Automated Dynamic Cube Attack on Block Ciphers: Cryptanalysis of SIMON and KATAN

Zahra Ahmadian, Shahram Rasoolzadeh, Mahmoud Salmasizadeh, and Mohammad Reza Aref

Sharif University of Technology, Tehran, Iran.
{ahmadian,sh_rasoolzadeh}@ee.sharif.edu,{salmasi,aref}@sharif.edu

Abstract. A few work has ever been performed in cryptanalysis of block ciphers using cube attacks. This paper presents a new framework for an efficient key recovery attack on block ciphers based on cube technique. In this method, a cube tester is positioned at the middle of the cipher which is extended in two directions over the maximum possible upper and lower rounds, given that some subkey bits are guessed. It is shown that an automated algorithm for this dynamic cube attack on block ciphers can be realized. Furthermore, we show its effectiveness on two lightweight block ciphers KATAN and SIMON. Our results shows that this method can break 117 and 152 out of 254 rounds of KATAN-32 in non-full-codebook and full-codebook attack scenarios, respectively. In the case of SIMON32/64, we succeed to cryptanalyse 16 and 18 out of 32 rounds, by the same scenarios. Both results show that although this method does not outperform all the existing attacks on these two ciphers, it can absolutely compete with the well-established and mature methods of cryptanalysis of block ciphers, such as linear, differential and meet in the middle attack families.

Keywords: block cipher, cryptanalysis, cube attack, SIMON, KATAN

1 Introduction

In 2007, Vielhaber proposed the idea of Algebraic IV Differential Attack (AIDA) [1], which was then continued by Dinur and Shamir under the new title *Cube attack* [2]. In this attack, the cipher is treated as a black box whose input consists of a public tweakable variable (e.g. IV in stream ciphers and plaintext in block ciphers) and a secret fixed parameter, (the key of the algorithm). The main idea behind this attack is to find some information about the secret key of the algorithm by choosing all values for an appropriate subset of the public input variables and summing up all the corresponding outputs.

This technique can also introduce a distinguisher, the so called *cube tester* in [3], where instead of retrieving some information from the key bits, a non-random property of the cipher is discovered by such an oracle access to the cipher. *Dynamic* cube attack can be regarded as a more advanced version of the cube attack, in which a cube tester is employed not to detect a non-random property,

but to determine if a specific guess for a subset of key bits could be correct or not [4]. In spite of the classic cube attack where the attacker sees the cipher as the black box and does not use the structural information of the algorithm, dynamic cube attack make use of such information, hence it can potentially reach better results than classic cube attack.

Cube attack family has shown to be very effective in cryptanalysis of lightweight stream ciphers. Grain [5] and Trivium [6] are two lightweight stream ciphers, for both of which the most successful attacks are a kind of cube attack family [1,2,3,4,7,8,9]. In contrast to its significant results in cryptanalysis of lightweight stream ciphers, apart from a combination of cube attack and algebraic attack [10], it has not yet been applied to block ciphers[1].

In this paper, we aim to take the first steps for an efficient application of cube technique for cryptanalysis of block ciphers. The technique that we propose is apparently similar to dynamic cube attack on stream ciphers [4], since it makes use of some distingushers to discard some wrong guesses of the secret key. Furthermore, some specific functions of the cube variables and secret key bits should be assigned to the input variables, similar to the so called *dynamic* variables in [4]. But, the mechanism by which these functions are derived are totally different. The main approach in the dynamic cube attack on stream ciphers is to use the recursive description of the cipher's output function in order to nullify/simplify some appropriate intermediate variables which consequently leads to a simplified algebraic function for the output bit. This process demands a "complex process that can not be fully automated and involves manual work to analyse the cipher" [4]. However, the situation is completely different in block ciphers, where a key-only-dependent function, i.e. the key schedule, is always available in any round of the cipher which enables the cryptanalyst to do partial encryption/decryption anywhere in the cipher, conditioned that she guesses the relevant subkeys. This option is not available in stream ciphers in any case, where the secret key as well as the IV are loaded into the state of the cipher at the first step, then they are mixed during the initialization phase so a key-only dependent function would never be conceivable then.

In spite of the very complex manual procedure of the dynamic cube attack on stream ciphers, our attack can be fully automated. We put the r_c-round cube tester (distinguisher) in the middle of the cipher. Hence, in spite of all the cube attacks, the cube variables in our attack are not among the plaintext bits but they are basically some bits of an intermediate state. Then, we extend the attack to r_u rounds before, and r_l rounds after the distinguisher. Both the extensions potentially involve guessing some key bits. In addition, the backward r_u-round extension determines exactly what each bit of the plaintext must be (i.e. its precise description in the guessed key bits and cube variables).

[1] There are some instances of combinations of cube and side channel attacks [10,11,12,13] and cube and fault injection attacks [14] in the literature. These attacks are defined in the *leakage* and *fault injection* attack models, both of which are out of our attacker model.

Table 1. Result of Previous Attacks on Katan32 and Simon32/64

Algorithm	Type	Round	Time	Data	Memory	Ref.
KATAN32	Cube/Algebraic	79	14.72 min	20	-	[10]
	Conditional Differantial	78	2^{22}	2^{22}	-	[17]
	MITM ASR	110	2^{77}	138	$2^{75.1}$	[19]
	Differential	114	2^{77}	$2^{31.9}$	-	[18]
	Dynamic Cube	**116**	$\mathbf{2^{78.83}}$	$\mathbf{2^{19}}$	$\mathbf{2^{6}}$	**Sec. 4.2**
	Dynamic Cube	**117**	$\mathbf{2^{78.77}}$	$\mathbf{2^{27}}$	$\mathbf{2^{7}}$	**Sec. 4.2**
	MITM ASR	119	$2^{79.1}$	144	$2^{79.1}$	[19]
	Matchbox MITM	121	$2^{77.5}$	4	2^{5}	[20]
	Dynamic Cube	**152**	$\mathbf{2^{78.92}}$	$\mathbf{2^{32}}$	$\mathbf{2^{32}}$	**Sec. 4.2**
	Matchbox MITM	153	$2^{78.5}$	2^{5}	2^{76}	[20]
SIMON32/64	Linear	11	-	2^{23}	-	[23]
	Linear (Matsui's 1st Alg.)	13	2^{32**}	2^{32}	-	[21]
	Impossible Differential	13	$2^{50.1}$	2^{30}	2^{20}	[23]
	Linear (Matsui's 2nd Alg.)	16	2^{54}	2^{32}	-	[21]
	Differential	16	$2^{26.48}$	$2^{29.48}$	2^{16}	[25]
	Dynamic Cube	**16**	$\mathbf{2^{51.5}}$	$\mathbf{2^{19}}$	$\mathbf{2^{10}}$	**Sec. 5.2**
	Dynamic Cube	**18**	$\mathbf{2^{51}}$	$\mathbf{2^{32}}$	$\mathbf{2^{32}}$	**Sec. 5.2**
	Differential*	18	2^{46}	$2^{31.2}$	2^{15}	[23]
	Multiple Linear	18	2^{32}	2^{32}	-	[21]
	Impossible Differential	19	$2^{62.56}$	2^{32}	2^{44}	[22]
	Differential	19	2^{32}	2^{31}	-	[24]
	Differential	20	2^{31}	2^{31}	-	[26]
	Linear Hull	20	$2^{59.69}$	$2^{31.69}$	-	[21]
	Differential	21	2^{46}	2^{31}	-	[26]

* Its probability of success is 63%.

** Its needed time for computing one bit of key. So, the total time complexity is 2^{63}.

Having introduced our attack framework in a general view, we examine its efficiency and flexibility on two lightweight block ciphers: the recently proposed NSA cipher, SIMON32/64 [15], and KATAN-32 [16]. We found them good targets for our attack since both has a low-degree round function compensated by relatively large number of rounds, similar to stream cipher potential targets of cube attack [5,6].

