IET Information Security

Research Article

IET Journals
The Institution of Engineering and Technology

New zero-sum distinguishers on full 24-round KECCAK-f using the division property

ISSN 1751-8709
Received on 28th May 2018
Revised 4th February 2019
Accepted on 25th February 2019
E-First on 16th April 2019
doi: 10.1049/iet-ifs.2018.5263
www.ietdl.org

Hailun Yan[1], Xuejia Lai[1,2] ✉, Lei Wang[1,2], Yu Yu[1,2], Yiran Xing[1]

[1]*Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China*
[2]*Westone Cryptologic Research Center, Beijing 100070, People's Republic of China*
✉ *E-mail: laix@sjtu.edu.cn*

Abstract: The authors analyse the security of KECCAK (the winner in SHA-3 competition) by focusing on the zero-sum distinguishers of its underlying permutation (named KECCAK-*f*). The authors' analyses are developed by using the division property, a generalised integral property that was initially used in the integral cryptanalysis of symmetric-key algorithms. Following the work pioneered by Todo at CRYPTO 2015, they first formalise and prove a more delicate propagation rule of the division property under the assumption that the S-box's specification is known to attackers. Then, they apply this rule to the inverse S-box in KECCAK-*f* with a further study on properties of its algebraic degree. They find that the rate of decline in the division property is gentler than that of a randomly chosen S-box. Meanwhile, they get the same results for the S-box in ASCON permutation. Thanks to this vulnerable property, they can improve the higher-order differential characteristics against the inverse of KECCAK-*f* in terms of the required number of chosen plaintexts. As an application, they give new zero-sum distinguishers on full 24-round KECCAK-*f* of size 2^{1573}. To the authors' knowledge, this is currently the best zero-sum distinguishers of full-round KECCAK-*f* permutation. Incidentally, they give the corresponding results for 12-round ASCON permutation.

1 Introduction

The KECCAK hash function family was proposed by Bertoni *et al.* [1], and follows the so-called sponge construction. It was selected as the winner of the SHA-3 competition [2] and subsequently standardised as SHA-3 standard [3]. Since the design was made public, KECCAK has attracted intensive cryptanalysis [4–11], including zero-sum distinguishers on its underlying permutation KECCAK-*f*.

Zero-sum distinguishers were introduced by Aumasson and Meier at CHES 2009 [4]. In order to construct zero-sum distinguishers of a permutation, attackers need to search for a set of zero-sum input values (called *zero-sum*), of which the corresponding output values also sum to zero. In other words, for a permutation P, they find a set V of values such that $\sum_{v\in V} v = 0$ and $\sum_{v\in V} P(v) = 0$. Given all inputs in V except one and the P-outputs of all inputs of V but one, the existence of zero-sum distinguishers allows to compute the missing input and output by simply summing over the known elements without calling P. The small size of zero-sum may give adversaries an advantage in an attack. In the public comment on the NIST Hash competition 2010 [12], zero-sum distinguishers are shown to be valid and qualitatively different from generic methods, albeit with a very small advantage. The existence of the zero-sum distinguisher over 16 rounds of KECCAK-f[1600] in [4] motivated the designers to increase the total number of rounds for KECCAK-*f* permutation in order to increase the security margin against possible attacks, since the hermetic sponge strategy adopted in the KECCAK sponge function impose the underlying permutation to be free from structural distinguishers [13, 14].

In practice, the search for zero-sum distinguishers on a hash function usually begins in the intermediate state and extends to both ends, since there is no secret parameter in the internal functions. Meanwhile, the analysis mainly uses the higher-order differential technique [15] combined with the degree estimation on the iterated round functions. More specifically, for an r-round iterated permutation $P = p^r$, suppose that the algebraic degrees of the $r-t$ forward rounds p^{r-t} and the t backward rounds p^{-t} are bounded by d. According to a well-known higher-order differential property, for a Boolean function f with $\deg(f) \le d$, the outputs sum to zero if the inputs form a linear subspace of dimension $d+1$. Then, by choosing a linear subspace V_M of size 2^{d+1} in the intermediate state after t rounds, both the inputs and outputs of P sum to zero, i.e. $\sum_{v\in V_M} p^{r-t}(v) = 0$ and $\sum_{v\in V_M} p^{-t}(v) = 0$. The *zero-sum* V of the r-round zero-sum distinguisher can be got by $V = \{p^{-t}(v) \,|\, v \in V_M\}$. See Fig. 1.

Zero-sum distinguishers on KECCAK-*f* permutation mainly exploit the following property of the only non-linear step in the round function: the algebraic degree of the 5-bit S-box is 2 and the algebraic degree of its inverse is 3, which is (relatively) low. Until now, there have been distinguishers of 9–15 rounds for KECCAK-*f* permutation of 1600 bits, with complexities bounded by 2^{800} (birthday bound) [4, 9]. Besides this, Guo *et al.* [16] formalised the idea of linear structures and inserted three rounds for free in the middle. Compared with the previous zero-sum distinguishers in [4], they generally increased the rounds of attack by two without increasing the complexity. Zero-sum distinguishers with more rounds on the KECCAK-*f* permutation were given by Boura *et al.* [17, 18], where the zero-sum of the full 24-round KECCAK-*f* is of size 2^{1590}. Then, Duan and Lai [19] improved Boura *et al.*'s work by lowering the size of the zero-sum to 2^{1575}.

Another important work related to our paper is the division property, which was proposed by Todo [20] at EUROCRYPT 2015. It generalised the integral property [21] by further exploiting the algebraic degree of block ciphers. Therefore, division property usually leads to more accurate higher-order differential characteristics for many cryptographic primitives. We briefly recall the concept of division property here and give the formal definition later. When we say the multi-set $\mathbb{X}$ whose elements take values from $\mathbb{F}_2^n$ has the division property $\mathcal{D}_k^n$, it satisfies such properties: for $0 \le i < k$, if we evaluate the product of any i bits of the elements in $\mathbb{X}$, then the sum of the evaluation over $\mathbb{X}$ is zero.

$$\text{zero-sum} \xleftarrow[\deg(p^{-t}) \le d]{t \text{ rounds backward}} V_M \xrightarrow[\deg(p^{r-t}) \le d]{r\text{-}t \text{ rounds forward}} \text{zero-sum}$$

Fig. 1 *Zero-sum distinguishers constructed with degree estimation*

Table 1 Data complexity of *r*-round higher-order differential (integral) characteristic on KECCAK-*f* and its inverse

Target application	$\log_2$(#texts)								Method	Reference
	$r=8$	$r=9$	$r=10$	$r=11$	$r=12$	$r=13$	$r=14$	$r=15$		
(5,2,320)-SPN KECCAK-*f*	257	513	1025	1409	1537	1579	1593	1598	degree	[18]
	130	258	515	1025	1410	1538	1580	1595	division property Algorithm 1 (Fig. 7)	[20]
(5,3,320)-SPN KECCAK-f^{-1}	1383	1492	1546	1573	1587	1594	1597	1599	degree division property Algorithm 1 (Fig. 7)	[18, 19]
	1310	1505	1570	1590	1599	—	—	—		[20]
	1165	*1384*	*1494*	*1546*	*1574*	*1589*	*1595*	1599	division property Algorithm 2 (Fig. 8)	Section 4.2

Table 2 Division property's propagation characteristic of the inverse S-box in KECCAK-*f* (the same for ASCON permutation)

$\mathcal{D}_k^5$ for input set $\mathbb{X}$	$\mathcal{D}_1^5$	$\mathcal{D}_2^5$	$\mathcal{D}_3^5$	$\mathcal{D}_4^5$	$\mathcal{D}_5^5$
$\mathcal{D}_k^5$ for output set $\mathbb{Y}$	$\mathcal{D}_1^5$	$\mathcal{D}_1^5$	$\mathcal{D}_1^5$	$\mathcal{D}_2^5$	$\mathcal{D}_5^5$
$\mathcal{D}_k^5$ for output set $\mathbb{Y}$ (improved)	$\mathcal{D}_1^5$	$\mathcal{D}_1^5$	$\mathcal{D}_1^5$	$\mathcal{D}_3^5$	$\mathcal{D}_5^5$

Moreover, for $k \le i \le n$, the sum becomes unknown. According to this definition, when the multi-set is of $\mathcal{D}_2^n$, the sum of all the elements in the multi-set is zero.

Todo gave some propagation rules (Substitution, Copy, Compression by XOR, Split and Concatenation) of the division property in [20]. Note that in the propagation rule of Substitution, it is assumed that nothing of the S-box's specification is known to the attackers but the algebraic degree. While in practice, the details of the S-box are usually public knowledge. For the 7-bit S-box S_7 in MISTY1 [22], Todo gave an improved propagation characteristic of the division property by further considering its specification [23]. Based on this, Todo found an integral distinguisher of 6-round on MISTY1 and successfully mounted the full-round integral attack on MISTY1. However, for the division property's propagation characteristic specific to S_7, no generic method was given or proved.

