Bit-Based Division Property and Application to Simon Family

Yosuke Todo[1,2(✉)] and Masakatu Morii[2]

[1] NTT Secure Platform Laboratories, Tokyo, Japan
todo.yosuke@lab.ntt.co.jp
[2] Kobe University, Kobe, Japan

Abstract. Ciphers that do not use S-boxes have been discussed for the demand on lightweight cryptosystems, and their round functions consist of `and`, `rotation`, and `xor`. Especially, the Simon family is one of the most famous ciphers, and there are many cryptanalyses again the Simon family. However, it is very difficult to guarantee the security because we cannot use useful techniques for S-box-based ciphers. Very recently, the division property, which is a new technique to find integral characteristics, was shown in Eurocrypt 2015. The technique is powerful for S-box-based ciphers, and it was used to break, for the first time, the full MISTY1 in CRYPTO 2015. However, it has not been applied to non-S-box-based ciphers like the Simon family effectively, and only the existence of the 10-round integral characteristic on Simon32 was proven. On the other hand, the experimental characteristic, which possibly does not work for all keys, covers 15 rounds, and there is a 5-round gap. To fill the gap, we introduce a bit-based division property, and we apply it to show that the experimental 15-round integral characteristic always works for all keys. Though the bit-based division property finds more accurate integral characteristics, it requires much time and memory complexity. As a result, we cannot apply it to symmetric-key ciphers whose block length is over 32. Therefore, we alternatively propose a method for designers. The method works for ciphers with large block length, and it shows "provable security" against integral cryptanalyses using the division property. We apply this technique to the Simon family and show that Simon48, 64, 96, and 128 probably do not have 17-, 20-, 25-, and 29-round integral characteristics, respectively.

Keywords: Integral cryptanalysis · Division property · Provable security · Simon family

1 Introduction

Non-S-box-based ciphers have been proposed for the demand on lightweight cryptosystems [2,3]. Such ciphers are superior in lightweight environments because they are implemented by logical operations and do not have a lookup table like S-boxes. In 2013, the NSA proposed a lightweight block cipher family,

T. Peyrin (Ed.): FSE 2016, LNCS 9783, pp. 357–377, 2016.
DOI: 10.1007/978-3-662-52993-5_18

Table 1. Integral characteristics on SIMON32

Methods	#Rounds	Balanced bit (right half)	Reference
Experiment (no proof)	15	(?b??,????,b???,???b)	[18]
Division	10	(bbbb,bbbb,bbbb,bbbb)	[17]
Conventional bit-based division	14	(bbbb,bbbb,bbbb,bbbb)	Sect. 3
Bit-based division using 3 subsets	15	(?b??,????,b???,???b)	Sect. 4

called the SIMON family, that follows this design principle [3]. However, it is too difficult to guarantee the security against several cryptanalyses because we cannot use many useful techniques for S-box-based ciphers. Therefore, many cryptanalyses have been proposed against the SIMON family, e.g., [1,5,6,10,15,18], and the designers recently summarized cryptanalyses in [4]. In this paper, we investigate the security of non-S-box-based ciphers against integral cryptanalyses and illustrate our methods on the SIMON family.

Division Property. Very recently, the division property, which is a new technique to find integral characteristics [9], was proposed in Eurocrypt 2015 [17]. The new technique permitted us to find a 6-round integral characteristic on MISTY1 in CRYPTO 2015, leading to the first complete theoretical cryptanalysis of the full MISTY1 [16]. Moreover, this technique was applied to generalized Feistel structures in [20], leading to improved integral cryptanalyses against LBlock and TWINE. The division property also proves integral characteristics on the SIMON family in [17], and SIMON32, 48, 64, 96, and 128 have 9-, 11-, 11-, 13-, 13-round integral characteristics, respectively[1]. However, the round function is regarded as any function of degree 2. Therefore, we can expect that integral characteristics can be extended to more rounds if one is able to exploit the concrete structure of the round function. In fact, the experimental integral characteristic, which possibly does not work for all keys, covers 15 rounds [18], and there is a large gap between the proved characteristic and experimental one.

Our Contribution. The round function of the SIMON family is regarded as any function of degree 2 in [17] because we cannot decompose the round function into several sub blocks like S-boxes. However, we can decompose the round function into every bit, and we call the division property that focuses on every bit a *bit-based division property*.

First, we apply the conventional bit-based division property to SIMON32, which is not against the definition of the division property. Therefore, we can directly use the propagation rules of the division property. As a result, the conventional bit-based division property proves that SIMON32 has a 14-round integral characteristic. However, there is still a gap of one round between the

[1] Since the round key is XORed after the round function in SIMON, we can trivially get one-round extended integral characteristics.

Table 2. Provable secure number of rounds for the SIMON family

Ciphers	SIMON48	SIMON64	SIMON96	SIMON128	reference
Vulnerable number	14 rounds	17 rounds	21 rounds	25 rounds	[21]
Provable security	17 rounds	20 rounds	25 rounds	29 rounds	this paper

proof and experiment. Namely, this means that either the experimental 15-round characteristic does not work for all keys or the conventional bit-based division property cannot find the accurate characteristic. As a result, we conclude that the conventional bit-based division property is insufficient to find the accurate characteristic. The conventional division property divides the set of $\boldsymbol{u}$ according to whether the parity becomes 0 or unknown [17]. However, we should divide the set of $\boldsymbol{u}$ according to whether the parity becomes 0, 1, or unknown because we can also exploit the fact that the parity is not only 0 but also 1. To exploit this fact, we newly introduce a variant of the bit-based division property, which divides the set of $\boldsymbol{u}$ into three subsets. Since the variant is completely different from the definition of the conventional division property, we show the propagation characteristic also. Finally, we apply the variant to SIMON32 and show that the experimental 15-round characteristic always works for all keys. The proved characteristic is the completely same as the experimental one including the position of balanced bits. Table 1 shows the comparison of integral characteristics, where balanced and unknown bits are labeled as `b` and `?`, respectively.

Although the bit-based division property can find more accurate integral characteristics, their propagations require much time and memory complexity. When we evaluate the propagation for n-bit block ciphers, it roughly requires 2^n complexity because the bit-based division property has to manage the set of n-dimensional vectors whose elements take values in $\mathbb{F}_2$. This is feasible for SIMON32 because the block length is 32 bits, but it is infeasible for other SIMON family members. Therefore, we introduce a new technique, which is useful for designers but is not useful for attackers. We call this technique a *lazy propagation*, where we evaluate only a part of all propagations. The lazy propagation cannot find the integral characteristic, but it can evaluate the number of rounds that the bit-based division property cannot find integral characteristics even if we can evaluate the accurate propagation. Namely, the technique shows "provable security" for the integral cryptanalysis using the division property, and we expect that it becomes a useful technique for designers. Our provable security guarantees the security against only the integral cryptanalysis using the division property, and it does not always guarantee the security against all integral-like cryptanalyses. However, for SIMON32, the bit-based division property can find the accurate integral characteristic. Therefore, we expect that it also finds the best integral characteristic for the other SIMON family if it is feasible. Table 2 shows the number of rounds of SIMON48, 64, 96, and 128, where

the division property never finds integral characteristics. As a result, we expect that SIMON48, 64, 96, and 128 do not have 17-, 20-, 25-, and 29-round integral characteristics, respectively[2]. Moreover, as the comparison, Table 2 also shows the number of rounds that SIMON48, 64, 96, and 128 have integral characteristics [21].

2 Preliminaries

2.1 Notations

We make the distinction between the addition of $\mathbb{F}_2^n$ and addition of $\mathbb{Z}$, and we use $\oplus$ and $+$ as the addition of $\mathbb{F}_2^n$ and addition of $\mathbb{Z}$, respectively. For any $a \in \mathbb{F}_2^n$, the ith element is expressed in $a[i]$, and the Hamming weight $w(a)$ is calculated as $w(a) = \sum_{i=1}^{n} a[i]$. For any $\boldsymbol{a} \in (\mathbb{F}_2^{n_1} \times \mathbb{F}_2^{n_2} \times \cdots \times \mathbb{F}_2^{n_m})$, the vectorial Hamming weight of $\boldsymbol{a}$ is defined as $W(\boldsymbol{a}) = (w(a_1), w(a_2), \ldots, w(a_m)) \in \mathbb{Z}^m$. Moreover, for any $\boldsymbol{k} \in \mathbb{Z}^m$ and $\boldsymbol{k'} \in \mathbb{Z}^m$, we define $\boldsymbol{k} \succeq \boldsymbol{k'}$ if $k_i \geq k'_i$ for all i. Otherwise, $\boldsymbol{k} \not\succeq \boldsymbol{k'}$. In this paper, we often treat the set of $\boldsymbol{k}$, and $\mathbb{K}$ denotes this set. Then, let $|\mathbb{K}|$ be the number of vectors. We simply write $\mathbb{K} \leftarrow \boldsymbol{k}$ when $\mathbb{K} := \mathbb{K} \cup \{\boldsymbol{k}\}$. Moreover, we simply write $\mathbb{K} \xleftarrow{\text{x}} \boldsymbol{k}$, where the new $\mathbb{K}$ computed as

$$\mathbb{K} := \begin{cases} \mathbb{K} \cup \{\boldsymbol{k}\} & \text{if the original } \mathbb{K} \text{ does not include } \boldsymbol{k}, \\ \mathbb{K} \setminus \{\boldsymbol{k}\} & \text{if the original } \mathbb{K} \text{ includes } \boldsymbol{k}. \end{cases}$$

2.2 Integral Attack

The integral attack was first introduced by Daemen et al. to evaluate the security of SQUARE [7], and then it was formalized by Knudsen and Wagner [9]. Attackers first prepare N chosen plaintexts and encrypt them R rounds. If the XOR of all encrypted texts becomes 0, we say that the cipher has an R-round integral characteristic with N chosen plaintexts. Finally, we analyze the entire cipher by using the integral characteristic. Therefore, it is very important to find integral characteristic. There are two main approaches to find integral characteristics. The first one is the propagation of the integral property [9] and the second one is based on the degree estimation [8,11].