Our attack algorithm is flexible enough to set the maximum allowable value for data and time complexities. So, we report our results in two scenarios: full-codebook and non-full-codebook attacks. In case of KATAN, we could analyse up to 152 and 117 rounds out of 254 rounds by full-codebook attack non-full-codebook scenarios, respectively which absolutely outperform the only instance of cube attack [10] which could break 79 rounds. In case of SIMON32/64, we could break 18-round and 16-round version of full 32-round cipher in full-codebook and non-full-codebook scenarios, respectively. Although our attacks does not yet reach the maximum rounds of the ciphers analysed by now, our results shows that in spite of the conventional view that cube attack is an appropriate tool just for cryptanalysis of stream ciphers, this newcomer attack can

absolutely compete with the well-established and accepted methods for cryptanalysis of block ciphers such as differential and meet in the middle attack families for KATAN-32 [17,18,19,20] and differential and linear attack families for SIMON32/64 [21,22,23,24] and overtake many of them.

This paper is organized as follows: In Section 2, we give some preliminaries and notations for cube attacks. In Section 3, the proposed framework for dynamic cube attack on block ciphers is explained in a general view. We present our results on KATAN-32 and Simon 32/64 in Sections 4 and 5, respectively. Finally we conclude our work in Section 6.

2 Preliminaries and Notations

Suppose $f : \{0,1\}^l \to \{0,1\}$ is the boolean function representing one output bit of the cipher based on the m-bit public input variable $P = \{p_{m-1}, ..., p_0\}$ and n-bit secret key $K = \{k_{n-1}, ..., k_0\}$ where $l = m + n$. Let $X = P \cup K$ be the set of all inputs of f. f can be represented as follows:

$$f(x_{l-1}, ..., x_0) = \sum_{i \in F_2^l} a_i \cdot x_{l-1}^{i_{l-1}} ... x_1^{i_1} x_0^{i_0} \tag{1}$$

where $\{i_{l-1}, ..., i_0\}$ is the binary representation of i, and a_i is a binary constant.

Now suppose that I is a certain index subset of $\{0, 1, ..., l-1\}$ of size d, namely *cube*. d is called the cube dimension. We denote the set of cube variables by $X_I = \{x_i | x_i \in I\}$. If we factorize function f by the monomial $t_I = \prod_{i \in I} x_i$, it can be written in the following form

$$f(x_{l-1}, ..., x_0) = t_I \cdot P_{S(I)} + q(x_{l-1}, ..., x_0) \tag{2}$$

where $P_{S(I)}$ is a polynomial that has no variable in common with t_I, and no monomial in q contains t_I. The $P_{S(I)}$ is called the *superpoly* of I.

The main idea behind all the variants of cube attack is to deal with the lower degree and simpler polynomial $P_S(I)$ rather than the potentially very complex polynomial f, where $X_I \subseteq P$. To do so, we enjoy the following property of Boolean functions:

$$\sum_{X_I \in \{0,1\}^d} f(x_{l-1}, ..., x_0) = P_{S(I)} \tag{3}$$

In other words, if the attacker query all the possible values of $X_I \in \{0,1\}^d$ from the encryption oracle where the other public variables are fixed, and sum up all the corresponding outputs, she will come up with the evaluation of $P_{S(I)}$ in which the other public variables have the same fixed values. The notations used in this paper are listed in Table 2.

3 Dynamic Cube Attack on Block ciphers

Apart from a distinguishing attack on hash function MD5 [3], all the major results published for the cube attack family implies its efficiency in crptanalysis

Table 2. Notations

Symbol	Definition
m/n	block/key size for block cipher
I	cube
d	cube dimension
X_I	set of cube variables $\{x_i \| i \in I\}$
$P_S(I)$	the superpoly for cube I
U/Z	index set for neutral/static bits
r_c	number of rounds in the cube tester part
r_u/r_l	number of rounds in the upper/lower extension part
$\mathcal{K}_u/\mathcal{K}_l$	set of key bits involved in the upper/lower extension part
$S^{(i)}$	intermediate state at the end of round i
S_j/S_J	bit j of intermediate state S, $S_J = \{S_j \| j \in J\}$
T	distinguisher bit
N	number of tests
$p^{(j)}(X_I, \mathcal{K}_u)$	j^{th} plaintext description in X_I and $\mathcal{K}_u$
$Time/Data$	time/(upper bound for) data complexity
$Time_{max}/Data_{max}$	maximum allowable time/data complexity
$enc(S^{(i)}, \mathcal{K})/dec(S^{(i)}, \mathcal{K})$	one-round symbolic encryption/decryption of $S^{(i)}$ under $\mathcal{K}$
$Enc(A, K, r_1, r_2)/$ $Dec(A, K, r_1, r_2)$	r_2-round partial encryption/decryption of $A \in \{0,1\}^m$ starting at round r_1 under key $K \in \{0,1\}^n$

of lightweight stream ciphers with a low degree round function compensated by a large number of initialization rounds. For Trivium [6], classic cube attack has led to cryptanalysis of the highest rounds analysed ever [2,3,7,8] and for Grain family [5], the best cryptanalytic results are reported by dynamic cube attack [4,9]. In spite of significant results in stream ciphers, no remarkable result on block ciphers using cube technique has appeared, by now. The only one is cryptanalysis of KATAN block cipher which analysed 79 rounds out of 254 round using a combination of cube and algebraic attack.

In this section we will present a framework for cryptanalysis of block ciphers using a kind of dynamic cube attack. We will show that, this method can also be applied efficiently in cryptanalysis block ciphers. Due to the algebraic essence of this attack, those block ciphers with a low-degree round function compensated by a large number of rounds are potentially good targets for this attack. SIMON32/64 [15] and KATAN-32 [16] are two instance of such ciphers that we have adopted to examine our method. For more details, see Tab 1.

3.1 Attack Framework

As for all methods of cube attack family, this attack proceeds in two phases: preprocessing phase and online phase. In the preprocessing phase, the attacker tries to find a cube tester for a subset of key bits as well as sufficiently many corresponding plaintext descriptions (in cube variables and that subset of key

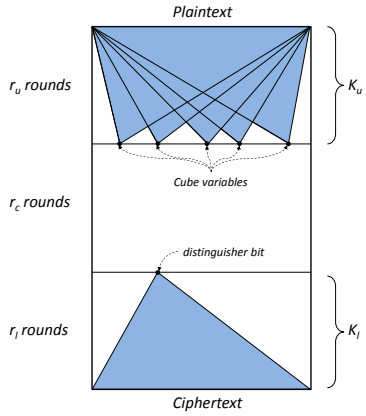

Fig. 1. An overview of the attack

bits). In the online phase, for each key guess, the attacker runs the cube tester by appropriate queries from the encryption oracle and checks if such collections of (P, C) pairs conform the cube tester. If so, the guessed key is treated as a correct key candidate to be rechecked then.

Preprocessing In this phase, inspiring from the distinguisher-based attacks on block ciphers, we position the cube tester at the *middle* of the cipher (e.g. exactly after round r_u) and *extend* it in two directions over the maximum possible upper and lower rounds. Such a position for the cube tester immediately implies that cube variables are defined among $S^{(r_u)}$ rather than the plaintext bits. Both the upper and lower extensions demands guessing some key bits. We denote the subset of key bits (or equivalent bits) that should be guessed in the upper extension by $\mathcal{K}_u$ and that of the lower extension by $\mathcal{K}_l$. A schematic view of the attack is shown in Fig. 1.

In general, the lower and upper extensions of a distinguisher for a block cipher is a straightforward procedure. At least, in none of the distinguisher-based attacks, the distinguisher itself is affected by those extensions. But in our cube attack, the upper extension procedure is not so straightforward insofar as it determines the role of the other-than-cube variables in $S^{(r_u)}$, hence it affects directly on the cube tester behavior. In other words, after determining the starting round of the distinguisher, r_u, and choosing the cube variables among $S^{(r_u)}$, we are not free enough to statically assign zero to the other variables (In fact, such a strategy is possible but so costly). It is actually the upper extension procedure that beside finding the specification of the plaintext bits in cube variables and $\mathcal{K}_u$, i.e. $P = p(X_I, \mathcal{K}_u)$, determines the status of the other-than-cube variables of $S^{(r_u)}$: statically zero or neutrally zero or one.

