Designs, Codes and Cryptography (2020) 88:2289–2311
https://doi.org/10.1007/s10623-020-00782-3

The phantom of differential characteristics

Yunwen Liu[1,2] · Wenying Zhang[4] · Bing Sun[1] · Vincent Rijmen[2,3] · Guoqiang Liu[1] · Chao Li[1] · Shaojing Fu[5] · Meichun Cao[4]

Received: 30 November 2019 / Revised: 2 June 2020 / Accepted: 20 June 2020 / Published online: 24 July 2020

Abstract

For differential cryptanalysis under the single-key model, the key schedules hardly need to be exploited in constructing the characteristics, which is based on the hypothesis of stochastic equivalence. In this paper, we study a profound effect of the key schedules on the validity of the differential characteristics. Noticing the sensitivity in the probability of the characteristics to specific keys, we label the keys where a characteristic has nonzero probability by *effective keys*. We propose the concept of *singular characteristics* which are characteristics with no effective keys, and exploit an algorithm to sieve them out by studying the key schedule. We show by a differential characteristic of PRINCE whose expected differential probability is much larger than that of a random permutation, i.e., 2^{-35} vs. 2^{-64}. Yet, it is indeed singular which could be mis-used to mount a differential attack. Singular characteristics are found for 3-round AES and 3-round Midori-128 as well. Furthermore, taking the possible mismatches of the effective keys in a number of differential characteristics into consideration, we present *singular clusters* which indicates an empty intersection of the corresponding effective keys, and this is evidenced by showing two differential characteristics of the 2-round AES. We also show that characteristics are tightly linked to the key schedule, as shown in the paper, a valid characteristic in the AES-128 can be singular for the AES-192. Our results indicate a gap over the perspectives of the designers and the attackers, which warns the latter to validate the theoretically-built distinguishers. Therefore, a closer look into the characteristics is inevitable before any attack is claimed.

Keywords Differential cryptanalysis · Key schedule · Effective keys · Singular characteristic · Singular cluster · AES · PRINCE

Mathematics Subject Classification 94A60

Communicated by R. Steinfeld.

✉ Wenying Zhang
zhangwenying@sdnu.edu.cn

✉ Bing Sun
happy_come@163.com

Extended author information available on the last page of the article

1 Introduction

1.1 Block ciphers and differential cryptanalysis

Block ciphers play a fundamental role in symmetric-key cryptosystems, forming the basis of various applications such as stream ciphers, hash functions and message authenticating codes (MACs). Instead of seeking for a good permutation directly, modern block ciphers iterate a cryptographically weak function called round function many times to achieve both security and efficient implementation. Prominent examples of iterated block ciphers are the Data Encryption Standard, DES [28] and the Advanced Encryption Standard, AES [12].

Symmetric-key designs are expected to resist known cryptanalytic methods based on distinguishers that distinguish a (round-reduced) block cipher from random permutations. Introduced by Biham and Shamir in 1990, differential cryptanalysis has been successfully applied to the analysis of numerous block ciphers, see for instance [2,3,17]. In this attack scenario, the adversary tries to find differences δ and Δ, such that the input pairs carrying a difference δ has a larger probability than random to be mapped into outputs with difference Δ. In order to find such a weakness that pushes forward the information of a difference at the first round to the rth round, taking the fact that the round functions are usually cryptographically weak, differences $\delta_i \rightarrow \delta_{i+1}$ with high probability are located for the ith round of the target cipher, then concatenated into an r-round *differential characteristic* ($\delta = \delta_0 \rightarrow \delta_1 \rightarrow \cdots \rightarrow \delta_r = \Delta$), and all the differential characteristics with the same input/output differences form a *differential*. The breaking of DES may be the most prominent example that shows the power of differential cryptanalysis on block ciphers. Variants of differential cryptanalysis have been proposed afterwards [4,5,7,20,22,35], many of which explore the congregated effect of characteristics [9,24,34,39].

Having been widely used in the evaluations of block ciphers and hash functions, the details of launching a differential attack take different approaches. In a single-key setting, the difference in a block cipher comes from the plaintexts, which play the same role as the chaining values in a hash function; while the difference in a hash function is injected through the messages, which resembles the round keys in a block cipher. Generally speaking, finding collisions for a hash function is somehow analogous to searching for related-key differentials of a block cipher. Examples are the collision attacks on MD5 and SHA1 [29,36,37], as well as the rebound attacks on the SHA3 candidates [27].