2.3 Division Property

The division property, which was proposed in [17], is a new method to find integral characteristics. This section briefly shows the definition and propagation rules. Please refer to [17] in detail.

[2] If we truly guarantee the security against integral attack, we have to consider the key recovery part.

Bit Product Function. The division property of a multiset is evaluated by using the bit product function defined as follows. Let $\pi_u : \mathbb{F}_2^n \to \mathbb{F}_2$ be a bit product function for any $u \in \mathbb{F}_2^n$. Let $x \in \mathbb{F}_2^n$ be the input, and $\pi_u(x)$ is the AND of $x[i]$ satisfying $u[i] = 1$, i.e., it is defined as

$$\pi_u(x) := \prod_{i=1}^{n} x[i]^{u[i]}.$$

Notice that $x[i]^1 = x[i]$ and $x[i]^0 = 1$. Let $\pi_{\boldsymbol{u}} : (\mathbb{F}_2^{n_1} \times \mathbb{F}_2^{n_2} \times \cdots \times \mathbb{F}_2^{n_m}) \to \mathbb{F}_2$ be a bit product function for any $\boldsymbol{u} \in (\mathbb{F}_2^{n_1} \times \mathbb{F}_2^{n_2} \times \cdots \times \mathbb{F}_2^{n_m})$. Let $\boldsymbol{x} \in (\mathbb{F}_2^{n_1} \times \mathbb{F}_2^{n_2} \times \cdots \times \mathbb{F}_2^{n_m})$ be the input, and $\pi_{\boldsymbol{u}}(\boldsymbol{x})$ is defined as

$$\pi_{\boldsymbol{u}}(\boldsymbol{x}) := \prod_{i=1}^{m} \pi_{u_i}(x_i).$$

The bit product function also appears in the Algebraic Normal Form (ANF) of a Boolean function. The ANF of a Boolean function f is represented as

$$f(x) = \bigoplus_{u \in \mathbb{F}_2^n} a_u^f \left(\prod_{i=1}^{n} x[i]^{u[i]} \right) = \bigoplus_{u \in \mathbb{F}_2^n} a_u^f \pi_u(x),$$

where $a_u^f \in \mathbb{F}_2$ is a constant value depending on f and u.

Definition of Division Property

Definition 1 (Division Property [17]). *Let $\mathbb{X}$ be a multiset whose elements take a value of $(\mathbb{F}_2^{n_1} \times \mathbb{F}_2^{n_2} \times \cdots \times \mathbb{F}_2^{n_m})$. When the multiset $\mathbb{X}$ has the division property $\mathcal{D}_{\mathbb{K}}^{n_1,n_2,\ldots,n_m}$, where $\mathbb{K}$ denotes a set of m-dimensional vectors whose ith element takes a value between 0 and n_i, it fulfils the following conditions:*

$$\bigoplus_{\boldsymbol{x} \in \mathbb{X}} \pi_{\boldsymbol{u}}(\boldsymbol{x}) = \begin{cases} \textit{unknown} & \textit{if there are } \boldsymbol{k} \in \mathbb{K} \textit{ s.t. } W(\boldsymbol{u}) \succeq \boldsymbol{k}, \\ 0 & \textit{otherwise.} \end{cases}$$

See [17] to better understand the concept in detail, and [14] and [16] help readers understand the division property. In this paper, the division property for $(\mathbb{F}_2^n)^m$ is referred to as $\mathcal{D}_{\mathbb{K}}^{n^m}$ for the simplicity[3]. If there are $\boldsymbol{k} \in \mathbb{K}$ and $\boldsymbol{k}' \in \mathbb{K}$ satisfying $\boldsymbol{k} \succeq \boldsymbol{k}'$ in the division property $\mathcal{D}_{\mathbb{K}}^{n_1,n_2,\ldots,n_m}$, $\boldsymbol{k}$ can be removed from $\mathbb{K}$ because the vector $\boldsymbol{k}$ is redundant.

Propagation Rules of Division Property. Some propagation rules for the division property are proven in [17], and the rules are summarized in [16] as follows.

[3] In [17], the division property was referred to as $\mathcal{D}_{\mathbb{K}}^{n,m}$.

Rule 1 (Substitution). Let F be a function that consists of m S-boxes, where the bit length and the algebraic degree of the ith S-box is n_i bits and d_i, respectively. The input and output take a value of $(\mathbb{F}_2^{n_1} \times \mathbb{F}_2^{n_2} \times \cdots \times \mathbb{F}_2^{n_m})$, and $\mathbb{X}$ and $\mathbb{Y}$ denote the input multiset and output multiset, respectively. Assuming that the multiset $\mathbb{X}$ has the division property $\mathcal{D}_{\mathbb{K}}^{n_1,n_2,\ldots,n_m}$, the division property of the multiset $\mathbb{Y}$ is $\mathcal{D}_{\mathbb{K}'}^{n_1,n_2,\ldots,n_m}$ as

$$\mathbb{K}' \leftarrow \left(\left\lceil \frac{k_1}{d_1} \right\rceil, \left\lceil \frac{k_2}{d_2} \right\rceil, \ldots, \left\lceil \frac{k_m}{d_m} \right\rceil \right), \quad \forall \boldsymbol{k} \in \mathbb{K}.$$

Here, when the ith S-box is bijective and $k_i = n_i$, the ith element of the propagated property becomes n_i not $\lceil n_i/d_i \rceil$.

Rule 2 (Copy). Let F be a copy function, where the input x takes a value of $\mathbb{F}_2^n$ and the output is calculated as $(y_1, y_2) = (x, x)$. Let $\mathbb{X}$ and $\mathbb{Y}$ be the input multiset and output multiset, respectively. Assuming that the multiset $\mathbb{X}$ has the division property $\mathcal{D}_k^n$, the division property of the multiset $\mathbb{Y}$ is $\mathcal{D}_{\mathbb{K}'}^{n,n}$ as

$$\mathbb{K}' \leftarrow (k - i, i), \quad \text{for } 0 \le i \le k.$$

Rule 3 (Compression by XOR). Let F be a function compressed by an XOR, where the input (x_1, x_2) takes a value of $(\mathbb{F}_2^n \times \mathbb{F}_2^n)$ and the output is calculated as $y = x_1 \oplus x_2$. Let $\mathbb{X}$ and $\mathbb{Y}$ be the input multiset and output multiset, respectively. Assuming that the multiset $\mathbb{X}$ has the division property $\mathcal{D}_{\mathbb{K}}^{n,n}$, the division property of the multiset $\mathbb{Y}$ is $\mathcal{D}_{k'}^n$ as

$$k' = \min_{(k_1,k_2)\in\mathbb{K}} \{k_1 + k_2\}.$$

Here, if the minimum value of k' is larger than n, the propagation characteristic of the division property is aborted. Namely, a value of $\oplus_{y\in\mathbb{Y}} \pi_v(y)$ is 0 for all $v \in \mathbb{F}_2^n$.

Rule 4 (Split). Let F be a split function, where the input x takes a value of $\mathbb{F}_2^n$ and the output is calculated as $x = y_1 \| y_2$, where (y_1, y_2) takes a value of $(\mathbb{F}_2^{n_1} \times \mathbb{F}_2^{n-n_1})$. Let $\mathbb{X}$ and $\mathbb{Y}$ be the input multiset and output multiset, respectively. Assuming that the multiset $\mathbb{X}$ has the division property $\mathcal{D}_k^n$, the division property of the multiset $\mathbb{Y}$ is $\mathcal{D}_{\mathbb{K}'}^{n_1,n-n_1}$ as

$$\mathbb{K}' \leftarrow (k - i, i), \quad \text{for } 0 \le i \le k.$$

Here, $(k - i)$ is less than or equal to n_1, and i is less than or equal to $n - n_1$.