Having determined the status of all bits of $S^{(r_u)}$ whether cube, static or neutral, we find the maximum length cube tester for such input variables at

round r_u. Suppose that the length of the cube tester is r_c. So, the distinguisher bit T is one bit among the the intermediate state $S^{(r_u+r_c)}$. In our analysis we have assumed only deterministic cube testers, but it can be any function with a non-random easily detectable property. Once r_c is known, we find the maximum possible rounds that the lower extension part covers. The limiting factor here is the number of key bits that should be guessed i.e. $|\mathcal{K}_l|$. This extension is not as complicated as the upper extension.

Therefore, the total rounds attacked by this method would be $r = r_u + r_c + r_l$. An optimum partitioning is the one maximizing r. To get such a partitioning, Algorithm 1 of Appendix C is proposed. In this algorithm, the cube tester *slides* along a pre-defined range, i.e. $r_{u,min} \leq r_u \leq r_{u,max}$, for which it tests randomly chosen cube sets of dimension d at round r_u and returns the one, whose attack covers the maximum rounds $r = r_u + r_c + r_l$, satisfying the complexity constraints $Data < Data_{max}, Time < Time_{max}$. For each r_u and cube set X_I of dimenssion d, this algorithm proceeds in 3 subroutines: Upper-Extension-Subroutine, Cube-Tester-Subroutine and Lower-Extension-Subroutine.

Upper-Extension-Subroutine

The input parameters for this subroutine are $\{r_u, I\}$ and it is supposed to return the followings.

- the subset of (equivalent) key bits, i.e. $\mathcal{K}_u$, which should be guessed.
- the plaintext description $p(X_I, \mathcal{K}_u)$ in such a way that its r_u-round partial encryption yields $S^{(r_u)}$ where

$$S_i^{(r_u)} = \begin{cases} x_i + f_i(\mathcal{K}) & i \in I \\ f_i(\mathcal{K}) & i \notin I \end{cases}, \quad i = 0, \ldots, m-1 \tag{4}$$

  where $S_i^{(r_u)}$ is the i^{th} bit of $S^{(r_u)}$, x_i is a cube variable and f_i is any function of the key only.
- the set of neutral and static variables of $S^{(r_u)}$, namely U and Z, respectively. It means that $Z = \{0 \leq i < m | f_i(\mathcal{K}) = 0\}$ and $U = \{0 \leq i < m | f_i(\mathcal{K}) \neq 0\}$.

So, it returns $\{\mathcal{K}_u, p(X_I, \mathcal{K}_u), U, Z\}$. The procedure of this subroutine is given in Algorithm 2 of Appendix C, where all the variables including cube variables, keys and intermediate states have symbolic values.

To see how this subroutine works, first assume that the attacker is going to extend the upper part for one round ($r_u = 1$). At first, all the bits of $S^{(r_u)}$ are set to zero except those whose indexes are in I which are assigned symbolic cube variables $\{x_i | i \in I\}$. Then, the attacker partially decrypts $S^{(r_u)}$ for one round while setting $K = 0$ to get the pure dependency of $S^{(r_u-1)}$ to cube variables. Then, she partially encrypt the resulted $S^{(r_u-1)}$ again, this time taking $\mathcal{K}$ into account, to find out how the subkey of round $r_u - 1$ affects on $S^{(r_u)}$. There would be no need to guess any bits of $\mathcal{K}$ as far as all the bits of $S^{(r_u)}$ can be written consistent with (4).

But, once there is a state bit whose representation violates (4), e.g. those containing AND of a x_i with a key bit, she has to zero an appropriate internal

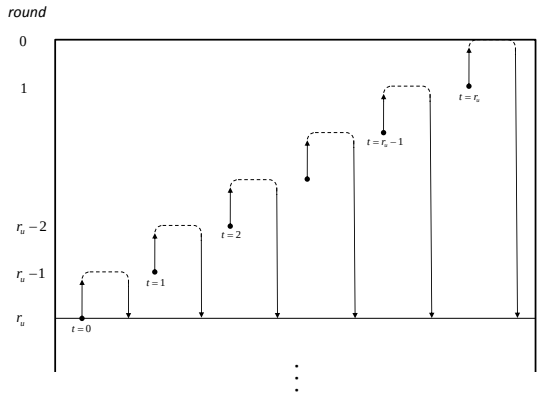

Fig. 2. schematic view of Upper-Extension-Subroutine

state bit to avoid such an event. This zero forcing is equivalent to one bit key guess which will be stored in $\mathcal{K}_u$ list. Having discovered such key bits, the attacker repeats the one-round partial decryption of $S^{(r_u)}$, this time with a subkey with nonzero symbolic value $\mathcal{K}_u$. So, $S^{(r_u-1)}$ would be a function of X_I and $\mathcal{K}_u$ while $S^{(r_u)}$ is a function of X_I and $\mathcal{K}$ consistent with (4).

The procedure for $r_u > 1$ is similar. Suppose that the attacker has already finished step t, $0 \leq t < r_u$ of upper extension. She has driven $S^{(r_u-t)}$ and a list $\mathcal{K}_u$ of key bits to be guessed by now, each of which corresponds to one internal state bit to be zero. $S^{(r_u-t)}$ is specified based on X_I and $\mathcal{K}_u$ which guaranties that all bits of $S^{(r_u)}$ do not violate the form given in (4). This necessitate also the following representation for the other internal states

$$S_i^{(r_u-t+j)} = h_i^{(j)}(X_I, \mathcal{K}_u) + g_i^{(j)}(K), \quad 1 \leq j < t, \quad 0 \leq i < m. \tag{5}$$

Now, for step $t+1$ the attacker performs a one-round partial decryption of $S^{(r_u-t)}$ under $\mathcal{K}_u$ to get $S^{(r_u-t-1)}$. This is followed by a $(t+1)$-round partial encryption of $S^{(r_u-t-1)}$ under $\mathcal{K}$. Throughout the partial encryption, she should check if any bit of intermediate state violates condition (5) for all $S^{(r_u-t-j)}$, $1 \leq j < t$ and condition (4) for $S^{(r_u)}$. Once such an event occurs, she breaks the partial encryption procedure, detects the internal state bit which should be forced to zero in order to avoid this event, finds the appropriate key bit to be guessed, and updates $\mathcal{K}_u$.

Having updated $\mathcal{K}_u$, the attacker repeats the one-round partial decryption of $S^{(r_u-t)}$ under the updated $\mathcal{K}_u$. Then, she performs the $(t+1)$-round partial encryption using $\mathcal{K}$, regularly checks conditions (4) and (5) and repeat the above scenario until these conditions are completely satisfied. In this way, at the end of step $t+1$, the the attacker has derived a list $\mathcal{K}_u$ of key bits to be guessed as well as the description of $S^{(r_u-t-1)}$ based on X_I and $\mathcal{K}_u$ which guaranties conditions (4) and (5).

This subroutine terminates at $t = r_u$, when all the required outputs $\{\mathcal{K}_u, p(X_I, \mathcal{K}_u), U, Z\}$ are provided. A schematic view can be seen in Fig. 2. It should be noted that, for a given $\{r_u, I\}$, this attack requires N different plaintext descriptions, denoted by $p^{(i)}(X_I, \mathcal{K}_u), i = 1, 2, \ldots N$ all for the same $\mathcal{K}_u, U$ and Z. As we will see in the case studies, given one $p(X_I, \mathcal{K}_u)$, it is an easy task to derive the other ones.

This subroutine solely determines the attack data complexity for a given $\{r_u, I\}$ and a number of N tests. Since each of the N tests requires 2^d plaintexts for $2^{|\mathcal{K}_u|}$ key guesses, the data complexity would be bounded to $D = N \cdot 2^{d+|\mathcal{K}_u|}$. After calling Upper-Extension-Subroutine for a given $\{r_u, I\}$ in Algorithm 1 of Appendix C, the (upper bound of) data complexity is computed and this algorithm will continue if the computed data complexity is less than D_{max}, otherwise it turns into the next $\{r_u, I\}$ pair.

Cube-Tester-Subroutine

Once the attacker defined the status of all bits of $S^{(r_u)}$, i.e. U and Z for a given I, she can find the longest cube tester. To do so, for a given I, Z, U she runs Cube-Tester-Subroutine given in Algorithm 3 of Appendix C. First, the potentially lowest algebraic degree bit in any round is identified as the distinguisher bit T (i.e. the bit whose non-randomness is supposed to be distinguished), based on the cipher internal structure.