A block cipher is designed with a specific key schedule which expands the master keys into several round keys. A key schedule often produces some randomness for the keys in successive rounds, meanwhile they cannot be too complicated to reduce the implementation costs. Therefore, in practical block ciphers designs, the operations of a key schedule are mostly linear. For example in the key schedule of AES-128, only 4 out of 16 bytes pass through the S-box operations. Linear key schedules are commonly adopted, especially in many lightweight ciphers.

The key schedules are hardly explored to construct a distinguisher for a (round-reduced) cipher. Indeed, cryptanalysis typically searches for the distinguishing properties that are independent of the key schedules. For example, truncated impossible differentials independent of the non-linear components and the key schedule [4,30] are constructed in most impossible differential attacks. Even if they have been used to mount attacks in the related-key settings [6], the key schedules are in most cases merely involved in the key recovery attack to reduce the guessing complexity by exploring the relations among the round keys. As a matter of fact, the key schedules are considered irrelevant to the distinguishers themselves in the single-key setting, such as the search of differential characteristics with an automated tool [32] and the differential enumerating technique developed for improving the meet-in-the-middle attack [14].

The overlooking of the key schedules leads to a bottle-neck in finding better distinguishers. As shown in a previous research [31], unless the details of the S-boxes and the key schedule are explored, the 4-round impossible differentials of the AES cannot be improved. Unfortunately though, the role of the key schedule in extending the distinguishers under the single-key model still remains an open problem.

1.2 Hypothesis of stochastic equivalence

Usually, the probability of a differential is the sum of that of the corresponding differential characteristics, hence, the estimation on the probabilities of differential characteristics is crucial to the validity of an analysis. However, predicting the probability for a given differential characteristic under a specific key schedule is known to be a difficult problem. A widely accepted solution is to assume that the probability varies negligibly for different keys, which is known as the *hypothesis of stochastic equivalence* [23]. Meanwhile, the primitive is assumed to be a Markov cipher and the round keys are independent and uniformly distributed despite the key schedule. Under these assumptions, the probability of a differential characteristic is estimated by the product of the probability in each round, which is the expected differential probability (EDP) of the characteristic, i.e., the averaged probability over all independent round keys. The assumptions enable the designers to provide security proofs against differential cryptanalysis by bounding the minimum number of active S-boxes, such as the wide trail strategy in the design of the AES [11]. Therefore, this model serves as the main methodology of evaluating block ciphers against differential cryptanalysis.

1.3 Motivations and contributions

Even though the hypothesis of stochastic equivalence and the Markov model provide reliable bounds for the designers, they encounter exceptions from the point of view of attackers [16,19,38]. For instance, there exists a discrepancy between the experiments and the theoretical estimation in a chain of modular additions for differential cryptanalysis and rotational cryptanalysis [18,26]. In the meantime, the existence of 2-round plateau characteristics in the AES indicates the mismatch of fixed-key and the expected differential probability in some SPN block ciphers [13]. Recently, studies have shown that the fixed-key probability largely improves the expected differential probability for a number of ciphers, when the key is zero or chosen from a subset of the key space, see for instance the work by Canteaut et al. [10] and Sun et al. [33]. Therefore, it is vital to test whether a differential characteristic with a non-zero EDP is a real differential characteristic in a block cipher with a specific key schedule. As a direct consequence, some differential attacks on block ciphers may be at stake. Moreover, if the characteristics can be shown to be of probability zero, an attack will probably fail since the techniques for the key recovery attack based on a differential and an impossible differential are essentially different.

Consider the following toy cipher. Let $S(\cdot)$ be the 8-bit S-box of the AES and $S_k^r(x) = S(S_k^{r-1}(x) \oplus k)$, where $S_k^1(x) = S(x \oplus k)$. For a fixed-key k, one can construct at most $C_{256}^2 \approx 2^{15}$ differential characteristics of $S_k^r(x)$ by naming the pairs of plaintexts. However, under the hypothesis of stochastic equivalence and the assumption of the Markov model, since for each input difference of the S-box, there are about 2^7 possible output differences, we can find $2^8 \times 2^7 \times 2^7 \cdots \times 2^7 = 2^{7r+8}$ differential characteristics with nonzero probability. Thus, for any characteristic, its probability of being valid is approximately $2^{8+15}/2^{7r+8} = 2^{-(7r-15)}$, which is marginal when r is large. From this point of view, it is probable that a differential characteristic with a nonzero EDP may turn out to be an impossible one. Therefore, *the results of characteristic-based differential cryptanalysis is suspicious unless one can claim the characteristic is a real one*.