Rule 5 (Concatenation). Let F be a concatenation function, where the input (x_1, x_2) takes a value of $(\mathbb{F}_2^{n_1} \times \mathbb{F}_2^{n_2})$ and the output is calculated as $y = x_1 \| x_2$. Let $\mathbb{X}$ and $\mathbb{Y}$ be the input multiset and output multiset, respectively. Assuming that the multiset $\mathbb{X}$ has the division property $\mathcal{D}_{\mathbb{K}}^{n_1,n_2}$, the division property of the multiset $\mathbb{Y}$ is $\mathcal{D}_{k'}^{n_1+n_2}$ as

$$k' = \min_{(k_1,k_2)\in\mathbb{K}} \{k_1 + k_2\}.$$

2.4 Simon Family

The Simon family is a lightweight block cipher family [3] based on the Feistel construction. Let Simon2n be the Simon block ciphers with $2n$-bit block length, where n is chosen from 16, 24, 32, 48, and 64. Moreover, Simon2n with mn-bit secret key is referred to as Simon2n/mn. Since we only care about integral characteristics on the Simon family, this paper only uses Simon2n.

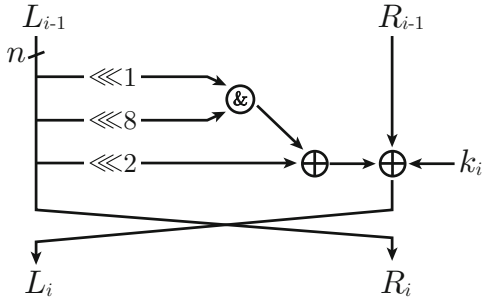

Fig. 1. Round function of Simon2n

The output of the ith round function is denoted by (L_i, R_i) and is calculated as

$$(L_i, R_i) = (L_{i-1}^{\lll 1} \wedge L_{i-1}^{\lll 8}) \oplus L_{i-1}^{\lll 2} \oplus R_{i-1} \oplus k_i, L_{i-1}),$$

where $L^{\lll j}$ denotes the j-bit left rotation of L, and k_i denotes the ith round key. Moreover, (L_0, R_0) denotes a plaintext. The round function consists of `and`, `rotation`, and `xor`, and Fig. 1 shows the round function. For more details, please refer to [3].

2.5 Known Integral Characteristic on Simon Family

It is difficult to find effective integral characteristics on ciphers which consist of `and`, `rotation`, and `xor`. In [18], authors experimentally showed that Simon32 has the 15-round integral characteristic with 2^{31} chosen plaintexts. Since their characteristic is confirmed under 2^{13} secret keys, they expected that the success probability of this characteristic is at least $1 - 2^{-13}$. Therefore, this approach does not guarantee that the characteristic works for all secret keys. Moreover, it is practically infeasible to find integral characteristics of other Simon family members because the block length is too large for proceeding to an experimental evaluation.

Integral characteristics proved under all secret keys are shown in [17], but in this approach the round function of Simon2n is seen as any n-bit function of degree 2. Therefore, the detailed structure of the round function is not exploited.

As a result, it shows that SIMON32, 48, 64, 96, and 128 has 9-, 11-, 11-, 13-, and 13-round integral characteristic, respectively. Since the round key is XORed after the round function, we can trivially get one-round extended integral characteristics using the same technique in [18]. Therefore, 10-, 12-, 12-, 14-, and 14-round integral characteristics are proved in SIMON32, 48, 64, 96, and 128, respectively. Thus, there is a 5-round gap between the proved characteristic and experimental one.

3 Conventional Bit-Based Division Property

This paper introduces a *bit-based division property*. When n-bit block ciphers are analyzed, the conventional division property uses $\mathcal{D}_{\mathbb{K}}^{\ell_1,\ell_2,\ldots,\ell_m}$, where ℓ_i and m are chosen by attackers in the range of $n = \sum_{i=1}^{m} \ell_i$. This section considers the conventional bit-based division property, i.e., $\mathcal{D}_{\mathbb{K}}^{1^n}$. Since it is not against the definition of the conventional division property, we can directly use the five propagation rules shown in Sect. 2.3.

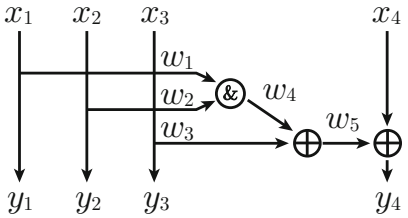

Fig. 2. Core operation of the SIMON family.

3.1 Comparison Between Conventional Bit-Based Division Property and Solving Algebraic Equations

Before the introduction of the conventional bit-based division property, we roughly show the relation between the bit-based division property and the resolution of algebraic equations by brute force. When entire ciphers are represented by algebraic equations, such equations involve both the plaintext and secret key. Therefore, if we solve such equations for an n-bit block cipher with a k-bit secret key, this roughly requires 2^{k+n} complexity. On the other hand, XORing with a constant value does not change the conventional bit-based division property because such XORing is a linear function [16]. Therefore, the propagation of the conventional bit-based division property does not involve the secret key. It may miss some useful cryptographic properties, but it dramatically reduces the complexity.

3.2 Propagation for Core Operation of Simon

As an example, we analyze SIMON2n by using the conventional bit-based division property. We focus on only one bit of the right half in SIMON2n. The core operation of the round function is represented by Fig. 2. Since the input and output bit length is 4 bits, we use the division property $\mathcal{D}_{\mathbb{K}}^{1^4}$.

We consider the propagation characteristic. For instance, let assume that the input multiset has $\mathcal{D}_{[k_1,k_2,k_3,1]}^{1^4}$, where k_i denotes any value, i.e., 0 or 1. Then, if the multiset of $(y_1, y_2, y_3, w_5, x_4)$ has $\mathcal{D}_{[*,*,*,1,1]}^{1^5}$, where $*$ is propagated values, the propagation always abort in the XOR, $x_4 \oplus w_5$. Consequently, the bit-based division property of (y_1, y_2, y_3, y_4) is the same as that of (x_1, x_2, x_3, x_4). On the other hand, assuming that the input multiset has $\mathcal{D}_{[k_1,k_2,k_3,0]}^{1^4}$, the output property is different from the input one.

Let $\mathcal{D}_{\mathbb{K}}^{1^4}$ and $\mathcal{D}_{\mathbb{K}'}^{1^4}$ be the division property of the input and output, respectively. When we get $\mathbb{K}'$ from $\mathbb{K}$, we first independently calculate vectors belonging to $\mathbb{K}'$ by evaluating the propagation from every vector in $\mathbb{K}$. Then, $\mathbb{K}'$ is represented as the union of all calculated vectors. Finally, if there are $\boldsymbol{k} \in \mathbb{K}'$ and $\boldsymbol{k}' \in \mathbb{K}'$ such that $\boldsymbol{k} \succeq \boldsymbol{k}'$, $\boldsymbol{k}$ is removed from $\mathbb{K}'$ because the vector is redundant.

Table 3 summarizes the propagation characteristics from $\mathcal{D}_{\boldsymbol{k}}^{1^4}$ to $\mathcal{D}_{\mathbb{K}}^{1^4}$. The round function of SIMON2n repeats the core operation for all n-bit values in the right half. Therefore, we use $\mathcal{D}_{\mathbb{K}}^{1^{2n}}$. In every core operation, we only focus on four bits and evaluate the propagation independent of other $(2n-4)$ bits.