This algorithm starts from $r_c = 1$, increases r_c in each step, and checks if a cube tester can be found for T at round $r_u + r_c$, while the input cube, static and dynamic variables are set at $S^{(r_u)}$ according to I, Z and U, respectively. The tool used here is similar to probabilistic linear test in [2] which is modified to tolerate a linear dependency to neutral bits (which is a function of key bits, in turn) as well as a linear dependency to key bits. So, this subroutine returns the maximum r_c, for which the distinguisher bit can be described as

$$\begin{aligned} T(r_c) &= L(K) + \sum_{i \in U} b_i \cdot S_i^{(r_u)} \\ &= L(K) + \sum_{i \in U} b_i \cdot f_i(K) = F(K) \end{aligned} \tag{6}$$

where, L is a linear function. If $F(K)$ is not a constant function, $\mathcal{K}_u$ should be updated as $\mathcal{K}_u = \mathcal{K}_u \cup F(K)$. Therefore, for a given $\{U, Z, I, r_u\}$, Cube-Tester-Subroutine returns the maximum possible r_c, for which the distinguisher bit has a non-random property.

The reader should be noticed that we called the variables whose index are in U neutral variables, since they contain an unknown constant value. Although the distinguisher should not naturally depend on a neutral variable by definition, our algorithm accepts a distinguisher as far as it is dependency on neutral bits is at most a linear dependency.

Lower-Extension-Subroutine

Once the cube tester length is determined, the only thing remaining in the

preprocessing phase is to trace the distinguisher bit at round $r_u + r_c$ towards the ciphertext to find a subset of (equivalent) key bits, $\mathcal{K}_l$, whose guess is sufficient for computation of the distinguisher bit from the ciphertext. This subroutine returns the maximum rounds r_l which can be covered by the lower extension part. The limiting factor here is the total time complexity of the attack which is directly influenced by $d, |\mathcal{K}_u|$ and $|\mathcal{K}_l|$ according to (7). The procedure of this subroutine is detailed in Algorithm 4 of Appendix C.

Having run the three subroutines for all candidate pairs of $\{r_u, I\}$, the pair which covers maximum rounds r, as well as all the relevant parameters $\{r, r_l, \mathcal{K}_u, \mathcal{K}_l, p^{(j)}(X_I, \mathcal{K}_u), 1 \leq j \leq N, Time, Data\}$ is returned as the output of preprocessing phase.

Online Phase This phase of the attack is detailed in Algorithm 5 of Appendix C. By this algorithm, the whole secret key is retrieved as follows. For each guess for $\mathcal{K}_u \cup \mathcal{K}_l$, the attacker computes N batch of plaintexts according to $p^{(j)}(X_I, \mathcal{K}_u), 1 \leq j \leq N$ and queries the encryption oracle to get the corresponding batch of ciphertexts. Then, she checks whether this specific guess for $\mathcal{K}_u \cup \mathcal{K}_l$ passes all the N tests. If so, this guess is regarded as a candidate for these subset of key bits and will be rechecked by another (P, C) pair while the remaining key bits have been guessed. So, the attack complexities are:

$$\begin{aligned} \text{Time Complexity} &= N \cdot 2^{d+|\mathcal{K}_u|} + \frac{r_l}{r_u + r_c + r_l} N \cdot 2^{d+|\mathcal{K}_u \cup \mathcal{K}_l|} + 2^{|\mathcal{K}|-N} \\ \text{Data Complexity} &= N \cdot 2^{d+|\mathcal{K}_u|} \\ \text{Memory Complexity} &= N \cdot 2^{d} \end{aligned} \tag{7}$$

According to Algorithm 5, this memory is required for storing all N batch of plaintext/ciphertext pairs.

4 Cryptanalysis of KATAN-32

4.1 Specification of KATAN-32

KATAN-32, the smallest member of KATAN family of block ciphers with a 32-bit and 80-bit block and key sizes. The plaintext is loaded into two registers $L1$, and $L2$ with 13 and 19 bits length, respectively. The least significant bit of each register, nunbered by 0, is the rightmost one, and the LSB of plaintext is loaded into the LSB of $L2$ while its MSB is loaded into the MSB of $L1$. Each round, $L1$ and $L2$ are shifted to the left for one bit, where the newly computed bits generated according to the following equations, are loaded into the LSB of $L2$ and $L1$, respectively.

$$\begin{aligned} f_a(L1) &= L1[12] + L1[7] + (L1[8] \cdot L1[5]) + (L1[3] \cdot IR) + a \\ f_b(L2) &= L2[18] + L2[7] + (L2[12] \cdot L2[10]) + (L2[8] \cdot L2[3]) + b \end{aligned} \tag{8}$$

where IR is a round-dependent constant, and a, b are two subkey bits generated by the key schedule of the cipher. After 254 rounds, the contents of the registers are then exported as the ciphertext.

Key schedule. KATAN-32 has a linear key schedule based on a LFSR structure which generates $2 \times 254 = 508$ subkey bits according to the following rule:

$$rk_i = \begin{cases} k_i & 0 \leq i < 80 \\ rk_{i-80} + rk_{i-61} + rk_{i-50} + rk_{i-13} & 80 \leq i < 508 \end{cases} \tag{9}$$

For round $i = 1, \ldots, 254$, a_i is defined to be rk_{2i-2}, whereas b_i is rk_{2i-1}.

4.2 Cube Attack on KATAN32

Selecting appropriate values for input parameters $r_{u,min}, r_{u,max}$, and d plays an important role in the efficiency of the resulted attack. Both the upper and lower extension parts, make a few key bits active in the first rounds of the extensions while it approach to 2 key bits per round in the higher rounds which we call the *saturated* domain for the upper/lower extensions. So, the best strategy is to make the best use of the both unsaturated domains. The optimum r_u locates the cube distinguisher in such a way that both of the unsaturated parts are included in the rounds covered by the attack. A too long upper or lower extension part makes the opposite side too short which does not allow an efficient use of the unsaturated part.

The other important choice is the cube dimension. Although the larger d makes r_c potentially larger, it also makes the second term of the attack time complexity (7) larger. In addition, A larger d increases the chance of guessing key bits in the upper extension part, i.e. increasing $|\mathcal{K}_u|$, which in turn increases the time complexity. Furthermore, be noticed that data complexity is exponentially dependent to d and $|\mathcal{K}_u|$, but independent of $|\mathcal{K}_l|$. The number of tests N, on the one hand reduces the last term of time complexity, but on the other hand it increases the second term. For $Time_{max} = 2^{79}$, we set $N = 2$. For KATAN-32, we select the MSB of register $L2$ as the distinguisher bit which had the lowest degree compared to the other state bits at the same round. We perform our attack in two scenarios: non-full-codebook and full-codebook attacks. In non-full-codebook attack we set $Data_{max} = 2^{20}$ and 2^{27} and searched through the small dimension cubes. The best results belongs to cubes of dimensions 5 and 6 which could break up to 116 and 117 rounds of KATAN-32, respectively. In full-codebook scenario, we set $Data_{max} = 2^{32}$ and searched through cubes of $d = 31$ one of which could attack 152 rounds of KATAN-32. The details of the best found cubes are reported in Table 3. As an example, all the parameters of attack Case I are reported in details in Appendix A.

Table 3. Details of the best dynamic cube attacks on KATAN-32

Case	d	N	I	r_u	r_c	r_l	$\lvert\mathcal{K}_u\rvert$	$\lvert\mathcal{K}_l\rvert$	r	Time	Data
I	5	2	$7, 13, 14, 17, 18$	22	50	45	14	59	116	$2^{78.83}$	2^{19}
II	6	2	$9, 13, 14, 15, 17, 19$	23	53	41	21	51	117	$2^{78.77}$	2^{27}
III	31	2	$1, 2, \ldots, 31$	17	101	34	11	37	152	$2^{78.92}$	2^{32}

5 Cryptanalysis of SIMON32/64

5.1 Specification of SIMON32/64

SIMON32/64 is a 32-round Feistel block cipher whose round function composed of AND, XOR and rotations. The internal state at the input and output of round i are denoted by $S^{(i-1)} = (L^{(i-1)}, R^{(i-1)})$ and $S^{(i)} = (L^{(i)}, R^{(i)})$ respectively, where

$$\begin{aligned} L^{(i)} &= R^{(i-1)} + F(L^{(i-1)}) + k^{(i)} \\ R^{(i)} &= L^{(i-1)} \end{aligned}$$

$$F(L^{(i-1)}) = (L^{(i-1)} \lll 1) \cdot (L^{(i-1)} \lll 8) + (L^{(i-1)} \lll 2)$$

where $\lll$ denotes the left rotation operation and $k^{(i)}$ is the subkey of round i, $1 \leq\ i \leq 32$.