Moreover, Todo proposed an algorithm in [20] based on the division property to search for higher-order differential (integral) distinguishers. Although the attack developed by his algorithm is *generic*, it can make an improvement on the higher-order differential distinguishers of several *specific* primitives, including KECCAK-*f*. Compared with the previous zero-sum distinguishers on KECCAK-*f* constructed by Boura *et al.* [18] using a degree estimation, Todo's algorithm can make a reduction on the number of the required chosen data (see Table 1). Then, a natural question is that, if such things also happened for the inverse of KECCAK-*f*, it might be possible to construct a better zero-sum distinguisher with a smaller size. However, for the higher-order differential distinguishers against the inverse of KECCAK-*f*, our experiments show that Todo's *generic* search algorithm has less advantage over Duan and Lai's degree estimation [19]. This is mainly because Duan and Lai have made an improvement in Boura *et al.*'s *generic* degree estimation by considering the inherent properties of KECCAK-*f*. In this paper, we consider how to improve Todo's search algorithm for the inverse of KECCAK-*f* permutation with a further study on properties of its non-linear layer.

Following the pioneering work of Todo in [20, 23], there have been many extensions and applications of the division property [24–30]. Particularly, in [31], Xiang *et al.* extended mixed-integer linear programming (MILP) method, which is used to search differential characteristics and linear trails of block ciphers, to search integral distinguishers of block ciphers based on division property. In [27], Sun *et al.* proposed automatic tools to detect ARX ciphers' division property relying on Boolean Satisfiability (SAT) problem. In [28], Todo *et al.* exploit the division property to mount cube attacks on non-blackbox polynomials, which is the first application of the division property to stream ciphers.

Our contribution: In general, there are mainly two different approaches to improve the zero-sum distinguishers, finding tighter upper bounds on the degree of iterated permutations [18, 19], or trying to insert more rounds for free in the middle [4, 16, 17]. In this paper, our work is in a new direction. We use the division property to improve the zero-sum distinguishers on KECCAK-*f*.

In our work, we first formalise and prove the propagation rule of the division property used by Todo [23], combined with the weight-degree table. We call it *table-based propagation rule of the division property*. Then, we apply this rule to the inverse of the S-box used in the round permutation of KECCAK-*f* (see Table 2). We find that, the rate of decline in the division property is gentler than that of a randomly chosen S-box. More specifically, for the inverse of the S-box in KECCAK-*f*, the deterioration of the division property $\mathcal{D}_4^5$ can be improved from $\mathcal{D}_2^5$ to $\mathcal{D}_3^5$.

Combined with the above property, we improve Todo's search algorithm for the inverse of KECCAK-*f*, getting better higher-order differential distinguishers. The results are summarised in Table 1, where the inverse of KECCAK-*f* is denoted by KECCAK-f^{-1}. Finally, we obtain new zero-sum distinguishers on 23-round KECCAK-*f* of size 2^{1546} with higher-order differential characteristics of 12 rounds forward and 11 rounds backward, which are connected by the propagation of the division property. By transforming the vectorial division property to the general division property in the intermediate part, we obtain zero-sum distinguishers on full 24-round KECCAK-*f* of size 2^{1573}. We have made an improvement compared with the previous best one of size 2^{1575} [19]. Although our improvement seems limited, it is a novel application of the division property to hash functions, and offers new possibilities to search for zero-sum distinguishers.

Our technique applies to other permutations. As a by-product of our research, we show our analytic results of the inner permutation of ASCON, which is one of the finalists in the CAESAR competition. For ASCON permutation, the propagation characteristic of the division property against its S-box is exactly the same as that against χ in KECCAK, evaluated by our table-based propagation rule of the division property (see Table 2). Therefore, we can also improve the higher-order differential characteristics against the inverse of ASCON permutation in terms of the required number of chosen plaintexts (see Table 3). We get zero-sum distinguishers on 12-round ASCON permutation of size 2^{130}, which is in accordance with distinguishers given by Dobraunig *et al.* at CT-RSA 2015 [32]. We have summarised the corresponding results of the ASCON permutation in Appendix 2. We do not show the details in this paper considering the technical similarities.

2 Preliminaries

2.1 Basic definitions

(n, d, m)-*SPN (substitute-permutation network)* [20]. SPN is a basic structure in designing block ciphers. The round function of an SPN iterated block cipher is composed of the S-Layer and the P-Layer. For an (n, d, m)-SPN, there are m n-bit parallel S-boxes in the S-Layer with degree at most d. Also, there is one $(m \cdot n)$-bit linear function in the P-Layer.

Hamming weight. Let $a = (a_1, a_2, \ldots, a_n) \in \mathbb{F}_2^n$. Define the Hamming weight of a as

$$w(a) = \sum_{i=1}^{n} a_i .$$

Let $\boldsymbol{a} = \left(a^{(1)}, \ldots, a^{(m)}\right) \in \left(\mathbb{F}_2^n\right)^m$. Define the extended Hamming weight of $\boldsymbol{a}$ as

$$W(\boldsymbol{a}) = \left(w\left(a^{(1)}\right), w\left(a^{(2)}\right), \ldots, w\left(a^{(m)}\right)\right).$$

Bit product function. Let $u = (u_1, u_2, \ldots, u_n)$, $x = (x_1, x_2, \ldots, x_n) \in \mathbb{F}_2^n$. Define the bit product function $\pi_u : \mathbb{F}_2^n \to \mathbb{F}_2$ as

$$\pi_u(x) = \prod_{i=1}^{n} x_i^{u_i} .$$

Let $\boldsymbol{u} = \left(u^{(1)}, \ldots, u^{(m)}\right) \in \left(\mathbb{F}_2^n\right)^m$, $\boldsymbol{x} = \left(x^{(1)}, \ldots, x^{(m)}\right) \in \left(\mathbb{F}_2^n\right)^m$. The bit product function $\pi_{\boldsymbol{u}} : \left(\mathbb{F}_2^n\right)^m \to \mathbb{F}_2$ is defined as

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

Algebraic normal form (ANF). For a Boolean function $f : \mathbb{F}_2^n \to \mathbb{F}_2$, its ANF is expressed as

$$f(x) = \bigoplus_{u \in \mathbb{F}_2^n} a_u^f \pi_u(x) .$$

The *algebraic degree* of f is defined as $\max\{w(u) \mid a_u^f \neq 0\}$, denoted by $\deg(f)$. Note that an n-bit S-box is a vectorial Boolean function from $\mathbb{F}_2^n$ to $\mathbb{F}_2^n$. The algebraic degree of an S-box is defined as the maximal algebraic degree of these n Boolean function coordinates.

2.2 Division property

In this part, we first give summary of the definition of the division property. Then we present several propagation rules of the division property related to this paper. For a better understanding of this part, please refer to [20, 23].

In [20], Todo gave three types of division properties, which are the (general) division property, vectorial division property and collective division property. In the following, we only formally introduce the definition of collective division property, where the other two types of division property can be regarded as its special case.

Definition 1 (Collective division property): Denote by $\mathbb{S}$ a multi-set whose elements take values in $\left(\mathbb{F}_2^n\right)^m$. Let $\boldsymbol{k}^{(i)} = \left(k_1^{(i)}, k_2^{(i)}, \ldots, k_m^{(i)}\right)$, where $k_j^{(i)}, 1 \le j \le m$ take values between 0 and n. When we say that the multi-set $\mathbb{S}$ has the *collective division property* $\mathcal{D}^{n,m}_{\boldsymbol{k}^{(1)}, \ldots, \boldsymbol{k}^{(q)}}$, it fulfils: $\bigoplus_{\boldsymbol{x} \in \mathbb{S}} \pi_{\boldsymbol{u}}(\boldsymbol{x}) = 0$ if $W(\boldsymbol{u}) \not\succeq \boldsymbol{k}^{(i)}, 1 \le i \le q\}$. Otherwise, $\bigoplus_{\boldsymbol{x} \in \mathbb{S}} \pi_{\boldsymbol{u}}(\boldsymbol{x})$ becomes unknown. Note that for $\boldsymbol{k} = (k_1, k_2, \ldots, k_m), \boldsymbol{k}' = (k_1', k_2', \ldots, k_m')$, $\boldsymbol{k} \succeq \boldsymbol{k}'$ means that $k_i \ge k_i'$ for all $1 \le i \le m$. Otherwise, $\boldsymbol{k} \not\succeq \boldsymbol{k}'$.

In particular, for $q = 1$, the multi-set $\mathbb{S}$ whose elements take values in $\left(\mathbb{F}_2^n\right)^m$ has the *vectorial division property* $\mathcal{D}^{n,m}_{\boldsymbol{k}}$. For $m = 1$, $q = 1$, the multi-set $\mathbb{S}$ whose elements take values in $\mathbb{F}_2^n$ has the *division property* $\mathcal{D}^n_k$. Namely, $\bigoplus_{x \in \mathbb{S}} \pi_u(x) = 0$ when $w(u) < k$ and $\bigoplus_{x \in \mathbb{S}} \pi_u(x)$ becomes unknown when $w(u) \ge k$.

Given the division property of the input multi-set of an S-box, in order to evaluate the division property of the output multi-set, Todo gave the following propagation characteristic [20], which we call the *generalised propagation rule of the division property*.