Table 3. Propagation of the conventional bit-based division property for the core operation in the SIMON family

Input $\mathcal{D}_{\boldsymbol{k}}^{1^4}$	Output $\mathcal{D}_{\mathbb{K}}^{1^4}$
$\boldsymbol{k} = [0,0,0,0]$	$\mathbb{K} = \{[0,0,0,0]\}$
$\boldsymbol{k} = [1,0,0,0]$	$\mathbb{K} = \{[1,0,0,0],[0,0,0,1]\}$
$\boldsymbol{k} = [0,1,0,0]$	$\mathbb{K} = \{[0,1,0,0],[0,0,0,1]\}$
$\boldsymbol{k} = [1,1,0,0]$	$\mathbb{K} = \{[1,1,0,0],[0,0,0,1]\}$
$\boldsymbol{k} = [0,0,1,0]$	$\mathbb{K} = \{[0,0,1,0],[0,0,0,1]\}$
$\boldsymbol{k} = [1,0,1,0]$	$\mathbb{K} = \{[1,0,1,0],[0,0,1,1],[1,0,0,1]\}$
$\boldsymbol{k} = [0,1,1,0]$	$\mathbb{K} = \{[0,1,1,0],[0,0,1,1],[0,1,0,1]\}$
$\boldsymbol{k} = [1,1,1,0]$	$\mathbb{K} = \{[1,1,1,0],[0,0,1,1],[1,1,0,1]\}$
$\boldsymbol{k} = [k_1,k_2,k_3,1]$	$\mathbb{K} = \{[k_1,k_2,k_3,1]\}$

Table 4. Size of $\mathbb{K}$ in $\mathcal{D}_{\mathbb{K}}^{1^{32}}$ for the integral characteristic on SIMON32

Round	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$\lvert\mathbb{K}\rvert$	1	1	3	11	65	774	18165	587692	5191387	1595164	95768	5894	682	136	32

3.3 Application to Simon32

We evaluate the propagation characteristic of the conventional bit-based division property on SIMON32. We prepare chosen plaintexts such that the first bit is constant and the others are active. Then, the set of chosen plaintexts has the division property $\mathcal{D}_{\mathbb{K}}^{1^{32}}$, where $\mathbb{K} = \{[0, 1, 1, \ldots, 1]\}$. Table 4 shows $|\mathbb{K}|$, which is the number of vectors, in every round, where we perfectly remove redundant vectors from $\mathbb{K}$. The output of the 14th round function has the division property $\mathcal{D}_{\mathbb{K}}^{1^{32}}$, where $\mathbb{K}$ has 32 distinct vectors whose Hamming weight is one. Therefore, the conventional bit-based division property cannot show whether or not the output of the 14th round function is balanced. On the other hand, the output of the 13th round function has the division property $\mathcal{D}_{\mathbb{K}}^{1^{32}}$, where $\mathbb{K}$ is represented as 16 vectors, whose Hamming weight of the left half is 1 and that of the right half is 0, and 120 $(= \binom{16}{2})$ vectors, whose Hamming weight of the left half is 0 and that of the right half is 2. This division property means that the output of the 13th round function takes the following integral property

$$(\texttt{????,????,????,????,}\ \texttt{bbbb,bbbb,bbbb,bbbb}),$$

where balanced and unknown bits are labeled as `b` and `?`, respectively. In the SIMON family, since round keys are XORed with the right half only after the round function is applied to the left half, we can easily get a 14-round integral characteristic from the 13-round one. The same technique is used in [18]. Therefore, we conclude that 14-round SIMON32 has the integral characteristic with 2^{31} chosen plaintexts.

4 Bit-Based Division Property Using Three Subsets

4.1 Motivation

The conventional bit-based division property proved the existence of the 14-round integral characteristic of SIMON32. However, the experimental characteristic covers 15 rounds [18], and there is still a one-round gap between the experiment and proof. In [18], the authors experimentally confirm the characteristic by randomly choosing 2^{13} secret keys. Therefore, they concluded that the success probability of the characteristic is at least $1 - 2^{-13}$. Thus, we consider that this gap derives from either the experimental result does not work for all keys or the conventional bit-based division property cannot find the accurate characteristic.

We first show that the conventional bit-based division property is insufficient to find integral characteristics on SIMON32, and we then introduce a new variant of the bit-based division property. The conventional bit-based division property focuses on that the parity $\bigoplus_{\boldsymbol{x} \in \mathbb{X}} \pi_{\boldsymbol{u}}(\boldsymbol{x})$ is 0 or unknown. On the other hand, the new variant focuses on that the parity $\bigoplus_{\boldsymbol{x} \in \mathbb{X}} \pi_{\boldsymbol{u}}(\boldsymbol{x})$ is 0, 1, or unknown. Therefore we call the new variant *the bit-based division property using three subsets*. The new variant can find more accurate integral characteristics and prove that the experimental characteristic shown in [18] works for all keys.

4.2 Characteristic that Conventional Bit-Based Division Property Cannot Find

The conventional division property divides the set of $\boldsymbol{u}$ according to whether the parity becomes 0 or unknown [17]. However, it sometimes overlooks useful characteristics. We show it by using a simple example.

We again evaluate the propagation of the conventional bit-based division property for the circuit in Fig. 2, and $F : \mathbb{F}_2^4 \rightarrow \mathbb{F}_2^4$ denotes the circuit. Moreover, let $\mathbb{X}$ and $\mathbb{Y}$ be the input and output multiset, respectively. Assuming that $\mathbb{X}$ has $\mathcal{D}^{1^4}_{\{[1,1,0,0],[0,0,1,0]\}}$, $\bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{[1,1,0,0]}(\boldsymbol{x})$ and $\bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{[0,0,1,0]}(\boldsymbol{x})$ are unknown. Then, the output multiset $\mathbb{Y}$ has $\mathcal{D}^{1^4}_{\{[1,1,0,0],[0,0,1,0],[0,0,0,1]\}}$ from Table 3.

Let us assume that both $\bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{[1,1,0,0]}(\boldsymbol{x})$ and $\bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{[0,0,1,0]}(\boldsymbol{x})$ are 1. Even if we know the parity is always one, the division property of $\mathbb{X}$ is $\mathcal{D}^{1^4}_{\{[1,1,0,0],[0,0,1,0]\}}$. However, we can get the following equation.

$$\begin{aligned}
\bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{[0,0,0,1]}(F(\boldsymbol{x})) &= \bigoplus_{\boldsymbol{x}\in\mathbb{X}} (x_1x_2 \oplus x_3 \oplus x_4) \\
&= \bigoplus_{\boldsymbol{x}\in\mathbb{X}} (x_1x_2) \bigoplus_{\boldsymbol{x}\in\mathbb{X}} (x_3) \bigoplus_{\boldsymbol{x}\in\mathbb{X}} (x_4) \\
&= \bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{[1,1,0,0]}(\boldsymbol{x}) \bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{[0,0,1,0]}(\boldsymbol{x}) \bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{[0,0,0,1]}(\boldsymbol{x}) \\
&= 1 \oplus 1 \oplus 0 = 0.
\end{aligned}$$

Therefore, $\oplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{[0,0,0,1]}(F(\boldsymbol{x}))$ is always 0 not unknown, and the division property of $\mathbb{Y}$ becomes $\mathcal{D}^{1^4}_{\{[1,1,0,0],[0,0,1,0],[0,1,0,1],[1,0,0,1]\}}$ not $\mathcal{D}^{1^4}_{\{[1,1,0,0],[0,0,1,0],[0,0,0,1]\}}$.

Since the conventional division property focuses on the case the parity becomes 0, it cannot find characteristics that appear by cancelling like the above example. Therefore, we newly introduce a variant of the bit-based division property to exploit this fact. The variant divides the set of $\boldsymbol{u}$ into three subsets, i.e., 0, 1, and unknown.

4.3 Definition of Bit-Based Division Property Using Three Subsets

The conventional division property uses the set $\mathbb{K}$ to represent the subset of $\boldsymbol{u}$ such that $\bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{\boldsymbol{u}}(\boldsymbol{x})$ is unknown. The bit-based division property using three subsets needs to represent not only the subset of $\boldsymbol{u}$ such that $\bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{\boldsymbol{u}}(\boldsymbol{x})$ is unknown but also the subset of $\boldsymbol{u}$ such that $\bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{\boldsymbol{u}}(\boldsymbol{x})$ is one. Therefore, we use the set $\mathbb{K}$ to represent the subset of $\boldsymbol{u}$ such that $\bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{\boldsymbol{u}}(\boldsymbol{x})$ is unknown, and we also use the set $\mathbb{L}$ to represent the subset of $\boldsymbol{u}$ such that $\bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{\boldsymbol{u}}(\boldsymbol{x})$ is one.

Definition 2 (Bit-based Division Property Using Three Subsets). *Let $\mathbb{X}$ be a multiset whose elements take a value of $(\mathbb{F}_2)^m$, and $\boldsymbol{k}$ is an m-dimensional vector whose ith element takes 0 or 1. When the multiset $\mathbb{X}$ has the bit-based division property using three subsets $\mathcal{D}^{1^m}_{\mathbb{K},\mathbb{L}}$, it fulfils the following conditions:*

$$\bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{\boldsymbol{u}}(\boldsymbol{x}) = \begin{cases} \text{unknown} & \textit{if there are } \boldsymbol{k}\in\mathbb{K} \textit{ s.t. } W(\boldsymbol{u}) \succeq \boldsymbol{k}, \\ 1 & \textit{else if there is } \boldsymbol{\ell}\in\mathbb{L} \textit{ s.t. } W(\boldsymbol{u}) = \boldsymbol{\ell}, \\ 0 & \textit{otherwise.} \end{cases}$$

If there are $\boldsymbol{k}\in\mathbb{K}$ and $\boldsymbol{k}'\in\mathbb{K}$ satisfying $\boldsymbol{k}\succeq\boldsymbol{k}'$, $\boldsymbol{k}$ can be removed from $\mathbb{K}$ because the vector $\boldsymbol{k}$ is redundant. Moreover, when there is $\boldsymbol{k}\in\mathbb{K}$ satisfying $W(\boldsymbol{u})\succeq\boldsymbol{k}$, $\bigoplus_{\boldsymbol{x}\in\mathbb{X}}\pi_{\boldsymbol{u}}(\boldsymbol{x})$ is unknown even if there is $\boldsymbol{\ell}\in\mathbb{L}$ satisfying $W(\boldsymbol{u})=\boldsymbol{\ell}$. Therefore, if there are $\boldsymbol{\ell}\in\mathbb{L}$ and $\boldsymbol{k}\in\mathbb{K}$ satisfying $\boldsymbol{\ell}\succeq\boldsymbol{k}$, the vector $\boldsymbol{\ell}$ is redundant. Notice that redundant vectors in $\mathbb{K}$ and $\mathbb{L}$ do not affect whether $\bigoplus_{\boldsymbol{x}\in\mathbb{X}}\pi_{\boldsymbol{u}}(\boldsymbol{x})$ becomes 0, 1, or unknown for any $\boldsymbol{u}$.