Key schedule. SIMON has a linear key schedule as follows. Assume that $K = K_3, K_2, K_1, K_0$ is the representation of secret key in four 16-bit words K_i. The subkey of SIMON32/64 at round i is generated as follows.

$$k^{(i)} = \begin{cases} K_{i+1} & 0 \leq i \leq 3 \\ k^{(i-4)} + Y + (Y \ggg 1) + c + (z_0)_i & 4 \leq i \leq 32 \end{cases}$$

where $Y = k^{(i-3)} + (k^{(i-1)} \ggg 3)$, c is a constatnt value and z_0 is a binary stream whose precise value is given in [15].

5.2 Cube Attack on SIMON32/64

Similar to KATAN, we consider the non-full-codebook and full-codebook scenarios, where in the former the cubes with small dimensions are searched and for the latter, cubes with dimension 31 are examined. For the non-full-codebook scenario, we set $D_{max} = 2^{20}$ and $N = 16$. Amongst all the the small dimension cubes which were examined, the best result belongs to a cube of dimension 5 which could break 16 round of SIMON32/64. In the full-codebook scenario, we could attack 18-round SIMON32/64 by a number of $N = 17$ tests. In this special case, we found many cubes of size 31 all for the same distinguisher bit, each of which can be regarded as tester for the same subset of $\mathcal{K}_l$. So, in this case instead of generating N different plaintext descriptions we make use of N different cube testers. The parameters related two our attacks are reported in Table 4. For Case II, only one of out of the 17 cubes is written. Furthermore, as an example all the parameters of attack Case I are given in details in Appendix B.

Table 4. Details of the best dynamic cube attacks on SIMON32/64

Case	d	N	I	r_u	r_c	r_l	$\lvert\mathcal{K}_u\rvert$	$\lvert\mathcal{K}_l\rvert$	r	Time	Data
I	5	16	$8, 14, 16, 22, 24$	3	8	5	10	34	16	$2^{51.5}$	2^{19}
II	31	17	$0, \ldots, 30$	1	13	4	0	18	18	2^{51}	2^{32}

6 Conclusions and Discussions

In this paper we proposed a new framework for dynamic cube attacks on block ciphers. The algorithm proposed for this attack can be made fully automated, in spite of its counterpart in stream ciphers. We analysed the efficiency of attack on two block ciphers KATAN and SIMON and the results shows that this newcomer method for cryptanalysis of block cipher can competes well with the mature and well-organized cryptanalysis methods in this domain.

Although our attack is similar to the dynamic cube attack on stream ciphers [4] in some properties, it is completely different in the tools, which are used. In more details, our attack shares the following properties with the dynamic cube attack on stream ciphers.

- Assigning the public variables some functions of cube variables and guessed keys (dynamic variables).
- Using a cube tester to check the validity of guessed key bits.

But, the main differentiations of our work with the dynamic cube attack on stream ciphers are as follows.

- The cipher is explicitly divided into three *distinct* parts: upper extension part, cube part, and lower extension part. Each part are treated in sequence and semi-independently. In dynamic cube attack on stream ciphers, the lower part is not present and the first two parts can not be separated.
- This partition enables the attacker to *fully automate* the attack and get rid of the manual complex process of dynamic cube attack for deriving the description of dynamic variables.
- The cube variables are defined among the intermediate state bits rather than the public variables.
- The keys guessed in the attack are not only those involved in the dynamic variables. They are also involved in the lower rounds of the cipher.

In fact, the main feature of block ciphers from which most of these differences arise, especially the possibility to automate the attack, is availability of an only-key-dependent function in any round of the cipher, i.e. the key schedule. Key schedule enables the attacker to perform partial encryption/decryption anywhere among the cipher rounds conditioned that she has guessed the required subkeys. Such a facility is not provided in stream ciphers, where the secret key along with the IV are loaded into the stream cipher state in the first clock and after that there would be no pure access to the secret key, either in the initialization or key stream generation phases.

References

1. M. Vielhaber, "Breaking ONE.FIVIUM by AIDA, an Algebraic IV Differential Attack", Cryptology ePrint Archive, Report 2007/413, 2007.
2. I. Dinur, and A. Shamir, "Cube Attacks on Tweakable Black Box Polynomials", EUROCRYPT'09, LNCS, vol. 5479, pp. 278-299, Springer, 2009.
3. J. Aumasson, I. Dinur, W. Meier, and A. Shamir, "Cube Testers and Key Recovery Attacks on Reduced-Round MD6 and Trivium", FSE'09. LNCS, vol. 5665, pp. 1-22, Springer, 2009.
4. I. Dinur, and A. Shamir, "Breaking Grain-128 with Dynamic Cube Attacks", FSE'11, LNCS, vol. 6733, pp. 167187, Springer, 2011.
5. M. Hell, T. Johansson, A. Maximov, and W. Meier, "The Grain Family of Stream Ciphers", New Stream Cipher Designs - The eSTREAM Finalists, LNCS, vol. 4986, pp.179190, Springer, 2008.
6. C. De Canniere, and B. Preneel, "Trivium", New Stream Cipher Designs - The eSTREAM Finalists. LNCS, vol. 4986, pp. 244-266, Springer, 2008.
7. P. -A. Fouque and T. Vannet, "Improving Key Recovery to 784 and 799 rounds of Trivium using Optimized Cube Attacks", FSE'13, LNCS, vol. 8424, pp. 502-517, Springer, 2013.
8. A. Vardasbi, M. Salmasizadeh, and J. Mohajeri, "Superpoly Algebraic Normal Form Monomial Test on Trivium", IET Information Security, vol. 7, no. 3, pp. 230-238, 2013.
9. M. Rahimi, M. Barmshory, M. H. Mansouri, and M. R. Aref, "Dynamic Cube Attack on Grain-v1", IACR Cryptology ePrint Archive, vol. 2013, p.268, 2013.
10. G. V. Bard, N. T. Courtois, J. Nakahara, P. Sepehrdad, and B. Zhang, "Algebraic, Aida/Cube and Side Channel Analysis of Katan Family of Block Ciphers", INDOCRYPT'10, LNCS, vol. 6498, pp. 176196, Springer, 2010.
11. I. Dinur and A. Shamir, "Side Channel Cube Attacks on Block Ciphers", IACR Cryptology ePrint Archive, vol. 2009, Report 2009/127, 2009.
12. Z. Li, B. Zhang, J. Fan, and I. Verbauwhede, "A New Model for Error-Tolerant Side-Channel Cube Attacks", CHES'13, LNCS, vol. 8086, pp. 453-470, 2013.
13. L. Yang, M. Wang, and S. Qiao, "Side Channel Cube Attack on PRESENT", CANS'09, LNCS, vol. 5888, pp. 379-391, 2009.
14. S. F. Abdul-Latip, R. Reyhanitabar, W. Susilo, and J. Seberry, "Fault Analysis of the Katan Family of Block Ciphers", ISPEC'12, LNCS, vol. 7232, pp.319-336, Springer, 2012.
15. R. Beaulieu, D. Shors, J. Smith, S. Treatman-Clark, B. Weeks, and L. Wingers, "The Simon and Speck Families of Lightweight Block Ciphers", IACR Cryptology ePrint Archive, vol. 2013, Report 2013/404, 2013.
16. C. D. Cannire, O. Dunkelman, and M. Knezevic, "Katan and Ktantan - a Family of Small and Efficient Hardware-oriented Block Ciphers", CHES'09, LNCS, vol. 5747, pp. 272-288, Springer, 2009.
17. S. Knellwolf, W. Meier, and M. Naya-Plasencia, "Conditional Differential Cryptanalysis of NLFSR-based Cryptosystems", ASIACRYPT'10, LNCS, vol. 6477, pp. 130-145, Springer, 2010.
18. M. R. Albrecht and G. Leander, "An all-in-one Approach to Differential Cryptanalysis for Small Block Ciphers, SAC'12, LNCS, vol. 7707, pp. 115, Springer, 2013.
19. T. Isobe and K. Shibutani, "Improved All-Subkeys Recovery Attacks on Fox, Katan and Shacal-2 Block Ciphers, in the proceeding of FSE'14, LNCS, Springer, 2014.