Property 1: For an S-box $S : \mathbb{F}_2^n \to \mathbb{F}_2^n$ with degree d, the division property $\mathcal{D}^n_k$ of the input multi-set propagates to the division property $\mathcal{D}^n_{\lceil \frac{k}{d} \rceil}$ of the output multi-set. Particularly, suppose that the S-box is a permutation, then $\mathcal{D}^n_n$ propagates to $\mathcal{D}^n_n$.

Rule 1 (Substitution) [23]*:* Let F: $\left(\mathbb{F}_2^n\right)^m \to \left(\mathbb{F}_2^n\right)^m$ be a function, which is composed of m n-bit S-boxes $(S_1, \ldots, S_m)$ defined over $\mathbb{F}_2^n$ with algebraic degree d. When the input multi-set has the collective division property $\mathcal{D}^{n,m}_{\boldsymbol{k}^{(1)}, \ldots, \boldsymbol{k}^{(q)}}$, the output multi-set has the collective division property $\mathcal{D}^{n,m}_{\boldsymbol{k}'^{(1)}, \ldots, \boldsymbol{k}'^{(q)}}$, which is calculated as

$$k_i'^{(j)} = \left\lceil \frac{k_i^{(j)}}{d} \right\rceil, \text{ where} 1 \le i \le m \text{ and } 1 \le j \le q .$$

Rule 2 (Concatenation) [23]: Let F be a concatenation function defined over $\left(\mathbb{F}_2^{n_1} \times \mathbb{F}_2^{n_2}\right)$, $F(x_1, x_2) = x_1 \parallel x_2$. Through F, the collective division property $\mathcal{D}^{n_1, n_2}_{\boldsymbol{k}^{(1)}, \ldots, \boldsymbol{k}^{(q)}}$ of the input multi-set propagates to the division property $\mathcal{D}^{n_1+n_2}_{k'}$ of the output multi-set, which is calculated as

$$k' = \min\left\{k_1^{(1)} + k_2^{(1)}, k_1^{(2)} + k_2^{(2)}, \ldots, k_1^{(q)} + k_2^{(q)}\right\} .$$

Rule 3 (Partition) [23]: Let F be a split function from $\mathbb{F}_2^n$ into $\left(\mathbb{F}_2^{n_1} \times \mathbb{F}_2^{n_2}\right)$, where $n = n_1 + n_2$. $F(x_1 \parallel x_2) = (x_1, x_2)$. Through F, the division property $\mathcal{D}^n_k$ of the input multi-set propagates to the collective division property $\mathcal{D}^{n_1, n_2}_{\boldsymbol{k}^{(1)}, \ldots, \boldsymbol{k}^{(q)}}$ of the output multi-set, which is calculated as

$$\boldsymbol{k}^{(i+1)} = (k - i, i), 0 \le i \le k .$$

Here $(k - i) \le n_1, i \le n_2$.

2.3 KECCAK-f permutation

The operation mode of KECCAK hash function is the sponge construction [14], which is depicted in Fig. 2. There are seven choices of the inherent permutation of KECCAK, indicated by KECCAK-$f[b]$, where $b = 25 \times 2^i$ $(i = 0, 1, \ldots, 6)$ is the input (and the output) size of the permutation. The KECCAK-f permutation used in SHA-3 is 1600 bits, i.e. KECCAK-$f[1600]$, which is the case we consider in this paper. In all the following, we denote KECCAK-$f[1600]$ by KECCAK-f for simplicity.

Table 3 Data complexity of r-round higher-order differential (integral) characteristic on ASCON permutation and its inverse

Target application	$\log_2$(#texts)											Method
	$r=1$	$r=2$	$r=3$	$r=4$	$r=5$	$r=6$	$r=7$	$r=8$	$r=9$	$r=10$	$r=11$	
(5,2,64)-SPN ASCON permutation	4	5	5	10	18	35	65	130	258	300	315	division property Algorithm 1 (Fig. 7)
(5,3,64)-SPN inverse of ASCON permutation	4	5	10	30	84	245	295	314	–	–	–	division property Algorithm 1 (Fig. 7)
	4	5	10	29	84	201	261	291	306	314	319	division property Algorithm 2 (Fig. 8)

The state of KECCAK-*f* permutation can be represented as a $5 \times 5 \times 64$ state, like that in Fig. 3. KECCAK-*f* has 24 rounds in total. The round permutation R consists of five operations:

$$R = \iota \circ \chi \circ \pi \circ \rho \circ \theta .$$

The first three steps θ, ρ, π and the last step ι are all linear transformations. The only non-linear step χ is composed of 320 parallel 5-bit S-box χ_0 in each row. We show the Boolean expression of χ_0 and χ_0^{-1} in Tables 4 and 5, respectively (see Appendix 1). Note that the degree of χ_0 is 2 and the degree of its inverse χ_0^{-1} is 3. The linear function provides diffusion in all the three dimensions of the state while the non-linear function is applied to each row to provide confusion.

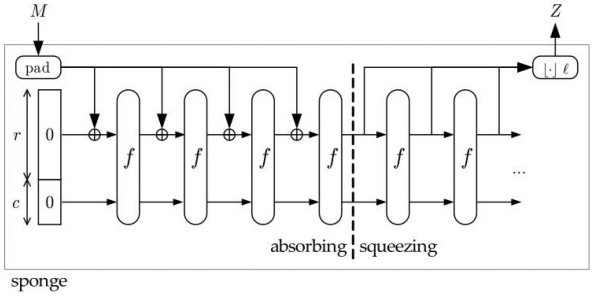

Fig. 2 *Sponge construction*

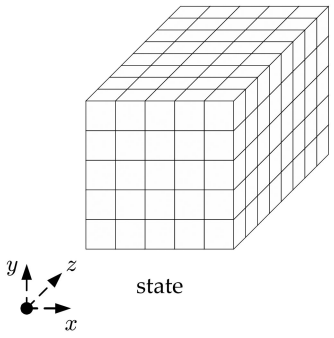

Fig. 3 *KECCAK state*

Table 4 Boolean expression of the 5-bit S-box χ_0 in KECCAK-*f*

Input	Output	Corresponding Boolean function
x_0	y_0	$x_0 + x_2 + x_1x_2$
x_1	y_1	$x_1 + x_3 + x_2x_3$
x_2	y_2	$x_2 + x_4 + x_3x_4$
x_3	y_3	$x_0 + x_3 + x_0x_4$
x_4	y_4	$x_1 + x_4 + x_0x_1$

Table 5 Boolean expression of the 5-bit S-box χ_0^{-1}

Input	Output	Corresponding Boolean function
y_0	x_0	$y_0 + y_2 + y_4 + y_1y_2 + y_1y_4 + y_3y_4 + y_1y_3y_4$
y_1	x_1	$y_0 + y_1 + y_3 + y_0y_2 + y_0y_4 + y_2y_3 + y_0y_2y_4$
y_2	x_2	$y_1 + y_2 + y_4 + y_0y_1 + y_1y_3 + y_3y_4 + y_0y_1y_3$
y_3	x_3	$y_0 + y_2 + y_3 + y_0y_4 + y_1y_2 + y_2y_4 + y_1y_2y_4$
y_4	x_4	$y_1 + y_3 + y_4 + y_0y_1 + y_0y_3 + y_2y_3 + y_0y_2y_3$

In [18], Boura *et al.* gave a bound on the degree of the iterated round permutations of KECCAK-*f* and of its inverse. Then Duan and Lai [19] improved the upper bounds on the degree of the inverse of KECCAK-*f*. We summarised the results in Table 6. Based on the degree estimation of the iterated round permutations in KECCAK-*f* and its inverse, Duan and Lai gave a zero-sum distinguisher of size 2^{1575} for the full KECCAK-*f* permutation. Note that the degree of the backward 11 (resp. forward 12) rounds for KECCAK-*f* is bounded by 1572 (resp. 1536), if we consider the internal states before the non-linear layer χ in the 12th round, we can construct zero-sum distinguishers by choosing any subspace V_M with active bits in 315 rows. Refer to Fig. 4.

3 Table-based propagation rule of the division property

In this part, we formalise the propagation rule of the division property under the assumption that the specification of an S-box is public. We call it *table-based propagation rule of the division property*, since the evaluation is supported by a weight-degree table.

Note that for the generalised propagation rule of the division property (Property 1), it is assumed that attackers treat the S-box as a black box, with the only knowledge of its algebraic degree. Such propagation rule of the division property against S-box is applicable to any S-box but does not exploit the specific feature of the S-box. In [23], Todo gave a more accurate evaluation of the division property's propagation against the 7-bit S-box S_7 in MISTY1 by regarding it as a public function. As a result, for the S-box S_7, the division property $\mathcal{D}_6^7$ can propagate to $\mathcal{D}_4^7$, with a smaller deterioration than that of a 7-bit black-box S-box, against which $\mathcal{D}_6^7$ propagates to $\mathcal{D}_2^7$. Thanks to the vulnerable property of S_7, 6-round integral distinguisher was found and the integral attack has been successfully mounted on full MISTY1.

Todo's work in [23] is instructive. When considering the specification of a concrete S-box, it might be possible to slow down the deterioration of the division property against the S-box, which potentially leads to a higher-order differential distinguisher of more rounds. However, Todo's work is customised for the S-box in MISTY1. A generic method has not yet been given. It is the first time that we formalise and prove such an improved propagation rule.

