fahrplan/events/559.en.html
41. Y. Tsunoo, E. Tsujihara, M. Shigeri, T. Saito, T. Suzaki, H. Kubo. Impossible Differential Cryptanalysis of CLEFIA. In: FSE'08, LNCS, vol. 5086, pp. 398–411, Springer-Verlag, 2008.
42. S. Vaudenay. *Decorrelation: A Theory for Block Cipher Security.* J. Cryptology, 16(4):249–286, Springer-Verlag, 2003.
43. D.J. Wheeler, R.M. Needham: TEA, a Tiny Encryption Algorithm. In: FSE'94, LNCS, vol. 1008, pp. 363–366, Springer-Verlag, 1995.
44. Y. Zheng, T. Matsumoto, H. Imai: On the Construction of Block Ciphers Provably Secure and Not Relying on Any Unproved Hypotheses. In: CRYPTO'89, LNCS, vol. 435, pp. 461–480, Springer-Verlag, 1989.

A Proof of Proposition 1 (distinguishing distributions)

Proof. Again, first assume that $\mu_0 < \mu_1$. The error probabilities β_0 and β_1 can be derived from the value of threshold t and the CDFs of the two normal distributions:

$$\begin{aligned} \beta_0 &= 1 - \Phi_{\mu_0,\sigma_0}(t), \\ \beta_1 &= \Phi_{\mu_1,\sigma_1}(t), \end{aligned} \tag{7}$$

where Φ_{μ_i,σ_i} is the CDF of the respective normal distribution. (7) can be rewritten as follows using the CDF of the standard normal distribution:

$$\begin{aligned} \beta_0 &= 1 - \Phi_{0,1}\left(\tfrac{t-\mu_0}{\sigma_0}\right), \\ \beta_1 &= \Phi_{0,1}\left(\tfrac{t-\mu_1}{\sigma_1}\right). \end{aligned} \tag{8}$$

By going to quantiles in (8), one obtains

$$\begin{aligned} z_{1-\beta_0} &= \tfrac{t-\mu_0}{\sigma_0}, \\ z_{\beta_1} \quad &= \tfrac{t-\mu_1}{\sigma_1}. \end{aligned}$$

Expressing and equating t in the two cases yields:

$$\mu_0 + \sigma_0 z_{1-\beta_0} = \mu_1 + \sigma_1 z_{\beta_1}$$

and, eventually, recalling that $z_{\beta_1} = -z_{1-\beta_1}$ gives the relation

$$\frac{\sigma_0 z_{1-\beta_0} + \sigma_1 z_{1-\beta_1}}{\mu_1 - \mu_0} = 1. \tag{9}$$

Considering $\mu_0 > \mu_1$ yields a change of denominator in (9) to $\mu_0 - \mu_1$. The claim of the theorem follows. □

B Proof of Property 1 (modular addition)

Proof. We denote the i-th bit for x, y and z as x_i, y_i and z_i, $0 \leq i \leq n-1$, respectively. From the modular addition, we have

$$\begin{aligned} &z_0 = x_0 \oplus y_0, c_0 = 0, \\ &z_1 = x_1 \oplus y_1 \oplus c_1, c_1 = f_1(x_0, y_0), \\ &\cdots, \\ &z_i = x_i \oplus y_i \oplus c_i, c_i = f_i(x_0, x_1, \ldots, x_{i-1}, y_0, y_1, \ldots, y_{i-2}). \cdots, \\ &z_{n-1} = x_{n-1} \oplus y_{n-1} \oplus c_{n-1}, c_{n-1} = f_{n-1}(x_0, x_1, \ldots, x_{n-2}, y_0, y_1, \ldots, y_{n-2}), \end{aligned}$$

where c_i is the carrying bit of the i-th bit and f_i is the non-linear carrying function of the i-th bit. From the above equations, the linear approximations with non-zero bias have the following form:

$$\begin{aligned} &z_0 = x_0 \oplus y_0, \\ &z_1 = x_1 \oplus y_1[\oplus x_0 \oplus y_0], \\ &z_2 = x_2 \oplus y_2[\oplus x_1 \oplus y_1 \oplus x_0 \oplus y_0], \\ &\cdots, \\ &z_i = x_i \oplus y_i[\oplus x_{i-1} \oplus y_{i-1} \oplus \ldots \oplus x_0 \oplus y_0], \\ &\cdots, \\ &z_{n-1} = x_{n-1} \oplus y_{n-1}[\oplus x_{n-2} \oplus y_{n-2} \oplus \ldots \oplus x_0 \oplus y_0], \end{aligned}$$

where the terms in the square brackets are optional. So any linear approximation with non-zero bias will be produced from any one or the combination from the above linear relations which can be denoted as follows,

$$z_i[\oplus z_{i-1} \oplus \cdots \oplus z_0] = x_i \oplus y_i[\oplus x_{i-1} \oplus y_{i-1} \oplus \ldots \oplus x_0 \oplus y_0].$$

If there is a linear approximation with the following form,

$$z_j \oplus z_i[\oplus z_{i-1} \oplus \cdots \oplus z_0] = x_i \oplus y_i[\oplus x_{i-1} \oplus y_{i-1} \oplus \ldots \oplus x_0 \oplus y_0], i < j < n. \tag{10}$$

We substitute the equation $z_j = x_j \oplus y_j \oplus c_j, c_j = f_i(x_0, x_1, \ldots, x_{j-1}, y_0, y_1, \ldots, y_{j-1})$ into Equation 10, we get

$$\begin{aligned} &x_j \oplus y_j \oplus f_i(x_0, x_1, \ldots, x_{j-1}, y_0, y_1, \ldots, y_{j-1}) \\ &\oplus z_i[\oplus z_{i-1} \oplus \cdots \oplus z_0] = x_i \oplus y_i[\oplus x_{i-1} \oplus y_{i-1} \oplus \ldots \oplus x_0 \oplus y_0], i < j < n. \end{aligned}$$

In the above equation, $x[j], x[j-1], \ldots x[i+1], y[j], y[j-1], \ldots y[i+1]$ are not related with $z_i, z_{i-1}, \ldots, z_0$, so they are independent random variables. The involved independent random variables will make the linear approximation Equation (10)be random, so the bias for Equation (10)will be zero. Similarly, the linear approximation with the following forms will also have zero bias,

$$\begin{aligned} z_i[\oplus z_{i-1} \oplus \cdots \oplus z_0] &= x[j] \oplus x_i \oplus y_i[\oplus x_{i-1} \oplus y_{i-1} \oplus \ldots \oplus x_0 \oplus y_0], i < j < n. \\ z_i[\oplus z_{i-1} \oplus \cdots \oplus z_0] &= y[j] \oplus x_i \oplus y_i[\oplus x_{i-1} \oplus y_{i-1} \oplus \ldots \oplus x_0 \oplus y_0], i < j < n. \end{aligned} \tag{11}$$

This means that the most non-zero significant bits for $\Gamma x, \Gamma y$ and Γz must be same, otherwise, the linear approximation will have zero bias. □

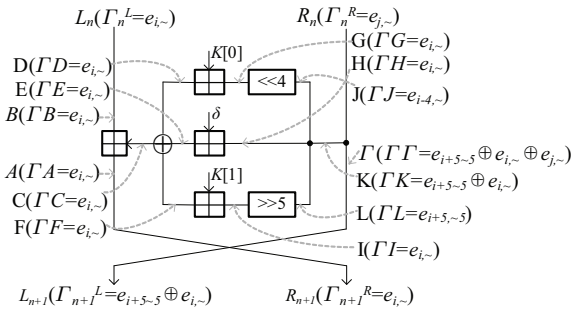

Fig. 7. Linear approximation of one TEA round

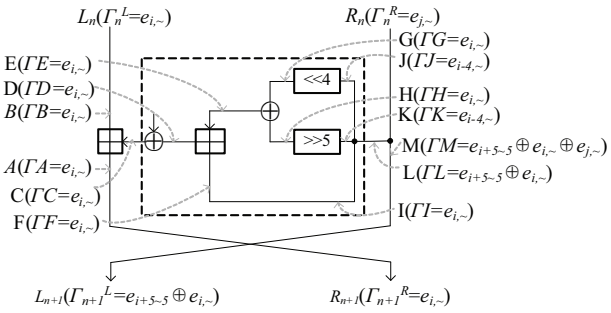

Fig. 8. Linear approximation of one XTEA round

C Proof of Property 2 (one round)