Example 1. Let $\mathbb{X}$ be a multiset whose elements take a value of $(\mathbb{F}_2)^4$. Assume the multiset $\mathbb{X}$ has the bit-based division property $\mathcal{D}^{1^4}_{\mathbb{K},\mathbb{L}}$, where $\mathbb{K} = \{[0,0,0,1],[0,1,1,0]\}$ and $\mathbb{L} = \{[1,0,0,0],[1,0,1,0],[0,0,1,0],[0,0,1,1]\}$. Then, every parity satisfies the following, where the value of $\boldsymbol{u}$ is represented as hexadecimal notation of $(u_1\|u_2\|u_3\|u_4)$.

$\boldsymbol{u}$	0×0	0×1	0×2	0×3	0×4	0×5	0×6	0×7	0×8	0×9	0xA	0xB	0xC	0xD	0xE	0xF
Parity	0	?	1	?	0	?	?	?	1	?	1	?	0	?	?	?

Notice that the parity of $\pi_{[0,0,1,1]}(\boldsymbol{x})$ over all $\boldsymbol{x}\in\mathbb{X}$ is unknown because there is $[0,0,0,1]\in\mathbb{K}$ and $W([0,0,1,1])\succeq W([0,0,0,1])$. Thus, $[0,0,1,1]\in\mathbb{L}$ is redundant.

4.4 Propagation Rules

We show propagation rules for the bit-based division property using three subsets. There rules are very similar to those of the conventional division property. Here, we show three rules, "Copy," "Compression by AND," and "Compression by XOR," because any Boolean function can be evaluated by using these three rules. We omit the proof of three propagation rules in this paper because of the page limit, and please see the full version of this paper.

Rule 1 (Copy). Let F be a copy function, where the input $(x_1,x_2,\ldots,x_m)$ takes values of $(\mathbb{F}_2)^m$, and the output is calculated as $(x_1,x_1,x_2,x_3,\ldots,x_m)$. Let $\mathbb{X}$ and $\mathbb{Y}$ be the input multiset and output multiset, respectively. Assuming that $\mathbb{X}$ has $\mathcal{D}^{1^m}_{\mathbb{K},\mathbb{L}}$, $\mathbb{Y}$ has $\mathcal{D}^{1^{m+1}}_{\mathbb{K}',\mathbb{L}'}$, where $\mathbb{K}'$ and $\mathbb{L}'$ are computed as

$$\mathbb{K}' \leftarrow \begin{cases} (0,0,k_2,\ldots,k_m), & \text{if } k_1=0 \\ (1,0,k_2,\ldots,k_m),(0,1,k_2,\ldots,k_m), & \text{if } k_1=1 \end{cases},$$

$$\mathbb{L}' \leftarrow \begin{cases} (0,0,\ell_2,\ldots,\ell_m), & \text{if } \ell_1=0 \\ (1,0,\ell_2,\ldots,\ell_m),(0,1,\ell_2,\ldots,\ell_m),(1,1,\ell_2,\ldots,\ell_m) & \text{if } \ell_1=1 \end{cases}.$$

from all $\boldsymbol{k}\in\mathbb{K}$ and all $\boldsymbol{\ell}\in\mathbb{L}$, respectively.

Rule 2 (Compression by AND). Let F be a function compressed by an AND, where the input $(x_1, x_2, \ldots, x_m)$ takes values of $(\mathbb{F}_2)^m$, and the output is calculated as $(x_1 \wedge x_2, x_3, \ldots, x_m)$. Let $\mathbb{X}$ and $\mathbb{Y}$ be the input multiset and output multiset, respectively. Assuming that $\mathbb{X}$ has $\mathcal{D}^{1^m}_{\mathbb{K},\mathbb{L}}$, $\mathbb{Y}$ has $\mathcal{D}^{1^{m-1}}_{\mathbb{K}',\mathbb{L}'}$, where $\mathbb{K}'$ is computed from all $\boldsymbol{k} \in \mathbb{K}$ as

$$\mathbb{K}' \leftarrow \left(\left\lceil \frac{k_1 + k_2}{2} \right\rceil, k_3, k_4, \ldots, k_m \right).$$

Moreover, $\mathbb{L}'$ is computed from all $\boldsymbol{\ell} \in \mathbb{L}$ s.t. $(\ell_1, \ell_2) = (0,0)$ or $(1,1)$ as

$$\mathbb{L}' \leftarrow \left(\left\lceil \frac{\ell_1 + \ell_2}{2} \right\rceil, \ell_3, \ell_4, \ldots, \ell_m \right).$$

Rule 3 (Compression by XOR). Let F be a function compressed by an XOR, where the input $(x_1, x_2, \ldots, x_m)$ takes values of $(\mathbb{F}_2)^m$, and the output is calculated as $(x_1 \oplus x_2, x_3, \ldots, x_m)$. Let $\mathbb{X}$ and $\mathbb{Y}$ be the input multiset and output multiset, respectively. Assuming that $\mathbb{X}$ has $\mathcal{D}^{1^m}_{\mathbb{K},\mathbb{L}}$, $\mathbb{Y}$ has $\mathcal{D}^{1^{m-1}}_{\mathbb{K}',\mathbb{L}'}$, where $\mathbb{K}'$ is computed from all $\boldsymbol{k} \in \mathbb{K}$ s.t. $(k_1, k_2) = (0,0)$, $(1,0)$, or $(0,1)$ as

$$\mathbb{K}' \leftarrow (k_1 + k_2, k_3, k_4, \ldots, k_m).$$

Moreover, $\mathbb{L}'$ is computed from all $\boldsymbol{\ell} \in \mathbb{L}$ s.t. $(\ell_1, \ell_2) = (0,0)$, $(1,0)$, or $(0,1)$ as

$$\mathbb{L}' \xleftarrow{\text{x}} (\ell_1 + \ell_2, \ell_3, \ell_4, \ldots, \ell_m).$$

4.5 Dependencies Between $\mathbb{K}$ and $\mathbb{L}$

Propagation for Public Function. In the propagation rules shown in Sect. 4.4, $\mathbb{K}'$ and $\mathbb{L}'$ are computed from $\mathbb{K}$ and $\mathbb{L}$, respectively. Therefore, we can evaluate the propagation from $\mathbb{K}$ and that from $\mathbb{L}$ independently. However, independent propagations generate many redundant vectors in $\mathbb{K}'$ and $\mathbb{L}'$. Note that redundant vectors in $\mathbb{K}'$ and $\mathbb{L}'$ do not affect whether the parity becomes 0, 1, or unknown for any $\boldsymbol{u}$. Therefore, when we consider the propagation for public functions, we do not need to care about the dependencies between $\mathbb{K}$ and $\mathbb{L}$. On the other hand, if there are many redundant vectors, the propagation requires much time complexity. Therefore, we should remove redundant vectors if possible because of the reason of only complexity.

XORing with Secret Round Key. For the public function, the propagation from $\mathbb{K}$ and that from $\mathbb{L}$ are independently evaluated. However, if the secret round key is XORed, every vector in $\mathbb{L}$ affects $\mathbb{K}$.

Let $\mathbb{X}$ and $\mathbb{Y}$ be the input and output multiset, respectively. Then, $\boldsymbol{y} \in \mathbb{Y}$ is computed as $\boldsymbol{y} = \boldsymbol{x} \oplus \boldsymbol{rk}$ for $\boldsymbol{x} \in \mathbb{X}$, where $\boldsymbol{rk}$ is the secret round key. Moreover, let $\mathcal{D}^{1^m}_{\mathbb{K},\mathbb{L}}$ and $\mathcal{D}^{1^m}_{\mathbb{K}',\mathbb{L}'}$ be the bit-based division property using three subsets on $\mathbb{X}$

and $\mathbb{Y}$, respectively. We want to get $\mathbb{K}'$ and $\mathbb{L}'$ from $\mathbb{K}$ and $\mathbb{L}$. We cannot know the secret round key. Therefore, the parity $\bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{\boldsymbol{v}}(\boldsymbol{x} \oplus \boldsymbol{rk})$ satisfying $\boldsymbol{v} \succ \boldsymbol{\ell}$ becomes unknown because the parity depends on the secret round key.