Before introducing our propagation rule, we first define the product degree of an S-box following [18].

Definition 2: For an n-bit S-box, define the product degree δ_i as the maximal algebraic degree of the product of any i ($1 \le i \le n$) output coordinates of the S-box.

We call the table that shows the product degree δ_i of an n-bit S-box for each i ($1 \le i \le n$) the *weight-degree table of the S-box*.

Example 1:

We calculate the product degree δ_i of χ_0^{-1} for $1 \le i \le 5$, and show its weight-degree table in Table 7. Recall that χ_0^{-1} is the

Table 6 Degree estimation of round-reduced KECCAK-*f* and its inverse

Number of rounds r	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
bound on $\deg(R^r)$	2	4	8	16	32	64	128	256	512	1024	1408	1536	1578	1592	1597	1599
bound on $\deg(R^{-r})$	3	9	27	81	243	729	1164	1382	1491	1545	1572	1586	1593	1596	1598	1599

$$\text{zero-sum} \xleftarrow[\deg(R^{-11}) \le 1572]{\text{11 rounds backward}} V_M \xrightarrow{\chi} V_M \xrightarrow[\deg(R^{12}) \le 1536]{\text{12 rounds forward}} \text{zero-sum}$$

Fig. 4 *Zero-sum distinguishers on full-round KECCAK-f*

inverse of the 5-bit S-box χ_0 in KECCAK-*f*, with the algebraic degree equal to 3. One important property which will be used in the following is that the algebraic degree of the product of any two output coordinates of χ_0^{-1} is 3, i.e. $\delta_2 = 3$. We show the Boolean expression in detail in Table 8. Note that this property has also been exploited by Duan and Lai [19].

Property 2: (Table-based propagation rule of division property): Let $S:\mathbb{F}_2^n \to \mathbb{F}_2^n$ be a function (S-box) and δ_i $(1 \le i \le n)$ be its product degree. Then the division property $\mathcal{D}_k^n$ of the input multi-set propagates to the division property $\mathcal{D}_{l_k}^n$ of the output multi-set, where $l_k = \min\{i|\delta_i \ge k, 1 \le i \le n\}$. Particularly, if S is a permutation, then $\mathcal{D}_n^n$ propagates to $\mathcal{D}_n^n$.

Proof: First, we consider the division property of the input multi-set $\mathbb{X}$. According to the definition of $\mathcal{D}_k^n$, we have

$$\begin{cases} \oplus_{x\in\mathbb{X}} \pi_u(x) = 0 & w(u) < k, \\ \oplus_{x\in\mathbb{X}} \pi_u(x) \text{isunknown} & w(u) \ge k. \end{cases}$$

Then for all functions $f:\mathbb{F}_2^n \to \mathbb{F}_2$ with $\deg(f) < k$, we have that $\oplus_{x\in\mathbb{X}} f(x) = 0$.

Next, we consider the division property of the output multi-set $\mathbb{Y}$. Suppose that $\mathbb{Y}$ has the division property $\mathcal{D}_{k'}^n$. For any $v \in \mathbb{F}_2^n$ such that $w(v) < k'$, to ensure that

$$\bigoplus_{y\in\mathbb{Y}} \pi_v(y) = \bigoplus_{x\in\mathbb{X}} \pi_v \circ S(x) = 0,$$

we only need to ensure that $\deg(\pi_v \circ S(x)) < k$. When $\deg(\pi_v \circ S(x)) \ge k$, the sum will become unknown. We illustrate such 'division' in the following figure:

$w(v) < k'$	$w(v) \ge k'$
$\deg(\pi_v \circ S(x)) < k$	$\deg(\pi_v \circ S(x)) \ge k$

According to the definition of the product degree δ_i of the S-box, the degree of $\pi_v \circ S(x)$ is bounded by $\delta_{w(v)}$. Let $k' = \min\{i|\delta_i \ge k, 1 \le i \le n\}$. Then, when $w(v) < k'$, $\deg(\pi_v \circ S(x)) \le \delta_{w(v)} < k$, and we always have that $\oplus_{x\in\mathbb{X}} \pi_v \circ S(x) = 0$. □

Remark 1: Some related results are stated in [33, 34]. We revisited the division property in different views. Moreover, it is the first time that we formalise the improved *propagation rule* of the division property, which is depicted in a clear way and quite easy to follow.

The difference between the table-based propagation rule in Property 2 and the generalised propagation rule in Property 1 is that, in Property 1, the degree of $\pi_v \circ S$ is roughly estimated by $\min\{n, w(v) \times d\}$ (see [20]), while in Property 2 it is calculated more accurately by $\delta_{w(v)}$ under the assumption that the S-box is public. Here we take χ_0^{-1} of KECCAK-*f* as an example. For χ_0^{-1} and $v \in \mathbb{F}_2^5$, when $w(v) = 2$, it is estimated in Property 1 that the degree of $\pi_v \circ \chi_0^{-1}$ is bounded by $\min\{5, w(v) \times 3\} = \min\{5, 6\} = 5$. In fact, the algebraic degree of the product of any two output coordinates of χ_0^{-1} is lower than 4. However, it does not mean that the table-based propagation rule would certainly be better than the generalised propagation rule, where 'better' refers to that the decline of the propagation characteristic is smaller. For example, for (5,3,320)-SPN KECCAK-*f* and (5,3,64)-SPN ASCON permutation [35], our propagation rule can improve the characteristic of the division property against their S-boxes, while for (8,7,16)-SPN AES [36] and (4,3,16)-SPN PRESENT [37], the S-boxes of which both have the highest algebraic degrees, the evaluations of the propagation are the same by using these two rules. So, the question is, when does the evaluation by using Property 2 have an advantage over that by using Property 1? We give the answer in Corollary 1.

Corollary 1: Let $S:\mathbb{F}_2^n \to \mathbb{F}_2^n$ be a public S-box with algebraic degree d. Let $t = \left\lceil \frac{k}{d} \right\rceil$, $1 \le k, t \le n$. Consider the propagation of the division property against S. If $\delta_t < k$, then the deterioration of the division property evaluated by Property 2 is smaller than that evaluated by Property 1.

Proof: The outline of the proof is shown in Fig. 5. According to Definition 2, if $\delta_t < k$, then $l_k > t$. □

3.1 Application to χ_0^{-1}: the inverse of the S-box used in KECCAK-*f*

We consider the propagation of the division property of χ_0^{-1}. Especially, we consider the propagation of $\mathcal{D}_4^5$ against χ_0^{-1}. Recall that the degree of the product of any two output components of χ_0^{-1} is bounded by 3 (see Table 7). When the input multi-set $\mathbb{X}$ which takes values over $\mathbb{F}_2^5$ fulfils $\mathcal{D}_4^5$, it means that the sum over the input multi-set of the output of any function with degree less than 4 is zero, i.e. $\oplus_{x\in\mathbb{X}} f(x) = 0$ for any function $f:\mathbb{F}_2^n \to \mathbb{F}_2$ with $\deg(f) < 4$. In this case, $\oplus_{x\in\mathbb{X}} \pi_v \circ \chi_0^{-1}(x) = 0$ when $w(v) = 2$, which implies that $k' \ge 3$ for the division property $\mathcal{D}_{k'}^5$ of the output multi-set.

If we only exploit the algebraic degree, then according to Property 1, we obtain

$$\mathcal{D}_4^5 = \mathcal{D}_k^n \xrightarrow{\chi_0^{-1}} \mathcal{D}_{\left\lceil \frac{k}{d} \right\rceil}^5 = \mathcal{D}_{\left\lceil \frac{4}{3} \right\rceil}^5 = \mathcal{D}_2^5.$$

Table 7 Weight-degree table of the inverse S-box in KECCAK-*f* (the same for ASCON permutation)

i	1	2	3	4	5
δ_i	3	3	4	4	5

Table 8 Product of any two output components of χ_0^{-1}

Output	Corresponding Boolean function
y_0y_1	$x_0 + x_0x_1 + x_0x_2 + x_0x_3 + x_0x_4 + x_0x_2x_3 + x_0x_2x_4$
y_0y_2	$x_2 + x_4 + x_0x_2 + x_0x_4 + x_1x_2 + x_1x_4 + x_3x_4 + x_0x_3x_4 + x_1x_3x_4$
y_0y_3	$x_0 + x_2 + x_0x_3 + x_0x_4 + x_1x_2 + x_2x_3 + x_2x_4 + x_3x_4 + x_1x_2x_3 + x_1x_2x_4$
y_0y_4	$x_4 + x_0x_3 + x_0x_4 + x_1x_4 + x_2x_4 + x_1x_2x_4 + x_1x_3x_4$
y_1y_2	$x_1 + x_0x_1 + x_1x_2 + x_1x_3 + x_1x_4 + x_0x_1x_3 + x_1x_3x_4$
y_1y_3	$x_0 + x_3 + x_0x_1 + x_0x_2 + x_0x_4 + x_1x_3 + x_2x_3 + x_0x_1x_4 + x_0x_2x_4$
y_1y_4	$x_1 + x_3 + x_0x_1 + x_0x_3 + x_1x_4 + x_2x_3 + x_3x_4 + x_0x_2x_3 + x_2x_3x_4$
y_2y_3	$x_2 + x_0x_2 + x_1x_2 + x_2x_3 + x_2x_4 + x_0x_2x_4 + x_1x_2x_4$
y_2y_4	$x_1 + x_4 + x_0x_1 + x_1x_2 + x_1x_3 + x_2x_4 + x_3x_4 + x_0x_1x_2 + x_0x_1x_3 + x_0x_3x_4 + x_2x_3x_4$
y_3y_4	$x_3 + x_0x_3 + x_1x_3 + x_2x_3 + x_3x_4 + x_0x_1x_3 + x_0x_2x_3$

Since $\delta_2 < 4$, according to Corollary 1, it is possible to improve the propagation characteristic against χ_0^{-1} by further considering the product degree in Table 7. According to our table-based propagation rule of the division property in Property 2, we obtain

$$\mathcal{D}_4^5 = \mathcal{D}_k^n \xrightarrow{\chi_0^{-1}} \mathcal{D}_{l_k}^n = \mathcal{D}_{l_4}^5 = \mathcal{D}_3^5.$$

Recall that l_4 is defined as $\min\{i | \delta_i \geq 4, 1 \leq i \leq n\}$.