This paper mainly focuses on the validity of differential characteristics while the details of the key schedule are exploited, especially the differential properties of the AES and some AES-like ciphers under this setting. The basic observation is that a characteristic of a block cipher may only be of nonzero probability for a fraction of keys which are called *effective keys* in this paper. Taking the effect into account, we get the following results:

(1) We propose the concept of singular characteristics which are characteristics with no effective keys, and exploit an algorithm by studying the key schedule to sieve them out. Note that unlike the so-called impossible characteristics, singular characteristics have no conflicts in the propagations of differences through the key-less round function. Moreover, taking the possible mismatch or contradiction of effective keys in a number of differential characteristics into consideration, we present singular clusters where the intersection of the corresponding effective keys forms an empty set.

(2) To show the fact that a differential characteristic is sensitive to the key schedule, we construct a valid differential characteristic of AES-128 while it is proved to be singular in AES-192. We also present two differential characteristics of AES-128 that form a singular cluster, which warns for possible flaws in estimating the congregating effect of differential characteristics. We find singular characteristics and singular clusters with a practical effect on the multiple differential cryptanalysis of PRINCE. The singular characteristic found in PRINCE has an expected differential probability of 2^{-35}, whereas it is invalid for all master keys. In addition, a 3-round singular characteristic is found for the lightweight block cipher Midori-128 as well.

(3) We define the effective vectors of diffusion functions, such that a sufficient condition is derived, under which the fixed-key differential probability of a 2-round characteristic equals the EDP for all keys. The result provides a new understanding on the relation between the branch number of a diffusion function and the effective keys of characteristics.

The probability of a singular characteristic is zero for all master keys, even if its expected differential probability is relatively high. Unlike the fluctuation in the fixed-key differential probabilities that leads to zero probability for a fraction of keys, singular characteristic reveals the profound influence of the key schedule on differential probability and the misleading information of expected differential probability on many characteristics in practical ciphers.

As a result, our results show a big gap between the practical security and the theoretical security with respect to the characteristic-based differential cryptanalysis, due to the different perspectives of the designer and the attacker towards the hypothesis of stochastic equivalence and the Markov model.

1.4 Organisation

The rest of this paper is organised as follows. Section 2 gives preliminaries of differential cryptanalysis and the SPN ciphers. We show our basic observation on effective keys of the characteristics in Sect. 3. The singular characteristic as well as singular cluster are proposed in Sect. 4. In Sect. 5, we give the deduction of a 3-round singular characteristic in PRINCE. Section 6 shows the properties of the singular characteristics and singular clusters in the AES. In Sect. 7, we give a 3-round singular characteristic in Midori-128. Section 8 further discusses the different views of the designer and the attacker, and provides a criteria for the EDP being equal to the fixed-key probability by studying the effective vectors of the diffusion function. Finally, we conclude in Sect. 9.

2 Preliminary

2.1 Vectors and block ciphers

Let $X = (x_0, x_1, \dots, x_{l-1})^T \in \mathbb{F}_{2^s}^l$ be a vector of length l, where l, s are positive integers. The elements over finite fields $\mathbb{F}_{2^s}$ are hexadecimal with `typewriter` font. We denote the block size of a block cipher by $n \times b$ where b is the size of the S-box. A cryptographic iterative function $f = f_{r-1} \circ \cdots \circ f_1 \circ f_0$ consists of a sequence of maps called *rounds*. If the state goes through a round function f_i with a layer of n S-boxes S, followed by a linear layer P, we call f an SPN-type cipher. To facilitate our discussion, we always assume the round keys are XORed to the output state of the diffusion layer.

Through out this paper, we will take AES-128/192 as the main examples of SPN-type ciphers. We refer to [12] for more details of these ciphers as well as the key schedules.