In many ciphers, round keys are XORed with a part of entire bits. Assuming a round key is XORed with the ith bit, $\mathbb{K}'$ is computed as

$$\mathbb{K}' \leftarrow (\ell_1, \ell_2, \ldots, \ell_i \vee 1, \ldots, \ell_m)$$

for all $\boldsymbol{\ell} \in \mathbb{L}$ satisfying $\ell_i = 0$.

4.6 Propagation for Core Operation of Simon

We search for integral characteristics on Simon32 by the bit-based division property using three subsets. Similar to the conventional bit-based division property, we focus on only one bit of the right half and consider the core operation of the Simon family (see Fig. 2).

The core operation is a public function and it does not involve any secret information. Therefore, we can evaluate the propagation from $\mathbb{K}$ and that from $\mathbb{L}$ independently. Table 5 summarizes the propagation characteristics from $\mathcal{D}^{1^4}_{\mathbb{K},\{\boldsymbol{\ell}\}}$ to $\mathcal{D}^{1^4}_{\mathbb{K}',\mathbb{L}'}$, where the propagation from $\mathbb{K}$ to $\mathbb{K}'$ is the same as that in Table 3. Next, the propagation on the round function can be evaluated by repeating for all bits of the right half. Finally, when round keys are XORed with the right half, new vectors are generated from $\mathbb{L}$, and the new vectors are inserted into $\mathbb{K}$.

Table 5. Propagation of the bit-based division property using three subsets for the core operation in the Simon family

Input $\mathcal{D}^{1^4}_{\mathbb{K},\{\boldsymbol{\ell}\}}$	Output $\mathcal{D}^{1^4}_{\mathbb{K}',\mathbb{L}'}$
$\boldsymbol{\ell} = [0,0,0,0]$	$\mathbb{L}' = \{[0,0,0,0]\}$
$\boldsymbol{\ell} = [1,0,0,0]$	$\mathbb{L}' = \{[1,0,0,0]\}$
$\boldsymbol{\ell} = [0,1,0,0]$	$\mathbb{L}' = \{[0,1,0,0]\}$
$\boldsymbol{\ell} = [1,1,0,0]$	$\mathbb{L}' = \{[1,1,0,0],[0,0,0,1],[1,0,0,1],[0,1,0,1],[1,1,0,1]\}$
$\boldsymbol{\ell} = [0,0,1,0]$	$\mathbb{L}' = \{[0,0,1,0],[0,0,0,1],[0,0,1,1]\}$
$\boldsymbol{\ell} = [1,0,1,0]$	$\mathbb{L}' = \{[1,0,1,0],[1,0,0,1],[1,0,1,1]\}$
$\boldsymbol{\ell} = [0,1,1,0]$	$\mathbb{L}' = \{[0,1,1,0],[0,1,0,1],[0,1,1,1]\}$
$\boldsymbol{\ell} = [1,1,1,0]$	$\mathbb{L}' = \{[1,1,1,0],[0,0,1,1],[1,0,1,1],[0,1,1,1],[1,1,0,1]\}$
$\boldsymbol{\ell} = [\ell_1,\ell_2,\ell_3,1]$	$\mathbb{L}' = \{[\ell_1,\ell_2,\ell_3,1]\}$

Table 6. Sizes of $\mathbb{K}$ and $\mathbb{L}$ in $\mathcal{D}^{1^{32}}_{\mathbb{K},\mathbb{L}}$ for the integral characteristic on Simon32

Round	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\lvert\mathbb{L}\rvert$	1	1	5	19	138	2236	89878	4485379	47149981	2453101	20360	168	8	0	0	0
$\lvert\mathbb{K}\rvert$	1	1	1	6	43	722	23321	996837	9849735	2524718	130724	7483	852	181	32	32

4.7 Application to Simon32

We evaluate the propagation characteristic of the bit-based division property using three subsets on SIMON32. We prepare chosen plaintexts such that the first bit is constant and the others are active, and the set of chosen plaintexts has $\mathcal{D}^{1^{32}}_{\{[1,1,1,\ldots,1]\},\{[0,1,1,\ldots,1]\}}$.

Table 6 shows $|\mathbb{K}|$ and $|\mathbb{L}|$ in every round, where we perfectly remove redundant vectors from $\mathbb{K}$ and $\mathbb{L}$. As a result, the output of the 14th round function has $\mathcal{D}^{1^{32}}_{\mathbb{K},\phi}$, where vector in $\mathbb{K}$ are represented by hexadecimal notation as

```
(0001 0000)(0002 0000)(0004 0000)(0008 0000)(0010 0000)(0020 0000)(0040 0000)(0080 0000)
(0100 0000)(0200 0000)(0400 0000)(0800 0000)(1000 0000)(2000 0000)(4000 0000)(8000 0000)
(0000 0002)(0000 0004)(0000 0008)(0000 0010)(0000 0020)(0000 0040)(0000 0081)(0000 0100)
(0000 0200)(0000 0400)(0000 0800)(0000 1000)(0000 2000)(0000 4001)(0000 4080)(0000 8000),
```

and ϕ denotes the empty set. This division property means that the output of the 14th round function takes the following integral property

$$(\texttt{????,????,????,????,\ ?b??,????,b???,???b}),$$

where balanced and unknown bits are labeled as `b` and `?`, respectively. In the SIMON family, we can easily get a 15-round integral characteristic from the 14-round one, and this proved integral characteristic is completely the same as the experimental one. Therefore, we conclude that the experimental characteristic is not probabilistic characteristic, and it works for all keys.

4.8 Application to Simeck32

Simeck was recently proposed in [19], and its round function is very similar to that of SIMON. Let (L_i, R_i) be the output of the ith round function, and it is calculated as

$$(L_i, R_i) = (L_{i-1} \wedge L_{i-1}^{\lll 5}) \oplus L_{i-1}^{\lll 1} \oplus R_{i-1} \oplus k_i, L_{i-1}).$$

The rotation number is changed from $(1, 8, 2)$ to (0,5,1). Similar to SIMON, Simeck has different parameters according to the block length. Let Simeck$2n$ be the Simeck block ciphers with $2n$-bit block length, where n is chosen from 16, 24, and 32.

We also evaluated the propagation of the bit-based division property using three subsets against Simeck32. As a result, the output of the 14th round function has $\mathcal{D}^{1^{32}}_{\mathbb{K},\phi}$, where vectors in $\mathbb{K}$ are represented by hexadecimal notation as

```
(0001 0000)(0002 0000)(0004 0000)(0008 0000)(0010 0000)(0020 0000)(0040 0000)(0080 0000)
(0100 0000)(0200 0000)(0400 0000)(0800 0000)(1000 0000)(2000 0000)(4000 0000)(8000 0000)
(0000 0002)(0000 0004)(0000 0008)(0000 0011)(0000 0021)(0000 0030)(0000 0040)(0000 0080)
(0000 0100)(0000 0201)(0000 0210)(0000 0220)(0000 0401)(0000 0410)(0000 0420)(0000 0600)
(0000 0800)(0000 1000)(0000 2000)(0000 4001)(0000 4010)(0000 4020)(0000 4200)(0000 4400)
(0000 8001)(0000 8010)(0000 8020)(0000 8200)(0000 8400)(0000 C000).
```

This division property means that the output of the 14th round function takes the following integral property

$$(\texttt{????,????,????,????,\ bb??,?bb?,??bb,???b}).$$

Since round keys are XORed after the round function in Simeck, we can trivially get the 15-round integral characteristic. Here, 2^{31} plaintexts are chosen as $(L_0, F(L_0) \oplus R_0)$, where the first bit of R_0 is constant and the others are active.

5 Provable Security Against Integral Cryptanalysis

We introduced the bit-based division property using three subsets in Sect. 4, and we proved that this method can find more accurate integral characteristics than those found by the conventional division property. In particular, we showed that the new method can discover the tight characteristic on SIMON32. However, a problem is left about the feasibility, i.e., the propagation of the division property requires much time and memory complexity. For instance, if we want to evaluate the propagation of the division property $\mathcal{D}_{\mathbb{K}}^{n^m}$, the time and memory complexity is upper-bounded by $(n+1)^m$. Therefore, if the upper bound is too large, e.g., $(n+1)^m \gg 2^{32}$, it is difficult to evaluate the propagation [4]. In the bit-based division property, the time and memory complexity is upper-bounded by 2^n, where n denotes the block length. Moreover, the bit-based division property using three subsets requires more complexity than that using two subsets. Therefore, we cannot apply the bit-based division property to the SIMON family except for SIMON32.