Following Property 2, we evaluate the table-based propagation rule of the division property against χ_0^{-1}. We list the results in Table 2.

Property 3: For χ_0^{-1}, the division property $\mathcal{D}_4^5$ propagates to $\mathcal{D}_3^5$. For an S-box randomly chosen from all 5-bit S-boxes, about which we only know that the algebraic degree is 3, $\mathcal{D}_4^5$ propagates to $\mathcal{D}_2^5$.

For the sake of descriptive completeness, we calculate the product degree δ_i of χ_0 for $1 \leq i \leq 5$, and show its weight-degree table in Table 9. The division property's propagation characteristic of χ_0 evaluated by our table-based propagation rule (Property 2) is the same as that evaluated by Todo's generic propagation rule (Property 1). See Table 10.

4 Improved higher-order differential characteristics for the inverse of KECCAK-*f*

In this section, we construct better higher-order differential characteristics of the inverse of KECCAK-*f* permutation, based on the improved propagation characteristic of the division property.

In [20], Todo showed generic attacks against (n, d, m)-SPN with an improvement to the higher-order differential (integral) characteristics. We say the attacks are *generic* because they are structural cryptanalyses assuming that attackers can only exploit the property of the structure but not the specific weaknesses of a specific cipher. For the higher-order differential characteristics of KECCAK-*f* permutation (R^r:(5,2,320)-SPN), Todo's attack can reduce data complexity compared with previous ones constructed by Boura *et al.*'s using degree estimation [18]. However, when it comes to the inverse of KECCAK-*f* (R^{-r}: (5,3,320)-SPN), Todo's method seems having no advantage (in terms of the data complexity) over Duan's degree estimation [18, 19]. The comparison is shown in Table 1. This is mainly because for the inverse of KECCAK-*f*, Duan has made an improvement to Boura *et al.*'s *generic* degree estimation by considering its particular properties. Therefore, we consider to exploit the inherent properties of the inverse of the non-linear layer.

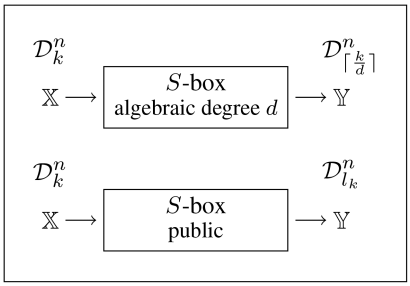

Fig. 5 *Propagation characteristic of division property*

Table 9 Weight-degree table of χ_0 in KECCAK-*f* (the same for ASCON permutation)

i	1	2	3	4	5
δ_i	2	4	4	4	5

Before introducing our work, we first recall the propagation characteristic for (n, d, m)-SPN and Todo's path search algorithm.

4.1 Propagation rule of division property against (n, d, m)-SPN [20]

The input set of the S-Layer: For the S-Layer, we prepare the input set satisfying the following properties: the input of the i-th n-bit S-box S_i consists of k_i active bits and $(n - k_i)$ constant bits. In this case, the input multi-set has the vectorial division property $\mathcal{D}_{\boldsymbol{k}}^{n,m}$, where $\boldsymbol{k} = (k_1, k_2, \ldots, k_m)$. The number of the chosen data is $2^{\sum_{i=1}^m k_i}$.

S-Layer: The vectorial division property $\mathcal{D}_{\boldsymbol{k}'}^{n,m}$ of the output set against the S-Layer is calculated according to the rule of Substitution (Rule 1), where $\boldsymbol{k}' = (k_1', k_2', \ldots, k_m')$.

$$\begin{cases} k_i' = \left\lceil \frac{k_i}{d} \right\rceil & \text{if } k_i < n, \\ k_i' = n & \text{if } k_i = n. \end{cases}$$

Concatenation (Conversion from S-Layer to P-Layer): From S-Layer to P-Layer, the vectorial division property $\mathcal{D}_{\boldsymbol{k}'}^{n,m}$ converts to the division property $\mathcal{D}_k^{nm}$, which follows Rule 2: $k = \sum_{i=1}^m k_i'$.

P-Layer: The division property does not change after applying the P-Layer.

Split (Conversion from P-Layer to S-Layer): The conversion of the division property $\mathcal{D}_k^{nm}$ from the P-Layer to the next S-Layer follows Rule 3. The S-Layer has the collective division property $\mathcal{D}_{\boldsymbol{k}'^{(1)}, \boldsymbol{k}'^{(2)}, \ldots, \boldsymbol{k}'^{(q)}}^{n,m}$, where q denotes the number of all possible vectors satisfying $k_1'^{(j)} + k_2'^{(j)} + \cdots + k_m'^{(j)} = k (1 \leq j \leq q)$.

Fig. 6 shows the propagation of the division property against (n, d, m)-SPN. With the increment of the number of rounds, the number q of all possible vectors in the Split step grows very fast. In this case, Todo gave a more efficient technique to control its size. If $k \leq (n-1)m$, the round function of (n, d, m)-SPN is regarded as one big S-box of nm bits, with algebraic degree d. And the propagation of the division property follows the rule of Substitution. If $k > (n-1)m$, then at least $k - (n-1)m = m - nm + k$ elements of $\boldsymbol{k}'^{(j)}$ have to become n, which are still n after applying the S-Layer. In this case, the rest elements have to become $n-1$, which will propagate to $\left\lceil \frac{n-1}{d} \right\rceil$ through the S-Layer. Then the division property $\mathcal{D}_{k'}^{nm}$ of the output multi-set is calculated as

$$k' = \begin{cases} \left\lceil \frac{n-1}{d} \right\rceil (nm - k) + n(m - nm + k), \\ \quad \text{for } k > (n-1)m, \\ \left\lceil \frac{k}{d} \right\rceil, \quad \text{for } k \leq (n-1)m. \end{cases}$$

The algorithm to construct the higher-order differential characteristics is shown in Algorithm 1 (see Fig. 7). When we use 2^D chosen plaintexts to construct the characteristics, we choose the initial vectorial division property $\mathcal{D}_{\boldsymbol{k}}^{n,m}$ satisfying the following properties:

$$k_i = \begin{cases} n & \text{for } i \cdot n \leq D, \\ D - (i-1)n & \text{for } (i-1)n \leq D < i \cdot n, \\ 0 & \text{for } D < (i-1)n. \end{cases}$$

Table 10 Propagation characteristic of χ_0 in KECCAK-*f* (the same for ASCON permutation)

$\mathcal{D}_k^5$ for input set $\mathbb{X}$	$\mathcal{D}_1^5$	$\mathcal{D}_2^5$	$\mathcal{D}_3^5$	$\mathcal{D}_4^5$	$\mathcal{D}_5^5$
$\mathcal{D}_k^5$ for output set $\mathbb{Y}$	$\mathcal{D}_1^5$	$\mathcal{D}_1^5$	$\mathcal{D}_2^5$	$\mathcal{D}_2^5$	$\mathcal{D}_5^5$

4.2 Specific search algorithm for the higher-order differential characteristics of the inverse of KECCAK-*f*

We improve the search algorithm of the higher-order differential characteristics for the inverse KECCAK-*f* permutation. Note that in Algorithm 1 (Fig. 7), the evaluation of the division property's propagation in the S-Layer only exploits the S-box's algebraic degree (in Lines 8 and 10). Compared with Algorithm 1 (Fig. 7), the major improvement is that $\mathscr{D}_4^5$ propagates to $\mathscr{D}_3^5$ instead of $\mathscr{D}_2^5$ against the S-Layer. More specifically, if $k > (n-1)m$, at least $(m - nm + k)$ elements of $\boldsymbol{k}'^{(j)}$ have to become n, then all the rest have to become $n-1$. n and $n-1$ propagate through the non-linear layer χ to n and l_{n-1}, respectively. Similarly, if $(n-2)m < k \le (n-1)m$, at least $k-(n-2)m = (2m - nm + k)$ elements of $\boldsymbol{k}'^{(j)}$ have to become $n-1$, then the rest of the $nm - m - k$ elements have to become $n-2$ and propagate to $\left\lceil \frac{n-2}{d} \right\rceil$. Otherwise, we also regard the round permutation as one big nm-bit S-box with degree d. We show the improved search algorithm targeted at the iterated round permutations of the inverse of KECCAK-*f* in Algorithm 2 (Fig. 8). We improve the evaluation of the division property's propagation characteristic for KECCAK-*f* (in Lines 11–17) by calculating k' as

$$k' = \begin{cases} l_{n-1}(nm-k) + n(k+m-nm), \\ \quad \text{for } k > (n-1)m, \\ \left\lceil \frac{n-2}{d} \right\rceil (nm-m-k) + l_{n-1}(k+2m-nm), \\ \quad \text{for } (n-2)m < k \le (n-1)m, \\ \left\lceil \frac{k}{d} \right\rceil, \quad \text{for } k \le (n-2)m\,. \end{cases}$$

We compared our distinguishers on the inverse of KECCAK-*f* with previous ones in Table 1. Thanks to the vulnerable property of χ_0^{-1} that $\mathscr{D}_3^5$ is provided from $\mathscr{D}_4^5$, we can improve the higher-order differential characteristics of the inverse of KECCAK-*f* with less chosen data than that in [19]. For example, for the 9-round distinguisher on the inverse of KECCAK-*f*, the number of the required chosen data is reduced from 2^{1505} to 2^{1384}.