20. T. Fuhr and B. Minaud, "Match Box Meet-in-the-Middle Attack Against Katan". (in C. Cid ed.), in the proceeding of FSE'14, LNCS, Springer, 2014.
21. J. Alizadeh, H. A. Alkhzaimi, M. R. Aref, N. Bagheri, P. Gauravaram, A. Kumar, M. M. Lauridsen, and S. K. Sanadhya, "Cryptanalysis of Simon variants with Connections". in the proceeding of RFIDsec14.
22. C. Boura, M. Naya-Plasencia, and V. Suder, "Scrutinizing and Improving Impossible Differential Attacks: Applications to Clefia, Camellia, Lblock and Simon", ASIACRYPT 2014.
23. Farzaneh Abed, E. List, S. Lucks, and J. Wenzel, "Differential Cryptanalysis of Round-Reduced Simon and Speck", in the proceeding of FSE'14, LNCS, Springer, 2014.
24. A. R. Alex Biryukov and V. Velichkov, "Differential Analysis of Block Ciphers Simon and Speck", in the proceeding of FSE'14, LNCS, Springer, 2014.
25. H. AlKhzaimi and M. M. Lauridsen, "Cryptanalysis of the SIMON family of block ciphers", IACR Cryptology ePrint Archive, Report 2013/543, 2013.
26. N. Wang, X. Wang, K. Jia, and J. Zhao, "Improved differential attacks on reduced SIMON versions", IACR Cryptology ePrint Archive, vol.2014, p.448, 2014.

Appendix A. Details of attack case I on KATAN-32

In this appendix the details of attack case I on KATAN-32 are reported. The set of upper extension key is $\mathcal{K}_u = \{u_i | 1 \leq i \leq 14\}$ where

$$
\begin{aligned}
&u_1 = a_3, \quad u_2 = a_4, \quad u_3 = a_{10}, \quad u_4 = a_{13}, \quad u_5 = b_5, \quad u_6 = b_1, \quad u_7 = a_7 + b_3,\\
&u_8 = a_6 + b_2, \quad u_9 = a_{11} + b_7 + b_3, \quad u_{10} = a_{14} + b_{10} + b_6, \quad u_{11} = a_{13} + b_7 \cdot b_4,\\
&u_{12} = a_{11} + a_7 + a_2 + b_7, \quad u_{13} = a_{10} + b_2 + b_4 \cdot (a_{14} + b_{10} + b_6),\\
&u_{14} = a_9 + a_{14} + b_{10} + b_6 + b_3 \cdot (a_{13} + b_7 \cdot b_4)
\end{aligned}
$$

The first descryption for plaintext is $p^{(1)}(X_I, \mathcal{K}_u) = S^{(0)} = (L1^{(0)}, L2^{(0)})$ where

$$
\begin{aligned}
L1^{(0)} = \{ \; & u_9 + u_{13} + u_8 + (\mathbf{x_3} + u_4) \cdot (\mathbf{x_2} + \mathbf{x_5} + u_{10} + u_3) + u_{11} \cdot (\mathbf{x_5} + u_{13})\\
& u_{12} + u_{10} \cdot u_9 + (\mathbf{x_3} + u_{14}) \cdot (\mathbf{x_5} + u_8 + u_9 + u_{11} \cdot (\mathbf{x_5} + u_{13}))\\
& \mathbf{x_3} + u_4 + u_1 + (u_7 + u_9 \cdot u_{10}) \cdot (\mathbf{x_5} + u_8)\\
& \mathbf{x_1} + \mathbf{x_4} + u_2 + u_{11} + u_{14} + (\mathbf{x_3} + u_4) \cdot u_9\\
& \mathbf{x_2} + \mathbf{x_5} + u_3 + u_{10}, \quad \mathbf{x_5} + u_8 + u_9 + u_{11} \cdot (\mathbf{x_5} + u_{13}), \quad u_7 + u_{10} \cdot u_9\\
& \mathbf{x_3} + u_4, \quad \mathbf{x_4} + u_{14}, \quad \mathbf{x_5} + u_{13}, \quad u_9, \quad 0, \quad u_{11}\}\\
L2^{(0)} = \{ \; & \mathbf{x_1} + u_{10} + u_6, \; \mathbf{x_5} + \mathbf{x_2}, \; 0, \; 0, \; \mathbf{x_3} + u_5, \; \mathbf{x_4}, \; 0, \; 0, \; 0, \; 0, \; 0,\\
& \mathbf{x_1}, \; \mathbf{x_2} + \mathbf{x_1} \cdot \mathbf{x_4}, \; 0, \; 0, \; \mathbf{x_3}, \; \mathbf{x_4}, 0, 0\}
\end{aligned}
\tag{10}
$$

which yields the following descrption for $S^{(22)} = (L1^{(22)}, L2^{(22)})$ which is completely consistent with (4).

$$
\begin{aligned}
L1^{(22)} = \{\ & b_{10},\ b_{11},\ b_{12},\ a_5 + b_{13} + b_2 + b_4 \cdot u_{10},\ b_{14},\ b_{15},\ a_8 + b_{16} + b_7 \cdot b_4,\\
& b_{17},\ b_{18},\ b_{19},\ a_1 + b_6 \cdot b_3 + b_4 + b_8 + a_{12} + b_{20},\\
& b_{21} + (b_6 \cdot b_3 + b_4 + b_8 + a_{12}) \cdot (b_4 + b_{11} \cdot b_8 + a_1 + b_9 + a_{17}),\ b_{22}\}\\
L2^{(22)} = \{\ & \mathbf{x_1}, \mathbf{x_2} + a_5 + b_2 + b_4 \cdot (a_{14} + b_{10} + b_6),\ 0,\ 0,\ \mathbf{x_3} + a_8 + b_7 \cdot b_4,\\
& \mathbf{x_4},\ 0,\ 0,\ b_6 \cdot b_3 + b_4 + b_8 + a_{12},\ 0,\ 0,\\
& \mathbf{x_5} + b_2 + b_6 \cdot (a_1 + b_9) + b_7 + a_{15},\ b_3 + b_{10} \cdot b_7 + b_8 + b_{12} + a_{16},\\
& b_4 + b_{11} \cdot b_8 + a_1 + b_9 + a_{17}, b_{12} \cdot (a_1 + b_9) + b_{10} + b_{14} + a_{18},\\
& b_6 + b_{10} \cdot (a_5 + b_{13} + b_2 + b_4 \cdot (a_{14} + b_{10} + b_6)) + b_{11} + a_{19},\\
& b_7 + b_{14} \cdot b_{11} + b_{12} + a_8 + b_16 + b_7 \cdot b_4 + a_{20},\\
& b_8 + b_{15} \cdot b_{12} + a_5 + b_{13} + b_2 + b_4 \cdot (a_{14} + b_{10} + b_6) + b_{17} + a_{21},\\
& b_9 + a_1 + (a_8 + b_{16} + b_7 \cdot b_4) \cdot (a_5 + b_{13} + b_2 + b_4 \cdot (a_{14} + b_{10} + b_6)) +\\
& b_{14} + b_{18} + a_{21}\}
\end{aligned}
$$

The set of neutral and static variables are as follows.

$$
\begin{aligned}
U &= \{0, 1, 2, 3, 4, 5, 6, 10, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31\}\\
Z &= \{8, 9, 11, 12, 15, 16\}]
\end{aligned}
$$

To get the second plaintext description, let $p^{(2)}(X_I, \mathcal{K}_u) = p^{(1)}(X_I, \mathcal{K}_u) + \Delta_2$ where Δ_2 and its corresponding difference at round 22, $\Delta_2^{(22)}$, are as follows.