2.2 Differential cryptanalysis

Let $\delta \in \mathbb{F}_2^{nb}$ be the input difference and $\Delta \in \mathbb{F}_2^{nb}$ be the output difference. For a vectorial boolean function $f : \mathbb{F}_2^{nb} \rightarrow \mathbb{F}_2^{nb}$, the differential probability of $\delta \rightarrow \Delta$ is defined as

$$\Pr(\delta \rightarrow \Delta) \triangleq \frac{\#\{x \in \mathbb{F}_2^{nb} | f(x) \oplus f(x \oplus \delta) = \Delta\}}{2^{nb}}.$$

We denote a differential over a map f by $\delta \overset{f}{\rightarrow} \Delta$ or simply $\delta \rightarrow \Delta$ if f is clear from the context. If $f(x \oplus \delta) \oplus f(x) = \Delta$, we call x a *right input* of $\delta \rightarrow \Delta$ and we denote by $RI(\delta, \Delta)$ all the right inputs of $\delta \rightarrow \Delta$. Similarly, we denote by $RO(\delta, \Delta)$ all the right outputs of $\delta \rightarrow \Delta$. Obviously, one has

$$RO(\delta, \Delta) = \{y = f(x) | x \in RI(\delta, \Delta)\}.$$

For an iterative function $f = f_{r-1} \circ \cdots \circ f_1 \circ f_0$, a sequence of differences

$$\Omega : \delta_0 \overset{f_0}{\rightarrow} \delta_1 \overset{f_1}{\rightarrow} \delta_2 \rightarrow \cdots \rightarrow \delta_{r-1} \overset{f_{r-1}}{\rightarrow} \delta_r$$

is called an r-round *differential characteristic* of f. The probability of a characteristic $\Pr(\Omega)$ is derived from the probability that δ_i is the difference after i rounds given that δ_{i-1} is the difference after $i-1$ rounds.

Throughout this paper, we are only interested in whether a differential characteristic is a real one or not and we will neglect the value of a nonzero probability. For a fixed permutation, the *character* of a differential characteristic Ω is defined as follows:

$$\chi(\Omega) = \begin{cases} 1 & \Pr(\Omega) > 0, \\ 0 & \Pr(\Omega) = 0. \end{cases}$$

The character function distinguishes whether a differential characteristic is *possible* or not. Furthermore, let E be a block cipher and k be a fixed key, if we can find a pair of plaintexts such that the characteristic Ω holds, we denote $\chi_k(\Omega) = 1$. Otherwise, $\chi_k(\Omega) = 0$.

3 Basic observation

For differential cryptanalysis and its variants, evaluating the probability of the characteristics is of special importance. To simplify the problem, it is often assumed that:

- The probabilities of a differential/differential characteristic under all keys are almost equal.
- The round keys are independent and uniformly distributed in a Markov cipher; the probability of a differential characteristic is the product of the probabilities for each round.

As illustrated before, a randomly-constructed characteristic under these assumptions may well be invalid in reality. If a differential characteristic with nonzero EDP is not a real characteristic in a block cipher, any key recovery attack based on such a distinguisher will fail to reflect the security. A probably more urgent concern is that the key schedule has a fundamental influence on the validity of the characteristics, contrary to the fact that the key schedules are ignored for a long time. For example, the AES encompasses three different key schedules to support variable key size. However in practice, a differential characteristic of AES-128 is naturally considered as a characteristic of AES-192 due to the same underlying round function.

Consider an r-round characteristic $\Omega : \alpha_0 \xrightarrow{S} \beta_0 \xrightarrow{P} \cdots \xrightarrow{P} \alpha_{r-1} \xrightarrow{S} \beta_{r-1}$ of an SPN cipher where $\alpha_{i+1} = P\beta_i, i = 0, \ldots, r-2$. The XOR of the round keys does not affect the differences, but it changes the intermediate values for sure. Consequently, the actual number of right pairs following such a characteristic may largely vary with the round keys. Denote the output y of the ith S-box-layer by $RO_S(\alpha_{i-1}, \beta_{i-1})$, the input z to the $(i+1)$th S-box-layer by $RI_S(\alpha_i, \beta_i)$, and the ith round key by K_i, respectively. The characteristic holds with a nonzero probability if $z = Py \oplus K_i$, which implies that

$$K_i \in [P \cdot RO_S(\alpha_{i-1}, \beta_{i-1})] \oplus RI_S(\alpha_i, \beta_i). \tag{1}$$

The observation of Eq. (1) is based on SPN ciphers; nevertheless, it can be extended to other structures as well. Thus we have the following claim:

Claim A differential characteristic of a block cipher always corresponds to some keys, i.e., the whole key space can be divided into two subsets K_0 and $K_1 = \overline{K_0}$ such that $\chi_k(\Omega) = 1$ if and only if $k \in K_0$.