5 New zero-sum distinguishers on full-round KECCAK-*f* permutation

In this section, we show improved zero-sum distinguishers of size 2^{1573} on full 24-round KECCAK-*f* permutation by using the division property. Meanwhile, we still use the inside-out technique. Our new zero-sum distinguishers have made an improvement compared with the previous best one of size 2^{1575}.

Let $F_{13} = R^{12} \circ \iota \circ \chi$ and (see (1)) .

According to our results got by implementing Algorithm 2 (Fig. 8) in Table 1, there exists a higher-order differential characteristic of F_{13} and G'_{11} with data complexity 2^{1538} and 2^{1546}, respectively. At first glance, it seems that we could combine the forward 13-round characteristic and backward 11-round characteristic together and obtain a zero-sum distinguisher of size $2^{\max\{1538,1456\}} = 2^{1538}$. However, we regret to say that it is impossible. This is mainly because, the chosen input sets of the two characteristics got by implementing Algorithm 2 (Fig. 8) have the vectorial division property $\mathscr{D}_{\boldsymbol{k}}^{5,320}$ and $\mathscr{D}_{\boldsymbol{k}'}^{5,320}$, respectively, with the 320-dimensional vector $\boldsymbol{k} = (\underbrace{5,\ldots,5}_{307}, 3, \underbrace{0,\ldots,0}_{12})$ and $\boldsymbol{k}' = (\underbrace{5,\ldots,5}_{309}, 1, \underbrace{0,\ldots,0}_{10})$. We cannot ensure that $\mathscr{D}_{\boldsymbol{k}}^{5,320}$ propagates to $\mathscr{D}_{\boldsymbol{k}'}^{5,320}$ through the intermediate P-layer ($\pi \circ \rho \circ \theta \circ \iota$).

23-round zero-sum distinguishers for KECCAK-f of size 2^{1546}. We evaluate the propagation the division property through the intermediate P-layer by using mixed-integer linear programming (MILP) method [31]. We get zero-sum distinguishers of size 2^{1546} on 23-round KECCAK-*f* permutation.

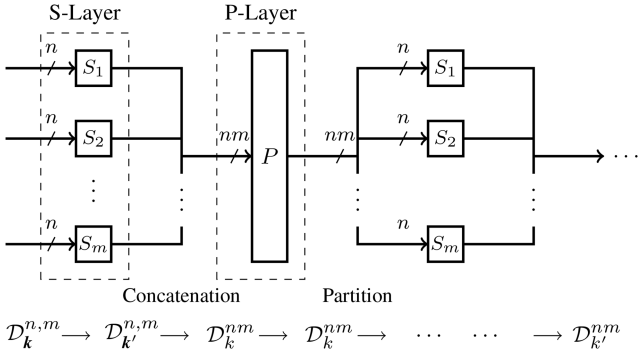

Fig. 6 *Propagation of the division property for* (n, d, m)*-SPN*

Require: $n, m, d, r = 0, k_1, k_2, \ldots, k_m$;
1: **if** $k_i < n$ **then**
2: $k_i \leftarrow \lceil \frac{k_i}{d} \rceil$
3: **end if**
4: $k \leftarrow \sum_{i=1}^{m} k_i$
5: **while** $1 < k$ **do**
6: $r \leftarrow r + 1$
7: **if** $k \le (n-1)m$ **then**
8: $k \leftarrow \lceil \frac{k}{d} \rceil$
9: **else**
10: $k \leftarrow \lceil \frac{n-1}{d} \rceil (nm-k) + n(m-nm+k)$
11: **end if**
12: **end while**
13: **return** r

Fig. 7 *Algorithm 1: Path search algorithm for higher-order differential (integral) characteristic of* (n, d, m)*-SPN [20]*

More specifically, we consider the internal states after the non-linear layer χ of the 11th round in KECCAK-*f*, and prepare a linear space of size 2^{1546}. Let $\mathbb{X}$ be the input multi-set of G'_{11}, the elements of which take values from $(\mathbb{F}_2^5)^{320}$. Suppose that $\mathbb{X}$ fulfils the division property $\mathscr{D}_{\boldsymbol{k}'}^{5,320}$, where $\boldsymbol{k}' = (\underbrace{5,\ldots,5}_{309}, 1, \underbrace{0,\ldots,0}_{10})$.

- *Backward 11 rounds.* A 11-round higher-order differential characteristic of G'_{11} can be ensured according to Table 1. $\sum_{x \in \mathbb{X}} G'_{11}(x) = 0$.
- *Forward 12 rounds.* Let $\mathbb{Y}$ be the input multi-set of F_{12} $(= R^{11} \circ \iota \circ \chi)$, the elements of which take values from $(\mathbb{F}_2^5)^{320}$. Suppose that $\mathbb{Y}$ fulfils the division property $\mathscr{D}_{\boldsymbol{k}''}^{5,320}$. Luckily, we find that $\mathscr{D}_{\boldsymbol{k}'}^{5,320}$ propagates to $\mathscr{D}_{\boldsymbol{k}''}^{5,320}$ against the P-Layer by using MILP method, where $\boldsymbol{k}'' = (\underbrace{5,\ldots,5}_{307}, 3, \underbrace{0,\ldots,0}_{12})$. Then, $\mathscr{D}_{\boldsymbol{k}''}^{5,320}$ propagates to $\mathscr{D}_2^{1600}$ through F_{12} (implementing Algorithm 1 (Fig. 7), as shown in the following equation:

$$\begin{aligned} \mathscr{D}_{\boldsymbol{k}''}^{5,320} &\to \mathscr{D}_{1438}^{1600} \to \mathscr{D}_{1114}^{1600} \to \mathscr{D}_{557}^{1600} \\ &\to \mathscr{D}_{279}^{1600} \to \mathscr{D}_{140}^{1600} \to \mathscr{D}_{70}^{1600} \to \mathscr{D}_{35}^{1600}. \\ &\to \mathscr{D}_{18}^{1600} \to \mathscr{D}_{9}^{1600} \to \mathscr{D}_{5}^{1600} \to \mathscr{D}_{2}^{1600} \end{aligned}$$

Therefore, $\sum_{y \in \mathbb{Y}} F_{12}(y) = 0$. The zero-sum V of the 23-round zero-sum distinguisher can be got by $V = \{G'_{11}(x) | x \in \mathbb{X}\}$. $|V| = |\mathbb{X}| = 2^{1546}$.

Remark 2: We have also considered the propagation of the division property against the P-Layer in the backward direction. Unfortunately one vectorial division property will propagate too many collective division property, which makes the following analysis very hard.

$$G_{11} = R^{-11} \circ \theta^{-1} \circ \rho^{-1} \circ \pi^{-1} \triangleq G'_{11} \circ \iota^{-1} \circ \theta^{-1} \circ \rho^{-1} \circ \pi^{-1} \triangleq G'_{11} \circ P\,. \tag{1}$$

24-round zero-sum distinguishers for KECCAK-f of size 2^{1573}. By transforming the vectorial division property to the general division property, we get zero-sum distinguishers of size 2^{1573} on the full 24-round KECCAK-*f* permutation. Refer to Fig. 9.

We consider the internal states before the non-linear layer χ of the 12th round in KECCAK-*f*, and prepare a linear space of size 2^{1573}. Let $\mathbb{X}$ be the input multi-set of F_{13}, the elements of which take values from $(\mathbb{F}_2^5)^{320}$. Suppose that $\mathbb{X}$ fulfils the division property $\mathcal{D}_{\boldsymbol{k}}^{5,320}$, where $\boldsymbol{k} = (\underbrace{5,\ldots,5}_{314},3,\underbrace{0,\ldots,0}_{5})$.