$$
\begin{aligned}
\Delta_2 &= \{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0\}\\
\Delta_2^{(22)} &= \{0,0,0,0,0,0,0,0,0,0,0,b_6 \cdot b_3 + b_4 + b_8 + a_{12} + b_4 + b_{11} \cdot b_8 + a_1 + b_9 +\\
&\quad a_{17} + 1, 0,0,0,0,0,0,0,0,0,0,1,0,0,0,1,b_{11},0,0,0,1,0\}
\end{aligned}
$$

which shows that all the nonzero differences are constant and confined to neutral bits. Finally, the set of key bits to be guessed in the lower extension part are as follows.

$$
\begin{aligned}
\mathcal{K}_l &= \{a_i | i \in A_l\} \cup \{b_i | i \in B_l\}\\
A_l &= \{77, 80, 81, 82, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101,\\
&\quad 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117\}\\
B_l &= \{98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114,\\
&\quad 115, 116, 117\}
\end{aligned}
$$

Appendix B. Details of attack case I on SIMON32/64

In this appendix the details of attack case I on SIMON32/64 are reported. The set of upper extension key is $\mathcal{K}_u = \{u_i | 1 \leq i \leq 10\}$ where

$$
\begin{aligned}
& u_1 = c_{15},\ \ u_2 = c_1,\ \ u_3 = c_5,\ \ u_4 = b_{15},\ \ u_5 = b_{13},\\
& u_6 = b_9,\ \ u_7 = b_7,\ \ u_8 = b_3,\ \ u_9 = b_1,\ \ u_{10} = b_5 + c_7,
\end{aligned}
$$

The first descryption for plaintext is $p^{(1)}(X_I, \mathcal{K}_u) = S^{(0)} = (L^{(0)}, R^{(0)})$ where

$$
\begin{aligned}
L^{(0)} &= \{u_1,\ \mathbf{x_2},\ 0,\ \mathbf{x_1},\ 0,\ \mathbf{x_5},\ 0,\ \mathbf{x_1}+\mathbf{x_4},\ u_{10},\ 0,\ u_3,\ 0,\ 0,\ \mathbf{x_2}+\mathbf{x_3},\ u_2,\ 0\}\\
R^{(0)} &= \{\mathbf{x_2}\cdot u_{10}+u_4,\ \mathbf{x_1},\ \mathbf{x_1}\cdot u_3+u_6,\ \mathbf{x_5},\ 0,\ \mathbf{x_4}, u_2\cdot(\mathbf{x_1}+\mathbf{x_4})+u_{10}+u_6,\\
&\quad \mathbf{x_5},\ u_3+u_7,\ \mathbf{x_4}+\mathbf{x_2}+u_3,\ 0,\ \mathbf{x_2}+\mathbf{x_3},\ u_2+u_8,\ \mathbf{x_5}\cdot u_2,\ u_1+u_9,\\
&\quad \mathbf{x_2}+u_1\cdot(\mathbf{x_1}+\mathbf{x_4})\}
\end{aligned}
$$

which yields the following descrption for $S^{(3)} = (L^{(3)}, R^{(3)})$ which is completely consistent with (4).

$$
\begin{aligned}
L^{(3)} = \{&k_5^{(1)}\cdot k_{12}^{(1)} + k_{11}^{(1)} + k_{13}^{(2)} + k_{15}^{(3)},\\
&(k_5^{(1)}\cdot k_{14}^{(1)} + k_4^{(1)} + k_6^{(2)})\cdot(k_5^{(1)}\cdot k_{12}^{(1)} + k_{11}^{(1)} + k_{13}^{(2)}) + \ k_{10}^{(1)} + k_4^{(1)}\cdot k_{11}^{(1)} +\\
&k_{12}^{(2)} + k_{14}^{(1)} + k_{14}^{(3)}k_{11}^{(2)} + k_{13}^{(3)},\ k_{11}^{(2)}\cdot(k_2^{(1)} + k_4^{(2)}) + k_8^{(1)} + k_{10}^{(2)} + k_{12}^{(1)} + k_{12}^{(3)},\\
&(k_8^{(1)} + k_{10}^{(2)})\cdot(k_2^{(1)}\cdot k_{11}^{(1)} + k_3^{(2)}) + k_9^{(2)} + k_{11}^{(1)} + k_{11}^{(3)},\\
&k_6^{(1)} + k_{10}^{(1)} + k_8^{(2)} + k_9^{(2)}\cdot(k_0^{(1)} + k_2^{(2)}) + k_{10}^{(3)},\ \ k_9^{(3)},\\
&\mathbf{x_5} + k_5^{(1)}\cdot k_{14}^{(1)} + k_4^{(1)} + k_6^{(2)} + k_8^{(1)} + k_8^{(3)},\ \ k_7^{(3)},\ \ \mathbf{x_4} + k_2^{(1)} + k_4^{(2)} + k_6^{(1)} + k_6^{(3)},\\
&(k_2^{(1)} + k_4^{(2)})\cdot(k_5^{(1)}\cdot k_{12}^{(1)} + k_{11}^{(1)} + k_{13}^{(2)}) + k_2^{(1)}\cdot k_{11}^{(1)} + k_3^{(2)} + k_5^{(1)} + k_5^{(3)},\\
&(k_2^{(1)}\cdot k_{11}^{(1)} + k_3^{(2)})\cdot(k_{10}^{(1)} + k_4^{(1)}\cdot k_{11}^{(1)} + k_{12}^{(2)}) + k_0^{(1)} + k_2^{(2)} + k_4^{(1)} + k_4^{(3)},\\
&k_{11}^{(2)}\cdot(k_0^{(1)} + k_2^{(2)}) + k_3^{(3)},\ \ k_{14}^{(1)} + k_0^{(2)} + k_2^{(1)} + k_2^{(3)},\\
&k_9^{(2)}\cdot(k_{14}^{(1)} + k_0^{(2)}) + k_1^{(3)},\\
&\mathbf{x_3} + k_0^{(1)} + k_{12}^{(1)} + k_{14}^{(2)} + k_0^{(3)}\}\\
R^{(3)} = \{&0,\ \ \mathbf{x_2} + k_{12}^{(1)} + k_{14}^{(2)},\ \ k_5^{(1)}\cdot k_{12}^{(1)} + k_{11}^{(1)} + k_{13}^{(2)},\ \ k_{10}^{(1)} + k_4^{(1)}\cdot k_{11}^{(1)} + k_{12}^{(2)},\\
&k_{11}^{(2)}, k_8^{(1)} + k_{10}^{(2)}, k_9^{(2)},\ \ \mathbf{x_1} + k_6^{(1)} + k_8^{(2)},\ \ 0,\ \ k_5^{(1)}\cdot k_1^{(1)}4 + k_4^{(1)} + k_6^{(2)},\ \ 0,\\
&k_2^{(1)} + k^{(2)},\ \ , k_2^{(1)}\cdot k_{11}^{(1)} + k_3^{(2)},\ \ k_0^{(1)} + k_2^{(2)},\ \ 0,\ \ k_{14}^{(1)} + k_0^{(2)}\}
\end{aligned}
$$

The set of neutral and static variables are identified as follows.

$$
\begin{aligned}
Z &= \{1, 5, 7, 15\}\\
U &= \{0, 2, 3, 4, 6, 9, 10, 11, 12, 13, 17, 18, 19, 20, 21, 22, 24, 25, 26, 27, 28, 29, 30, 31\}
\end{aligned}
$$

To get the remaining 15 plaintext descriptions, let $p^{(j)}(X_I, \mathcal{K}_u) = p^{(1)}(X_I, \mathcal{K}_u) + \Delta_j$, $j = 2, \ldots, 16$. Δ_j could be any difference in the right half of $p^{(1)}(X_I, \mathcal{K}_u)$ since it would result in a difference at round 3, $\Delta_j^{(3)}$, with nonzero constant values confined to U and I indexes. For example see the following pair of differences:

$$
\begin{aligned}
\Delta_j &= \{0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1\}\\
\Delta_j^{(3)} &= \{0,0,0,0,0,0,0,0,0,0,0,1,k_{11}^{(2)},0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0\}
\end{aligned}
$$

Finally, the set of key bits to be guessed in the lower extension part are as follows.