Proof. The linear approximation for the encryption round function of TEA and XTEA have been shown in Fig.7 and Fig. 8, we use the notations $A, B, C, D, E, F, G, H, I, J, K, L, M$ to denote the intermediate variables and the notation ΓX to denote the respective mask value for $X \in \{A, B, C, D, E, F, G, H, I, J, K, L, M\}$. Next, we will only give the proof for the linear approximation of TEA. The proof for XTEA is similar as that for TEA and we will not describe it due to the limited space.

From Theorem 1, $C = A + B$, $MA = e_{i,\sim}$, so we have $\Gamma B = e_{i,\sim}$ and $\Gamma C = e_{i,\sim}$. Then from Lemma 1 in [5] and $C = D \oplus E \oplus F$, we can get $\Gamma D = \Gamma E = \Gamma F = \Gamma C = e_{i,\sim}$. From

Theorem 1, $D = G + K[0]$ and $\Gamma D = e_{i,\sim}$, so $\Gamma G = e_{i,\sim}$; $E = H + \delta$ and $\Gamma E = e_{i,\sim}$, so $\Gamma H = e_{i,\sim}$; $F = I + K[1]$ and $\Gamma F = e_{i,\sim}$, so $\Gamma I = e_{i,\sim}$. As $G = J \lll 4$ and $I = L \ggg 5$, then $\Gamma J = \Gamma G \ggg 4 = e_{i-4,\sim}$ and $\Gamma L = \Gamma I \lll 5 = e_{i+5\sim 5}$. From Lemma 2 in [5], we have $\Gamma K = \Gamma G \oplus \Gamma H \oplus \Gamma I = \Gamma H \oplus \Gamma J \oplus \Gamma L = e_{i,\sim} \oplus e_{i-4,\sim} \oplus e_{i+5\sim 5} = e_{i,\sim} \oplus e_{i+5\sim 5}$. As $j < i$, $\Gamma M = \Gamma K \oplus \Gamma_n^R = e_{i,\sim} \oplus e_{i+5\sim 5} \oplus e_{j,\sim} = e_{i,\sim} \oplus e_{i+5\sim 5}$.

The proof for the linear approximation of the decryption round function can be proved in the same way. □

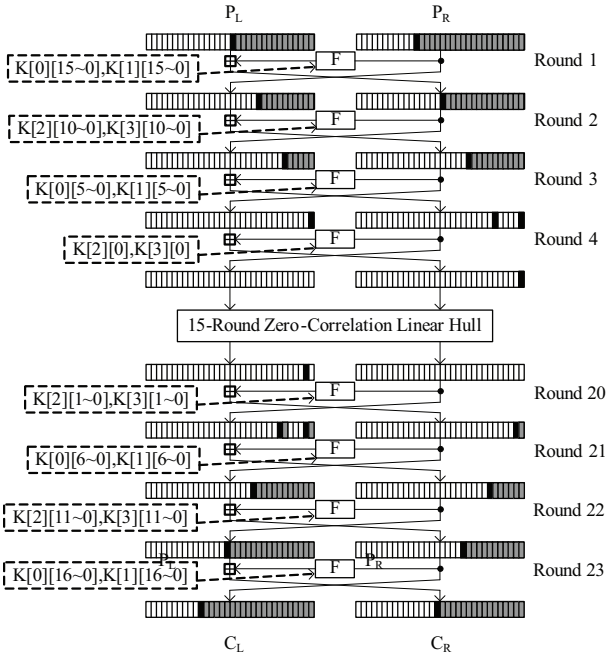

Fig. 9. Key recovery for 23 rounds of TEA For the estimation of correlation, grey and black bits need to be computed and white bits are irrelevant. Uses the zero correlation approximation of Figure 4.