- *Forward 13 rounds.* A 13-round higher-order differential characteristic of F_{13} can be ensured according to Table 1. $\sum_{x\in\mathbb{X}} F_{13}(x) = 0$.
- *Backward 11 rounds.* Let $\mathbb{Y}$ be the input multi-set of G_{11}, the elements of which take values from $\mathbb{F}_2^{1600}$. Suppose that $\mathbb{Y}$ fulfils division property $\mathcal{D}_k^{1600}$. The conversion from $\mathcal{D}_{\boldsymbol{k}}^{5,320}$ to $\mathcal{D}_k^{1600}$ follows the rule of Concatenation, thereby $k = 1573$. The division property does not change in the P-Layer. Then, $\mathcal{D}_{1573}^{1600}$ propagates to $\mathcal{D}_2^{1600}$ through G'_{11} (implementing lines 9–19 of Algorithm 2 (Fig. 8), as shown in the following equation:

$$\begin{aligned}\mathcal{D}_{1573}^{1600} &\to \mathcal{D}_{1550}^{1600} \to \mathcal{D}_{1500}^{1600} \to \mathcal{D}_{1400}^{1600}\\ &\to \mathcal{D}_{1200}^{1600} \to \mathcal{D}_{800}^{1600} \to \mathcal{D}_{207}^{1600} \to \mathcal{D}_{89}^{1600}\,.\\ &\to \mathcal{D}_{30}^{1600} \to \mathcal{D}_{10}^{1600} \to \mathcal{D}_{4}^{1600} \to \mathcal{D}_{2}^{1600}\end{aligned}$$

Therefore, $\sum_{y\in\mathbb{Y}} G_{11}(y) = 0$. The zero-sum V of the 24-round zero-sum distinguisher can be got by $V = \{G_{11}(y)|y \in \mathbb{Y}\}$. $|V| = |\mathbb{Y}| = 2^{1573}$.

Require: $n = 5, d = 3, m = 320, r = 0, l_{n-1} = 3, k_1, k_2, \ldots, k_m$;
1: **if** $k_i < n$ **then**
2: **if** $k_i = n - 1$ **then**
3: $k_i \leftarrow l_{n-1}$ ▷ 1-st round S-Layer
4: **else**
5: $k_i \leftarrow \lceil \frac{k_i}{d} \rceil$ ▷ 1-st round S-Layer
6: **end if**
7: **end if**
8: $k \leftarrow \sum_{i=1}^{m} k_i$ ▷ 1-st round Concatenation and P-Layer
9: **while** $1 < k$ **do**
10: $r \leftarrow r + 1$
11: **if** $k \le (n-2)m$ **then**
12: $k \leftarrow \lceil \frac{k}{d} \rceil$ ▷ (r+1)-th round
13: **else if** $(n-2)m < k \le (n-1)m$ **then**
14: $k \leftarrow \lceil \frac{n-2}{d} \rceil (nm - m - k) + l_{n-1}(2m - nm + k)$
15: ▷ (r+1)-th round
16: **else if** $k > (n-1)m$ **then**
17: $k \leftarrow l_{n-1}(nm - k) + n(k + m - nm)$
18: ▷ (r+1)-th round
19: **end if**
20: **end while**
21: **return** r

Fig. 8 *Algorithm 2: Searching for higher-order differential (integral) characteristics of the inverse KECCAK-f*

6 Conclusions and discussions

In this paper, we first formalise and prove a more delicate propagation rule of the division property under the assumption that the S-box's specification is known to attackers. Based on this rule, we find an interesting property of the inverse of the S-box in KECCAK-*f*: the rate of decline in the division property $\left(\mathcal{D}_4^5 \to \mathcal{D}_3^5\right)$ is gentler than that of a randomly chosen S-box $\left(\mathcal{D}_4^5 \to \mathcal{D}_2^5\right)$. Thanks to this vulnerable property, we can improve the higher-order differential characteristics against the inverse of KECCAK-*f* in terms of the required number of chosen plaintexts. As an application, we gave zero-sum distinguishers on 23-round KECCAK-*f* of size 2^{1546} and zero-sum distinguishers on full 24-round KECCAK-*f* of size 2^{1573}.

Discussion 1: Our technique applies to other permutations. What we need to emphasise is that, it is not certain whether or not the table-based propagation rule of the division property takes an advantage over Todo's generic propagation rule. It depends on the specific S-box (Table 2 versus Table 10), as the table-based propagation rule is closely related to the product degree of the S-box. We analysed the inner permutation of ASCON [35], which uses approximately the same 5-bit S-Box as KECCAK (refer to Tables 11 and 12 in Appendix 2). Our results show that the propagation characteristic of the division property against its S-box is exactly the same as that against χ in KECCAK, evaluated by our table-based propagation rule of the division property (see Table 2). Another point worth noting is that, with an improvement in the division property's propagation characteristic of the S-box, it is also not certain whether we can improve the higher-order differential distinguishers and zero-sum distinguishers on the corresponding SPN permutation, as the situation varies from round to round. In general, distinguishers covering more rounds have a greater advantage. For the inverse of ASCON permutation, we improve the higher-order differential characteristics in terms of the required number of chosen plaintexts when it covers more than five rounds (see Table 3). We get zero-sum distinguishers on 12-round ASCON permutation of size 2^{130} (refer to Appendix 2), which is in accordance with distinguishers given by Dobraunig *et al.* at CT-RSA 2015 [32].

Discussion 2: While this paper was under submission, we noticed that Wang *et al.* released a paper on the ePrint Archive (subsequently published at CT-RSA 2018 [38]), in which they also construct zero-sum distinguishers based on the division property and applied it to the full permutation of some PHOTON variants. Unlike our method, their search for the forward and backward

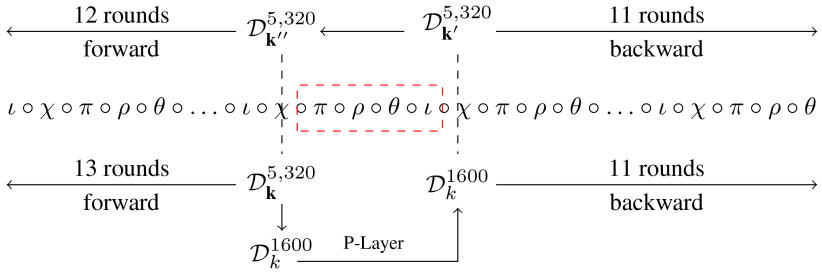

Fig. 9 *Constructing zero-sum distinguishers on full 24-round KECCAK-f permutation*

Table 11 Boolean expression of the 5-bit S-box S in ASCON permutation

Input	Output	Corresponding Boolean function
x_0	y_0	$x_0 + x_1 + x_2 + x_3 + x_0x_1 + x_1x_2 + x_1x_4$
x_1	y_1	$x_0 + x_1 + x_2 + x_3 + x_4 + x_1x_2 + x_1x_3 + x_2x_3$
x_2	y_2	$x_1 + x_2 + x_4 + x_3x_4 + 1$
x_3	y_3	$x_0 + x_1 + x_2 + x_3 + x_4 + x_0x_3 + x_0x_4$
x_4	y_4	$x_1 + x_3 + x_4 + x_0x_1 + x_1x_4$

Table 12 Boolean expression of the inverse of the above 5-bit S-box S

Input	Output	Corresponding Boolean function
y_0	x_0	$y_1 + y_2 + y_3 + y_0y_1 + y_2y_3 + y_0y_2y_3 + y_0y_3y_4 + y_1y_3y_4 + y_2y_3y_4 + 1$
y_1	x_1	$y_0 + y_1 + y_4 + y_0y_2 + y_2y_3 + y_0y_2y_4$
y_2	x_2	$y_0 + y_1 + y_2 + y_4 + y_0y_2 + y_1y_2 + y_1y_3$
		$+y_2y_3 + y_2y_4 + y_3y_4 + y_0y_1y_2$
		$+y_0y_1y_3 + y_0y_2y_4 + y_1y_2y_4 + y_1y_3y_4 + 1$
y_3	x_3	$y_1 + y_3 + y_4 + y_0y_2 + y_1y_2$
		$+y_1y_4 + y_2y_4 + y_0y_2y_4 + y_1y_2y_4$
y_4	x_4	$y_3 + y_0y_1 + y_0y_2 + y_1y_2 + y_2y_3 + y_2y_4$
		$+y_0y_2y_3 + y_0y_2y_4 + y_1y_2y_4 + y_2y_3y_4$

higher-order differential characteristics is based on the bit-based division property and MILP method. In fact, we have also considered the MILP method to characterise the division property propagation for full-round KECCAK-f. But the computational complexity is too high to get satisfactory results since the 1600-bit permutation size of KECCAK-f is relatively large.

Future work: Further work might be done in two directions: Excavating special properties in KECCAK's linear layer such that the propagation of the division property in large-size permutation can be evaluated effectively; applying Wang *et al.*'s technique to some other small-size permutations like the inner permutation of ASCON. Besides, it is interesting to consider whether it is possible to improve the degree bound of the iterated permutations by using the division property.

7 Acknowledgments

The authors thank the anonymous reviewers for their helpful comments. This work was supported by the National Natural Science Foundation of China (61702331, 61472251, U1536101, 61602302, 61472250, 61672347), 13th five-year National Development Fund of Cryptography (MMJJ20170105, MMJJ20170114), Natural Science Foundation of Shanghai (16ZR1416400), Shanghai Excellent Academic Leader Funds (16XD1401300) and Science and Technology on Communication Security Laboratory.