It has already been noticed by previous studies that a differential characteristic may be of zero probability under certain keys, for instance, the plateau characteristics in the AES.

Definition 1 ([13]) A characteristic Q is a plateau characteristic with height $height(Q)$ if and only if:

- For a fraction $2^{nb-(weight(Q)+height(Q))}$ of the keys, the differential probability is $DP[k](Q) = 2^{height(Q)-nb}$;
- For all other keys, $DP[k](Q) = 0$.

The weight $weight(Q)$ of a possible differential or a characteristic is minus the binary logarithm of their EDP.

Different from the plateau characteristics, our focus here is on the division of the key space based on the possibility of a characteristic, rather than its exact probability. Recall that the character function $\chi_k(\Omega)$ indicates whether for a given key k the probability of a characteristic Ω is zero. We introduce the following definition of the *effective keys*.

Definition 2 The keys for which a differential characteristic Ω holds with nonzero probability are called the effective keys of Ω. The set containing all effective keys of the characteristic Ω is denoted by K_Ω, i.e., $\chi_k(\Omega) = 1$ if and only if $k \in K_\Omega$. For differential characteristics $\Omega_0, \ldots, \Omega_{t-1}$, their effective keys are defined as the intersection of all K_{Ω_i}, i.e., $K_{\Omega_0,\ldots,\Omega_{t-1}} = \cap_{i=0}^{t-1} K_{\Omega_i}$.

Usually, due to the complex relation between plaintexts and the keys, as well as a detailed key schedule, computing the effective keys of a differential characteristic is difficult. Nevertheless, insightful results can be obtained, given that the differential is planar, i.e., the right inputs and right outputs both form affine subspaces. More specifically, we focus on the effective keys of a 2-round characteristic where the differential propagations in each of the S-boxes are planar.

Theorem 1 *Let $\mathcal{E}$ be a 2-round SPN cipher. Assume that $\Omega = (\delta_0, \delta_1, \delta_2)$ is a differential characteristic of $\mathcal{E}$, where (δ_0, δ_1) and (δ_1, δ_2) are both planar. Then, the effective keys K_Ω form a linear (affine) subspace.*

Proof Given a characteristic $\Omega = (\delta_0, \delta_1, \delta_2)$ with the differentials (δ_0, δ_1) and (δ_1, δ_2) being planar, the right inputs/outputs of each active S-box form a linear/affine subspace. In addition, it is trivial that if an S-box is not active, the right inputs form exactly the whole space.

Then, according to Eq. (1), the effective keys form the following set:

$$[P \cdot RO_S(\alpha_0, \beta_0)] \oplus RI_S(\alpha_1, \beta_1).$$

Since $RO_S(\alpha_0, \beta_0)$ and $RI_S(\alpha_1, \beta_1)$ are linear/affine subspaces, the effective keys form a linear/affine subspace. □

As shown by our experiments, diffusion layer with a tight algebraic structure, such as the MixColumns operation in the AES, leads to a rather small or even empty set of effective keys. In contrast, a 2-round characteristic of PRESENT is likely to process a large amount of effective keys, as we observed in characteristics with low Hamming weights.

In fact, we can give a representation of the effective-key set of a 2-round AES characteristic through a linear system.