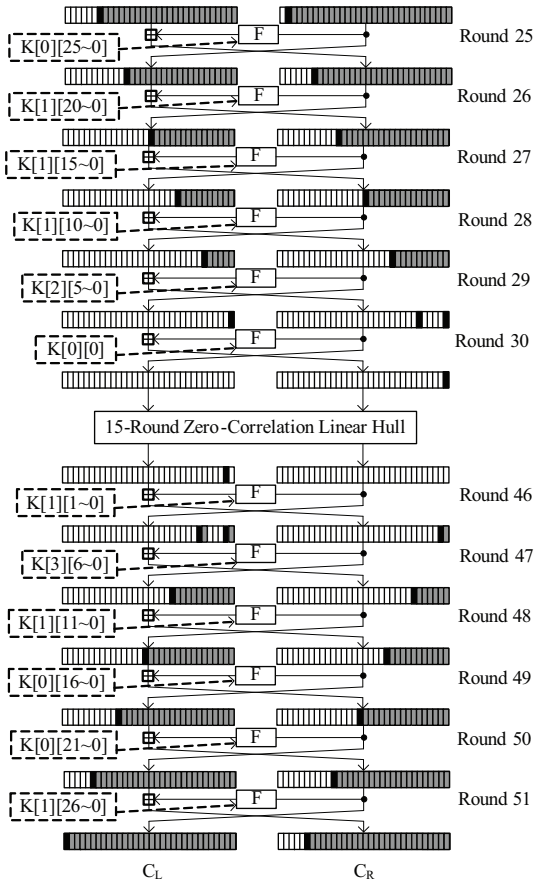

Fig. 10. Key recovery for 27 rounds of XTEA. For the estimation of correlation, grey and black bits need to be computed and white bits are irrelevant. Uses the zero correlation approximation of Figure 4.

D Key Recovery for More Rounds of (X)TEA with the Full Codebook

D.1 Key Recovery for 23 Rounds of TEA

We use the 15-round zero correlation linear approximations of Figure 4 to attack 23-round TEA, see Figure 9 for an illustration.

Now we use the basic zero correlation linear cryptanalysis with the full code book. To compute the parity of the approximation, we need to guess 58 bits: $(K_0^{16\sim 0}|K_1^{16\sim 0}|K_2^{11\sim 0}|K_3^{11\sim 0})$. For each guess, we partially encrypt 4 rounds and decrypt 4 rounds for the whole code book to get $R_5^0|L_{20}^1$ and verify if the equation holds. The computational complexity is dominated by those

computations: $2^{58+64} \cdot (1 + 0.5 \cdot 3 + 0.5 \cdot 4)/23 \simeq 2^{119.64}$. Memory complexity is negligible. Data complexity is 2^{64}. Success probability is 1.

D.2 Key Recovery for 27 Rounds of XTEA

We use the same 15-round zero correlation linear approximation to attack 27-round XTEA, see Figure 10 for the attack. Again, using the full codebook, we rely on the basic zero correlation linear cryptanalysis procedure of [5].

The attack is proceeded as follows:

For each possible 59 bits value of $(K_0^{25\sim 0}|K_1^{26\sim 0}|K_2^{5\sim 0})$:

- Allocate and set to zero the 64-bit counter $V[x]$ for each of 2^{22} possible values of

$$x = (R_{30}^5|R_{30}^0||L_{30}^0|R_{48}^{6\sim 0}|L_{48}^{11\sim 0}).$$

- Partially encrypt 5 rounds from round 25 and partially decrypt 4 rounds from round 51 for the whole code book, get 22-bit $(R_{30}^5|R_{30}^0||L_{30}^0|R_{48}^{6\sim 0}|L_{48}^{11\sim 0})$ and add one to $V[x]$.
- For each possible 7 bits value of $K_3^{6\sim 0}$:
  - Partially encrypt 1 round from round 30 and decrypt 2 rounds from 47 for 2^{22} possible values for x to get 2 bits of $(R_{31}^0|L_{46}^1)$ and verify if the linear approximation holds.
  - If the counter equals to zero, it means that the guessed value for key bits is right with high probability.

The computational complexity of the attack is dominated by the partial encryption and decryption: $2^{59} \cdot 2^{64} \cdot (2 + 0.5 \cdot 3 + 2 + 0.5 \cdot 2)/27 = 2^{120.71}$ 27-round XTEA encryptions. The memory complexity is 2^{23} 32-bit words. Data complexity is 2^{64}. Success probability is 1.