5.1 Provable Security for Designers

We cannot apply the bit-based division property to the SIMON family except for SIMON32, but we can show the "provable security" alternatively. When we design new symmetric-key primitives, we have to guarantee the security against several cryptanalyses. Provable security has been discussed in detail for differential and linear cryptanalyses [12,13], but such tools do not exist for integral cryptanalysis.

Let $\mathcal{D}_{\mathbb{K}_i,\mathbb{L}_i}^{1^m}$ denotes the division property of the output set of the ith round function. We want to find R-round integral characteristics. Then, for any $\boldsymbol{u}$ with $w(\boldsymbol{u}) = 1$, we have to evaluate that there are not $\boldsymbol{k} \in \mathbb{K}_R$ satisfying $W(\boldsymbol{u}) \succeq \boldsymbol{k}$ and $\boldsymbol{\ell} \in \mathbb{L}_R$ satisfying $W(\boldsymbol{u}) = \boldsymbol{\ell}$. Therefore, we have to get all vectors in $\mathbb{K}_R$ and $\mathbb{L}_R$, and such vectors are searched by an algorithm like breadth-first search. On the other hand, we want to show that an R-round integral characteristic cannot exist. Then, it is enough to show that $\mathbb{K}_R$ has m distinct vectors whose Hamming weight is one, i.e., there is not balanced bits, and such vectors are searched by an algorithm like depth-first search. In our provable security, we aim to get such number of rounds efficiently, and a *lazy propagation* is useful to find such number of rounds.

[4] In [16], the propagation for MISTY1 was evaluated, and the division property $\mathcal{D}_{\mathbb{K}}^{7,2,7,7,2,7,7,2,7,7,2,7}$ was used. Then, $|\mathbb{K}|$ is upper bounded by $8^8 \times 3^4 = 1358954496 \approx 2^{30.3}$, and it is feasible.

Table 7. Accurate propagations up to six rounds

#rounds	SIMON48		SIMON64		SIMON96		SIMON128	
	$\min_w(\mathbb{L})$	$\min_w(\mathbb{K})$	$\min_w(\mathbb{L})$	$\min_w(\mathbb{K})$	$\min_w(\mathbb{L})$	$\min_w(\mathbb{K})$	$\min_w(\mathbb{L})$	$\min_w(\mathbb{K})$
0	47	48	63	64	95	96	127	128
1	47	48	63	64	95	96	127	128
2	46	47	62	63	94	96	126	128
3	45	46	61	62	93	94	125	126
4	43	44	59	60	91	92	123	124
5	40	41	56	57	88	89	120	121
6	35	36	51	52	83	84	115	116

Definition 3 (Lazy Propagation). *Let $\mathcal{D}^{1^m}_{\mathbb{K}_{i-1},\mathbb{L}_{i-1}}$ be the bit-based division property of the input set of the ith round function. The ith round function is applied, and let $\mathcal{D}^{1^m}_{\bar{\mathbb{K}}_i,\bar{\mathbb{L}}_i}$ be the bit-based division property from the lazy propagation. Then, $\bar{\mathbb{K}}_i$ is computed from only a part of vectors in $\mathbb{K}_{i-1}$, and $\bar{\mathbb{L}}_i$ always becomes the empty set ϕ.*

The lazy propagation first removes all vectors from $\mathbb{L}_{i-1}$. Moreover, it only evaluates the propagation from vectors with low Hamming weight in $\mathbb{K}_{i-1}$ because such vectors are more close to unknown. Therefore, it is more efficiently evaluated than the accurate propagation.

Let us consider the meaning of the lazy propagation. Assuming the input set of the $(i-1)$th round function has $\mathcal{D}^{1^m}_{\mathbb{K}_{i-1},\mathbb{L}_{i-1}}$, we get $\mathcal{D}^{1^m}_{\bar{\mathbb{K}}_i,\mathbb{L}_i}$ and $\mathcal{D}^{1^m}_{\bar{\mathbb{K}}_i,\phi}$ by the accurate propagation and the lazy propagation, respectively. Then, the set of $\boldsymbol{u}$ that the parity is unknown is represented as

$$\mathbb{S}_{\mathbb{K}} := \{\boldsymbol{u} \in (\mathbb{F}_2)^m \mid \text{there are } \boldsymbol{k} \in \mathbb{K}_i \text{ satisfying } W(\boldsymbol{u}) \succeq \boldsymbol{k}\}\,.$$

On the other hand, $\mathbb{S}_{\bar{\mathbb{K}}_i}$ cannot completely represent the set of $\boldsymbol{u}$ that the parity is unknown. However, $\mathbb{S}_{\bar{\mathbb{K}}_i} \subseteq \mathbb{S}_{\mathbb{K}_i}$ always holds.

Next, we repeat the lazy propagation, and we assume that $\mathcal{D}^{1^m}_{\bar{\mathbb{K}}_{i+1},\phi}$ is propagated from $\mathcal{D}^{1^m}_{\bar{\mathbb{K}}_i,\phi}$ by the lazy propagation. Similarly, assuming that $\mathcal{D}^{1^m}_{\mathbb{K}_{i+1},\mathbb{L}_{i+1}}$ is the division property from $\mathcal{D}^{1^m}_{\mathbb{K}_i,\mathbb{L}_i}$ by the accurate propagation, $\mathbb{S}_{\bar{\mathbb{K}}_{i+1}} \subseteq \mathbb{S}_{\mathbb{K}_{i+1}}$ always holds because $\mathbb{S}_{\bar{\mathbb{K}}_i} \subseteq \mathbb{S}_{\mathbb{K}_i}$. Therefore, if the lazy propagation creates $\mathcal{D}^{1^m}_{\bar{\mathbb{K}}_R,\phi}$, where $\bar{\mathbb{K}}_R$ has m distinct vectors whose Hamming weight is one, the accurate propagation also creates the same m distinct vectors in the same round.

5.2 Application to Simon Family

We evaluate the lazy propagation of the bit-based division property on SIMON48, SIMON64, SIMON96, and SIMON128. Here, we only evaluate integral characteristics when they use chosen plaintexts that only one bit of the left half is constant and the other bits are active. We calculate the accurate propagation up to 6

Table 8. Lazy propagation of the bit-based division property for the SIMON family

#rounds	SIMON48		SIMON64		SIMON96		SIMON128	
	$\min_w(\mathbb{K})$	Limit	$\min_w(\mathbb{K})$	Limit	$\min_w(\mathbb{K})$	Limit	$\min_w(\mathbb{K})$	Limit
7	30	33	46	61	78	81	110	113
8	20	23	35	38	68	71	100	103
9	11	14	23	26	55	57	87	88
10	7	10	13	15	40	41	71	71
11	5	8	9	10	27	28	59	59
12	3	8	6	8	17	17	42	42
13	2	5	4	7	11	11	32	32
14	2	3	3	7	8	9	21	21
15	1	2	2	7	5	6	15	15
16	1(u)	1	2	4	4	6	10	10
17			1	3	3	6	8	8
18			1	1	2	6	5	6
19			1(u)	1	2	6	4	6
20					1	6	3	6
21					1	6	2	6
22					1	6	2	6
23					1	1	2	6
24					1(u)	1	1	6
25							1	6
26							1	6
27							1	6
28							1(u)	1

rounds[5] Table 7 shows $\min_w(\mathbb{L})$ and $\min_w(\mathbb{K})$ in the accurate propagation of $\mathcal{D}^{1^{2n}}_{\mathbb{K},\mathbb{L}}$ up to 6 rounds, where $\min_w(\mathbb{L})$ and $\min_w(\mathbb{K})$ are calculated as

$$\min_w(\mathbb{K}) = \min_{\boldsymbol{k}\in\mathbb{K}} \left(\sum_{i=1}^{2n} w(k_i) \right), \quad \min_w(\mathbb{L}) = \max_{\boldsymbol{\ell}\in\mathbb{L}} \left(\sum_{i=1}^{2n} w(\ell_i) \right).$$

From the 7th round function, we repeat the lazy propagation. We first remove all vectors from $\mathbb{L}$, and then the bit-based division property is represented as $\mathcal{D}^{1^{2n}}_{\mathbb{K},\phi}$, where ϕ denotes the empty set. Moreover, we remove vectors with high

[5] In our implementation, we could not calculate the accurate propagation up to 7 rounds because of the limitation of the memory size.

Hamming weight from $\mathbb{K}$. Table 8 shows the lazy propagation of the bit-based division property $\mathcal{D}^{1^{2n}}_{\mathbb{K},\phi}$, where we only store vectors $\boldsymbol{k} \in \mathbb{K}$ satisfying

$$\min_w(\mathbb{K}) \leq \sum_{i=1}^{2n} w(k_i) \leq \text{Limit}.$$

Here, $\mathtt{u}$ means that the $\mathbb{K}$ has $2n$ distinct vectors whose Hamming weight is one, and then, we simply say that the propagation reaches the unknown.