Example 1 Let a 2-round characteristic of the AES be

$$\begin{pmatrix} \mathtt{1} & \mathtt{0} & \mathtt{0} & \mathtt{0} \\ \mathtt{0} & \mathtt{0} & \mathtt{0} & \mathtt{0} \\ \mathtt{0} & \mathtt{0} & \mathtt{0} & \mathtt{0} \\ \mathtt{0} & \mathtt{0} & \mathtt{0} & \mathtt{0} \end{pmatrix} \xrightarrow{S} \begin{pmatrix} \mathtt{1} & \mathtt{0} & \mathtt{0} & \mathtt{0} \\ \mathtt{0} & \mathtt{0} & \mathtt{0} & \mathtt{0} \\ \mathtt{0} & \mathtt{0} & \mathtt{0} & \mathtt{0} \\ \mathtt{0} & \mathtt{0} & \mathtt{0} & \mathtt{0} \end{pmatrix} \xrightarrow{P} \begin{pmatrix} \mathtt{2} & \mathtt{0} & \mathtt{0} & \mathtt{0} \\ \mathtt{1} & \mathtt{0} & \mathtt{0} & \mathtt{0} \\ \mathtt{1} & \mathtt{0} & \mathtt{0} & \mathtt{0} \\ \mathtt{3} & \mathtt{0} & \mathtt{0} & \mathtt{0} \end{pmatrix} \xrightarrow{S} \begin{pmatrix} \mathtt{3} & \mathtt{0} & \mathtt{0} & \mathtt{0} \\ \mathtt{1} & \mathtt{0} & \mathtt{0} & \mathtt{0} \\ \mathtt{1} & \mathtt{0} & \mathtt{0} & \mathtt{0} \\ \mathtt{2} & \mathtt{0} & \mathtt{0} & \mathtt{0} \end{pmatrix}.$$

The right pairs of the differences propagating through the active S-boxes are $RI_S(\mathtt{1}, \mathtt{1}) = \{\mathtt{ce}, \mathtt{cf}\}$, $RO_S(\mathtt{1}, \mathtt{1}) = \{\mathtt{8a}, \mathtt{8b}\}$, $RI_S(\mathtt{2}, \mathtt{3}) = \{\mathtt{10}, \mathtt{12}\}$, $RI_S(\mathtt{3}, \mathtt{2}) = \{\mathtt{18}, \mathtt{1b}\}$. Hence, one gets that the first column of the effective keys falls into a linear subspace as follows:

$$\left\{ \mathcal{A} \cdot x \oplus \begin{pmatrix} \mathtt{m}_2 & \mathtt{m}_3 & \mathtt{m}_1 & \mathtt{m}_1 \\ \mathtt{m}_1 & \mathtt{m}_2 & \mathtt{m}_3 & \mathtt{m}_1 \\ \mathtt{m}_1 & \mathtt{m}_1 & \mathtt{m}_2 & \mathtt{m}_3 \\ \mathtt{m}_3 & \mathtt{m}_1 & \mathtt{m}_1 & \mathtt{m}_2 \end{pmatrix} \begin{pmatrix} \mathtt{8a} \\ \mathtt{0} \\ \mathtt{0} \\ \mathtt{0} \end{pmatrix} \oplus \begin{pmatrix} \mathtt{10} \\ \mathtt{ce} \\ \mathtt{ce} \\ \mathtt{18} \end{pmatrix}, x \in \mathbb{F}_2^{29} \right\},$$

where $\mathcal{A} = \begin{pmatrix} 2 & \mathsf{m}_3 & \mathsf{m}_1 & \mathsf{m}_1 & 2 & 0 & 0 & 0 \\ 1 & \mathsf{m}_2 & \mathsf{m}_3 & \mathsf{m}_1 & 0 & 1 & 0 & 0 \\ 1 & \mathsf{m}_1 & \mathsf{m}_2 & \mathsf{m}_3 & 0 & 0 & 1 & 0 \\ 3 & \mathsf{m}_1 & \mathsf{m}_1 & \mathsf{m}_2 & 0 & 0 & 0 & 3 \end{pmatrix}$.

The hexadecimal numbers in the above linear system represent column vectors of 8 bits. The MixColumns matrix of the AES is denoted as

$$\begin{pmatrix} \mathsf{m}_2 & \mathsf{m}_3 & \mathsf{m}_1 & \mathsf{m}_1 \\ \mathsf{m}_1 & \mathsf{m}_2 & \mathsf{m}_3 & \mathsf{m}_1 \\ \mathsf{m}_1 & \mathsf{m}_1 & \mathsf{m}_2 & \mathsf{m}_3 \\ \mathsf{m}_3 & \mathsf{m}_1 & \mathsf{m}_1 & \mathsf{m}_2 \end{pmatrix}$$