8 References

[1] Bertoni, G., Daemen, J., Peeters, M., *et al.*: 'The Keccak reference'. https://keccak.team/files/Keccak-reference-3.0.pdf, (January 2011) Version 3.0

[2] NIST.: Sha-3 competition. http://csrc.nist.gov/groups/ST/hash/sha-3/index.html, 2002–2012

[3] Dworkin, M.J.: 'Sha-3 standard: permutation-based hash and extendable-output functions', Federal Inf. Process. Stds. (NIST FIPS)-202, 2015

[4] Aumasson, J.P., Meier, W.: 'Zero-sum distinguishers for reduced keccak-f and for the core functions of Luffa and Hamsi'. Rump Session of Cryptographic Hardware and Embedded Systems-CHES, 2009, vol. 2009, p. 67

[5] Daemen, J., Van Assche, G.: 'Differential propagation analysis of keccak'. Fast Software Encryption, Washington, USA, 2012, vol. 7549, pp. 422–441

[6] Dinur, I., Dunkelman, O., Shamir, A.: 'New attacks on keccak-224 and keccak-256'. Fast Software Encryption, Washington, USA, 2012, vol 12, pp. 442–461

[7] Dinur, I., Dunkelman, O., Shamir, A.: 'Collision attacks on up to 5 rounds of SHA-3 using generalized internal differentials'. Fast Software Encryption, Singapore, Singapore, 2013, pp. 219–240

[8] Dinur, I., Morawiecki, P., Pieprzyk, J., *et al.*: 'Cube attacks and cube-attack-like cryptanalysis on the round-reduced keccak sponge function'. Advances in Cryptology – EUROCRYPT 2015, Sofia, Bulgaria, 2015, pp. 733–761

[9] Jean, J., Nikolić, I.: 'Internal differential boomerangs: practical analysis of the round-reduced keccak-f permutation'. Fast Software Encryption, Istanbul, Turkey, 2015, pp. 537–556

[10] Mendel, F., Nad, T., Schläffer, M.: 'Finding SHA-2 characteristics: searching through a minefield of contradictions'. Advances in Cryptology – ASIACRYPT 2011, Seoul, Korea, 2011, pp. 288–307

[11] Naya Plasencia, M., Röck, A., Meier, W.: 'Practical analysis of reduced-round Keccak'. INDOCRYPT, Chennai, India, 2011, Vol. 7107, pp. 236–254

[12] Bertoni, G., Daemen, J., Peeters, M., *et al.*: 'Note on zero-sum distinguishers of Keccak-f. Public comment on the NIST Hash competition (2010)', https://keccak.team/files/NoteZeroSum.pdf

[13] Bertoni, G., Daemen, J., Peeters, M., *et al.*: 'On the indifferentiability of the sponge construction'. Annual Int. Conf. on the Theory and Applications of Cryptographic Techniques, Berlin, Heidelberg, 2008, pp. 181–197

[14] Bertoni, G., Daemen, J., Peeters, M., *et al.*: 'Cryptographic sponges', http://sponge.noekeon.org/, 2009

[15] Lai, X.: 'Higher order derivatives and differential cryptanalysis' *Communications and Cryptography* (Springer, Boston, MA, 1994), pp. 227–233

[16] Guo, J., Liu, M., Song, L.: 'Linear structures: applications to cryptanalysis of round-reduced keccak'. ASIACRYPT 2016, Hanoi, Vietnam, 2016, vol. 2016, pp. 249–274

[17] Boura, C., Canteaut, A.: 'Zero-sum distinguishers for iterated permutations and application to keccak-f and hamsi-256'. Selected Areas in Cryptography, Waterloo, Canada, 2010, pp. 1–17

[18] Boura, C., Canteaut, A., De Canniere, C.: 'Higher-Order differential properties of keccak and Luffa'. Fast Software Encryption, Lyngby, Denmark, 2011, vol. 6733, pp. 252–269

[19] Duan, M., Lai, X.J.: 'Improved zero-sum distinguisher for full round keccak-f permutation', *Chin. Sci. Bull.*, 2012, **57**, (6), pp. 694–697

[20] Todo, Y.: 'Structural evaluation by generalized integral property'. Advances in Cryptology – EUROCRYPT 2015, Sofia, Bulgaria, 2015, pp. 287–314

[21] Knudsen, L., Wagner, D.: 'Integral cryptanalysis'. Fast Software Encryption, Leuven, Belgium, 2002, pp. 629–632

[22] Matsui, M.: 'New block encryption algorithm MISTY'. Fast Software Encryption, Haifa, Israel, 1997, Vol. 1267, pp. 54–68

[23] Todo, Y.: 'Integral cryptanalysis on full MISTY1', *J. Cryptol.*, 2017, **30**, (3), pp. 920–959

[24] Todo, Y., Morii, M.: 'Bit-based division property and application to simon family'. Int. Conf. on Fast Software Encryption, Berlin, Heidelberg, 2016, pp. 357–377

[25] Xiang, Z., Zhang, W., Lin, D.: 'On the division property of SIMON48 and SIMON64'. Int. Workshop on Security, Cham, 2016, pp. 147–163

[26] Todo, Y.: 'Division property: efficient method to estimate upper bound of algebraic degree'. Int. Conf. on Cryptology in Malaysia, Cham, 2016, pp. 553–571

[27] Sun, L., Wang, W., Wang, M.: 'Automatic search of bit-based division property for ARX ciphers and word-based division property'. Int. Conf. on the Theory and Application of Cryptology and Information Security, Cham, 2017, pp. 128–157

[28] Todo, Y., Isobe, T., Hao, Y., *et al.*: 'Cube attacks on non-blackbox polynomials based on division property', *IEEE Trans. Comput.*, 2018, **67**, (12), pp. 1720–1736

[29] Wang, S., Hu, B., Guan, J., *et al.*: 'MILP method of searching integral distinguishers based on division property using three subsets', IACR ePrint Report 2018/1186, https://eprint.iacr.org/2018/1186.pdf

[30] Hu, K., Wang, M.: 'Automatic Search for a Variant of Division Property Using Three Subsets'. IACR ePrint Report 2018/1187, https://eprint.iacr.org/2018/1187.pdf

[31] Xiang, Z., Zhang, W., Bao, Z., *et al.*: 'Applying MILP method to searching integral distinguishers based on division property for 6 lightweight block

ciphers'. Int. Conf. on the Theory and Application of Cryptology and Information Security, Berlin, Heidelberg, 2016, pp. 648–678
[32] Dobraunig, C., Eichlseder, M., Mendel, F., *et al.*: 'Cryptanalysis of ascon'. Cryptographersąf Track at the RSA Conf., Cham, 2015, pp. 371–387
[33] Göloglu, F., Rijmen, V., Wang, Q.: 'On the division property of S-boxes', IACR Cryptology ePrint Archive, 2016, **2016**, p. 188
[34] Boura, C., Canteaut, A.: 'Another view of the division property'. Annual Cryptology Conf., Berlin, Heidelberg, 2016, pp. 654–682
[35] Dobraunig, C., Eichlseder, M., Mendel, F., *et al.*: 'Ascon v1. 2', Submission to the CAESAR Competition, https://competitions.cr.yp.to/round3/asconv12.pdf, 2016
[36] Daemen, J., Rijmen, V.: '*The design of Rijndael: AES-the advanced encryption standard*' (Springer Science and Business Media, Berlin, Heidelberg, 2013)
[37] Bogdanov, A., Knudsen, L.R., Leander, G., *et al.*: 'PRESENT: An ultra-lightweight block cipher'. Cryptographic Hardware and Embedded Systems, Vienna, Austria, 2007, vol. 4727, pp. 450–466
[38] Wang, Q., Grassi, L., Rechberger, C.: 'Zero-sum partitions of PHOTON permutations'. Cryptographersąf Track at the RSA Conf., Cham, 2018, pp. 279–299

9 Appendix

9.1 Appendix 1

We show the Boolean expression of χ_0 and its inverse in Tables 4 and 5, respectively.

9.2 Appendix 2

We show the Boolean expression of the S-box in ASCON permutation and its inverse in Tables 11 and 12, respectively. The higher-order differential characteristics against the inverse of ASCON permutation is summarised in Table 3. Since the propagation characteristic of the division property against the S-box in ASCON is exactly the same as that against χ in KECCAK, our specific search algorithm can get better results than Todo's generic search algorithm in terms of the data complexity for certain rounds.

Zero-sum distinguishers on ASCON permutation: Combined with higher-order differential characteristics of eight rounds forward and four rounds backward, we get zero-sum distinguishers of size 2^{130} on 12th-round ASCON permutation. We consider the internal states before the nonlinear layer in the 5th round of ASCON permutation. In order to construct 12th-round zero-sum distinguishers, we prepare a linear space which takes values from $(\mathbb{F}_2^5)^{64}$, fulfilling the division property $\mathcal{D}_{\boldsymbol{k}}^{5,64}$, where $\boldsymbol{k} = (\underbrace{5,\ldots,5}_{26},\underbrace{0,\ldots,0}_{38})$.