Even if there is a vector $\boldsymbol{k} \in \mathbb{K}$ satisfying Limit $< \sum_{i=1}^{2n} w(k_i)$, we do not evaluate the propagation from the $\boldsymbol{k}$. Therefore, if the propagation from the removed vector $\boldsymbol{k}$ immediately reaches the unknown, there is a gap between the accurate propagation and the lazy propagation. However, if the lazy propagation reaches the unknown in a specific number of rounds, the accurate propagation at least reaches the unknown in the same number of rounds. Therefore, the lazy propagation is not useful for attackers, but it guarantees the number of rounds that the bit-based division property cannot find integral characteristics.

As a result, the lazy propagation shows that 16-, 19-, 24-, and 28-round SIMON48, 64, 96, and 128 probably do not have integral characteristics, respectively. However, we can get $(r+1)$-round integral characteristics from r-round integral characteristics because round keys are XORed after the round function. Therefore, we expect that 17-, 20-, 25-, and 29-round SIMON48, 64, 96, and 128 probably do not have integral characteristics, respectively.

5.3 Characteristics that Bit-Based Division Property Cannot Find

We consider characteristics that the bit-based division property cannot find. Our provable security supposes that all round keys are randomly and secretly chosen. However, practical ciphers generate round keys from the secret key using the key scheduling algorithm. Therefore, our provable security does not suppose integral characteristics that exploit the key scheduling algorithm.

The bit-based division property using three subsets focuses on the parity $\bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{\boldsymbol{u}}(\boldsymbol{x})$, and divide the set of $\boldsymbol{u}$ into three subsets. Then, the propagation simply regard $\bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{\boldsymbol{u}_1}(\boldsymbol{x}) \oplus \pi_{\boldsymbol{u}_2}(\boldsymbol{x})$ as unknown if either $\bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{\boldsymbol{u}_1}(\boldsymbol{x})$ or $\bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{\boldsymbol{u}_2}(\boldsymbol{x})$ is unknown. For instance, if $\bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{\boldsymbol{u}_1}(\boldsymbol{x}) \oplus \pi_{\boldsymbol{u}_2}(\boldsymbol{x})$ is always 0 or 1 although $\bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{\boldsymbol{u}_1}(\boldsymbol{x})$ and $\bigoplus_{\boldsymbol{x}\in\mathbb{X}} \pi_{\boldsymbol{u}_2}(\boldsymbol{x})$ are unknown, the bit-based division property cannot exploit such property.

6 Conclusions

The division property is a useful technique to find integral characteristics, but it has not been applied to non-S-box-based ciphers effectively. This paper focused on the bit-based division property. More precisely, this paper proposed a new variant using three subsets. The conventional bit-based division property divides the set of $\boldsymbol{u}$ into two subsets, but the new variant divides the set of $\boldsymbol{u}$ into three subsets. The bit-based division property using three subsets can prove that the

experimental integral characteristic for SIMON32 shown in [18] works for all keys. Moreover, we focused on the propagation of the division property. Then, we showed that the lazy propagation is useful to guarantee the security against integral cryptanalyses using the division property. As a result, we showed that 17-, 20-, 25-, and 29-round SIMON48, 64, 96, and 128 probably do not have integral characteristics, respectively.

Acknowledgments. The authors would like to thank the anonymous referees for their helpful comments.

References

1. Abed, F., List, E., Lucks, S., Wenzel, J.: Differential cryptanalysis of round-reduced Simon and speck. In: Cid, C., Rechberger, C. (eds.) FSE 2014. LNCS, vol. 8540, pp. 525–545. Springer, Heidelberg (2015)
2. Aumasson, J.P., Jovanovic, P., Neves, S.: Norx v2.0, submission to CAESAR competition (2015)
3. Beaulieu, R., Shors, D., Smith, J., Treatman-Clark, S., Weeks, B., Wingers, L.: The SIMON and SPECK families of lightweight block ciphers. IACR Cryptology ePrint Archive 2013/404 (2013). http://eprint.iacr.org/2013/404
4. Beaulieu, R., Shors, D., Smith, J., Treatman-Clark, S., Weeks, B., Wingers, L.: SIMON and SPECK: block ciphers for the internet of things. IACR Cryptology ePrint Archive 2015/585 (2015). http://eprint.iacr.org/2015/585
5. Biryukov, A., Roy, A., Velichkov, V.: Differential analysis of block ciphers SIMON and SPECK. In: Cid, C., Rechberger, C. (eds.) FSE 2014. LNCS, vol. 8540, pp. 546–570. Springer, Heidelberg (2015)
6. Boura, C., Naya-Plasencia, M., Suder, V.: Scrutinizing and improving impossible differential attacks: applications to CLEFIA, Camellia, LBlock and SIMON. In: Sarkar, P., Iwata, T. (eds.) ASIACRYPT 2014. LNCS, vol. 8873, pp. 179–199. Springer, Heidelberg (2014)
7. Daemen, J., Knudsen, L.R., Rijmen, V.: The block cipher SQUARE. In: Biham, E. (ed.) FSE 1997. LNCS, vol. 1267, pp. 149–165. Springer, Heidelberg (1997)
8. Knudsen, L.R.: Truncated and higher order differentials. In: Preneel, B. (ed.) FSE. LNCS, vol. 1008, pp. 196–211. Springer, Heidelberg (1994)
9. Knudsen, L.R., Wagner, D.: Integral cryptanalysis. In: Daemen, J., Rijmen, V. (eds.) FSE 2002. LNCS, vol. 2365, pp. 112–127. Springer, Heidelberg (2002)
10. Kölbl, S., Leander, G., Tiessen, T.: Observations on the SIMON block cipher family. In: Gennaro, R., Robshaw, M. (eds.) CRYPTO 2015. LNCS, vol. 9215, pp. 161–185. Springer, Heidelberg (2015)
11. Lai, X.: Higher order derivatives and differential cryptanalysis. In: Blahut, R.E., Costello Jr., D.J., Maurer, U., Mittelholzer, T. (eds.) CC. SISECS, vol. 276, pp. 227–233. Springer, Heidelberg (1994)
12. Matsui, M.: New structure of block ciphers with provable security against differential and linear cryptanalysis. In: Gollmann, D. (ed.) FSE. LNCS, vol. 1039, pp. 205–218. Springer, Heidelberg (1996)
13. Nyberg, K., Knudsen, L.R.: Provable security against a differential attack. J. Cryptol. **8**(1), 27–37 (1995)

14. Sun, B., Hai, X., Zhang, W., Cheng, L., Yang, Z.: New observation on division property. IACR Cryptology ePrint Archive 2015/459 (2015). http://eprint.iacr.org/2015/459
15. Sun, S., Hu, L., Wang, P., Qiao, K., Ma, X., Song, L.: Automatic security evaluation and (Related-key) differential characteristic search: application to SIMON, PRESENT, LBlock, DES(L) and other bit-oriented block ciphers. In: Sarkar, P., Iwata, T. (eds.) ASIACRYPT 2014. LNCS, vol. 8873, pp. 158–178. Springer, Heidelberg (2014)
16. Todo, Y.: Integral Cryptanalysis on Full MISTY1. In: Gennaro, R., Robshaw, M. (eds.) CRYPTO 2015. LNCS, vol. 9215, pp. 413–432. Springer, Heidelberg (2015)
17. Todo, Y.: Structural evaluation by generalized integral property. In: Oswald, E., Fischlin, M. (eds.) EUROCRYPT 2015. LNCS, vol. 9056, pp. 287–314. Springer, Heidelberg (2015)
18. Wang, Q., Liu, Z., Varici, K., Sasaki, Y., Rijmen, V., Todo, Y.: Cryptanalysis of reduced-round SIMON32 and SIMON48. In: Meier, W., Mukhopadhyay, D. (eds.) INDOCRYPT 2014. LNCS, vol. 8885, pp. 143–160. Springer, Heidelberg (2014)
19. Yang, G., Zhu, B., Suder, V., Aagaard, M.D., Gong, G.: The simeck family of lightweight block ciphers. In: Güneysu, T., Handschuh, H. (eds.) CHES 2015. LNCS, vol. 9293, pp. 307–329. Springer, Heidelberg (2015)
20. Zhang, H., Wu, W.: Structural evaluation for generalized feistel structures and applications to Lblock and TWINE. In: Biryukov, A., Goyal, V. (eds.) INDOCRYPT 2015. LNCS, vol. 9462, pp. 218–237. Springer, Heidelberg (2015)
21. Zhang, H., Wu, W., Wang, Y.: Integral attack against bit-oriented block ciphers. In: Kwon, S., Yun, A. (eds.) ICISC 2015. LNCS, vol. 9558, pp. 102–118. Springer, Heidelberg (2016). doi:10.1007/978-3-319-30840-1_7