$$\mathcal{K}_l = \{k_i^{(14)} | i \in I_{14}\} \cup \{k_i^{(15)} | i \in I_{15}\} \cup \{k_i^{(16)} | i \in I_{16}\} \cup \{k_i^{(17)} | i \in I_{17}\}$$
$$I_{14} = \{7, 14\}, \quad I_{15} = \{5, 6, 12, 13, 15\}$$
$$I_{16} = \{3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15\}, \quad I_{17} = \{0, 1, \ldots, 15\}$$

Appendix C. Attack algorithms

Algorithm 1 Preprocessing Phase

Input: $\{r_{u,min}, r_{u,max}, d, N, Time_{max}, Data_{max}\}$.
Output: $\{r, r_l, \mathcal{K}_u, \mathcal{K}_l, p^{(j)}(X_I, \mathcal{K}_u), 1 \leq j \leq N, Time, Data\}$.
$r = 0$; $Out' = \{\}$;
for $r_u = r_{u,min}$ **to** $r_{u,max}$ **do**
  for randomly chosen I of size d **do**
    $(\mathcal{K}_u, p(X_I, \mathcal{K}_u), U, Z) =$ Upper-Extension-Subroutine(r_u, I)
    $Data = N \cdot 2^{d+|\mathcal{K}_u|}$;
    if $Data \leq Data_{max}$ **then**
      $r_c =$ Cube-Tester-Subroutine(U, Z, r_u, d);
      $(r_l, \mathcal{K}_l, Time)$=Lower-Extension-Subroutine$(r_c + r_u, d, \mathcal{K}_u, Time_{max}, N)$;
      if $r_u + r_c + r_l > r$ **then**
        $r = r_u + r_c + r_l$;
        $Out' \leftarrow \{r, r_l, \mathcal{K}_u, \mathcal{K}_l, p(X_I, \mathcal{K}_u), Time, Data\}$;
      end if
    end if
  end for
end for
$p^{(1)}(X_I, \mathcal{K}_u) = p(X_I, \mathcal{K}_u)$;
for $j = 2$ **to** N **do**
  $p^{(j)}(X_I, \mathcal{K}_u) = p^{(1)}(X_I, \mathcal{K}_u) + \Delta_j$; ▷ for some appropriate Δ_j chosen by the attacker.
end for
$Out \leftarrow \{r, r_l, \mathcal{K}_u, \mathcal{K}_l, p^{(j)}(X_I, \mathcal{K}_u), 1 \leq j \leq N, Time, Data\}$; **return** Out;

Algorithm 2 Upper-Extension-Subroutine

Input: $\{r_u, I\}$.
Output: $\{\mathcal{K}_u, p(X_I, \mathcal{K}_u), U, Z\}$.
$X_I = \{x_i | i \in I\}$; ▷ x_i is symbolic.
$\mathcal{K} = \{k_{n-1}, \ldots, k_0\}$; ▷ k_i is symbolic.
$S^{(r_u)} = 0$, $S_I^{(r_u)} = X_I$, $\mathcal{K}_u = \{\}$;
for $t = 0$ **to** $r_u - 1$ **do**
    $\tilde{S}^{(r_u - t - 1)} = dec(S^{(r_u - t)}, \mathcal{K}_u)$;
    for $j = 0$ **to** t **do**
        $\tilde{S}^{(r_u - t + j)} = enc(\tilde{S}^{(r_u - t + j - 1)}, \mathcal{K})$;
        if conditions (4) and (5) are not satisfied due to the key bit k_f **then**
            update $\mathcal{K}_u$ by k_f;
            go to 7;
        end if
    end for
    $S^{(r_u - t - 1)} = \tilde{S}^{(r_u - t - 1)}$;
end for
$p(X_I, \mathcal{K}_u) = S^{(0)}$;
$S^{(r_u)} = \tilde{S}^{(r_u)}$;
Identify sets U and Z from $S^{(r_u)}$;
return $\{\mathcal{K}_u, p(X_I, \mathcal{K}_u), U, Z\}$;

Algorithm 3 Cube-Tester-Subroutine

Input: $\{U, Z, I, r_u, d\}$.
Output: $\{r_c\}$.
for $r_c = 1$ **to** $r_{c,max}$ **do** ▷ $r_{c,max}$ is set by the attacker.
    $T_0 = 0$;
    for $i = 0$ **to** $2^d - 1$ **do**
        $S = 0, S_I = i$; ▷ writes the binary representation of i in bits I of S.
        $T_0 = T_0 + Enc(S, 0, r_u, r_c)$;
    end for
    for $j = 1$ **to** M **do** ▷ M is the number of linearity tests, set by the attacker
        Randomly choose $x, y \in \{0,1\}^{|U|}$ and $k_1, k_2 \in \{0,1\}^{|\mathcal{K}|}$;
        $S_{1_Z} = 0$, $S_{1_U} = x$;
        $S_{2_Z} = 0$, $S_{2_U} = y$;
        $S_{3_Z} = 0$, $S_{3_U} = x + y$;
        $T_1 = 0, T_2 = 0, T_3 = 0$;
        for $i = 0$ **to** $2^d - 1$ **do**
            $S_{1_I} = i$, $S_{2_I} = i$, $S_{3_I} = i$;
            $T_1 = T_1 + Enc(S_1, k_1, r_u, rc)$;
            $T_2 = T_2 + Enc(S_2, k_2, r_u, rc)$;
            $T_3 = T_3 + Enc(S_3, k_1 + k_2, r_u, rc)$;
        end for
        if $T_1 + T_2 + T_3 + T_0 \neq 0$ **then return** $r_c - 1$; halt.
        end if
    end for
end for

Algorithm 4 Lower-Extension-Subroutine

Input: $\{r', \mathcal{K}_u, d, Time_{max}\}$.
Output: $\{r_l, \mathcal{K}_l, Time\}$
$\mathcal{K}_l = \{\}, r_l = 0, Time = N \times 2^{d+|\mathcal{K}_u|} + 2^{|\mathcal{K}|-N}$;
while $Time < Time_{max}$ **do**
  $\tilde{\mathcal{K}_l} \leftarrow \mathcal{K}_l, \tilde{Time} \leftarrow Time$;
  Trace T in forward direction for one round;
  Update $\mathcal{K}_l$;
  $Time = N \cdot 2^{d+|\mathcal{K}_u|} + \frac{r_l}{r'+r_l} N \cdot 2^{d+|\mathcal{K}_u \cup \mathcal{K}_l|} + 2^{|\mathcal{K}|-N}$;
  $r_l = r_l + 1$;
end while
return $\{r_l - 1, \tilde{\mathcal{K}_l}, \tilde{Time}\}$;

Algorithm 5 Online Phase

Input: $d, N, r_l, \mathcal{K}_u, \mathcal{K}_l, p^{(j)}(X_I, \mathcal{K}_u), 1 \le j \le N$
Output: secret key K.
for $k_1 = 0$ **to** $2^{|\mathcal{K}_u|} - 1$ **do**
  for $i = 0$ **to** $2^d - 1$ **do**
    for $j = 0$ **to** N **do**
      Compute plaintext $P(j, i, k_1) = p^{(j)}(X_I, \mathcal{K}_u)|_{X_I = i, \mathcal{K}_u = k_1}$;
      Query the r-round encryption oracle and save the corresponding ciphertext $C(j, i, k_1)$;
    end for
  end for
  for $k_2 = 0$ **to** $2^{|\mathcal{K}_l - \mathcal{K}_l \cap \mathcal{K}_u|} - 1$ **do**
    $fail = 0$; $j = 0$;
    while $j < N$ **and** $fail = 0$ **do**
      $T = 0$;
      for $i = 0$ **to** $2^d - 1$ **do**
        $T = T + Dec(C(j, i, k_1), (k_1, k_2), r, r_l)$;
      end for
      if $T \neq 0$ **then**
        $fail = 1$;
      end if
      $j = j + 1$;
    end while
    if $fail = 0$ **then**
      Randomly choose P and query its ciphertext C;
      for $k_3 = 0$ **to** $2^{|\mathcal{K} - \mathcal{K}_l \cup \mathcal{K}_u|} - 1$ **do**
        if $C = Enc(P, (k_1, k_2, k_3), 0, r)$ **then return** (k_1, k_2, k_3);
        end if
      end for
    end if
  end for
end for