with each $\mathsf{m}_i \in \mathbb{F}_2^{8\times 8}$ being the matrix representation of the multiplication with i over $\mathbb{F}_2^8$. To be specific, with the primitive representation [30] of a diffusion function, m_1 is the identity matrix and

$$\mathsf{m}_2 = \begin{pmatrix} 0&1&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 1&0&0&0&1&0&0&0 \\ 1&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&1&0 \\ 1&0&0&0&0&0&0&1 \\ 1&0&0&0&0&0&0&0 \end{pmatrix}, \mathsf{m}_3 = \mathsf{m}_1 \oplus \mathsf{m}_2 = \begin{pmatrix} 1&1&0&0&0&0&0&0 \\ 0&1&1&0&0&0&0&0 \\ 0&0&1&1&0&0&0&0 \\ 1&0&0&1&1&0&0&0 \\ 1&0&0&0&1&1&0&0 \\ 0&0&0&0&0&1&1&0 \\ 1&0&0&0&0&0&1&1 \\ 1&0&0&0&0&0&0&1 \end{pmatrix}.$$

Meanwhile, the effective keys of the remaining columns have no constraints hence simply be in $\mathbb{F}_2^{32}$.

4 Singular characteristic and singular cluster

4.1 Singular characteristic

Though the round keys are often assumed to be independently random for the sake of simplicity, all of them depend on the master key by the key schedule with no doubt. For a key schedule with an n-bit master key generating r round keys of n bits each, the proportion of genuine keys out of all the independently random round keys is $2^{(1-r)n}$. Typically, the size of a round key is 128 bits or 64 bits, which means the fraction of genuine keys produced by the key schedule is marginal within all independently random round keys.

Recall that the set of effective keys of a differential characteristic is often a subset of all the possible keys. In this section, our focus is a seemingly peculiar but general phenomenon of characteristics, namely, the sets containing the effective keys for 2-round fragments of a characteristic are unfortunately inconsistent to each other with respect to the key schedule. We will demonstrate the existence of such "phantom" characteristics which have nonzero EDP while their probability is zero for all master keys.

Definition 3 Let Ω be an r-round characteristic of a cipher E. If $K_\Omega = \emptyset$, we call Ω a singular characteristic.

The difference between singular characteristics and the so-called impossible characteristics is that impossible characteristics are of probability zero due to the mismatch of difference propagations in the S-boxes or the linear layers, while a singular characteristic appears to be a valid characteristic if the information of the key schedule is not taken into consideration. Especially, techniques to enumerate differential characteristics with an automatic search may generate a large number of characteristic which are in fact singular. Usually, detecting a singular characteristic is not trivial. However, for those ciphers which have well-structured diffusion layers, we are able to mathematically describe the set of the effective keys. Furthermore, taking the key schedule into consideration, we execute the following strategy.

Exploring the key schedule. Suppose that the key schedule updates the $(i+1)$th round keys with the ith one by $k_{i+1} = F(k_i)$, where F is the key-expansion function. Consider a 3-round differential characteristic as follows:

$$\Omega : \alpha_0 \xrightarrow{S} \beta_0 \xrightarrow{P} P\beta_0 = \alpha_1 \xrightarrow{S} \beta_1 \xrightarrow{P} P\beta_1 = \alpha_2 \xrightarrow{S} \beta_2 \xrightarrow{P} P\beta_2 = \alpha_3.$$

According to Theorem 1, the effective keys of

$$\alpha_0 \xrightarrow{S} \beta_0 \xrightarrow{P} \alpha_1 \xrightarrow{S} \beta_1$$

can be written as a linear subspace, denote by $\mathcal{K}_1 = \{A_1 x_1 \oplus B_1\}$, where A_1 (resp. B_1) is a matrix (resp. vector) determined by the characteristic, x_1 is a vector of free variables. Similarly, denote the effective keys of

$$\alpha_1 \xrightarrow{S} \beta_1 \xrightarrow{P} \alpha_2 \xrightarrow{S} \beta_2$$

by $\mathcal{K}_2 = \{A_2 x_2 \oplus B_2\}$. According to the key schedule, we have that the keys satisfying $k_2 = F(k_1)$, for $k_1 \in \mathcal{K}_1, k_2 \in \mathcal{K}_2$, are the candidates of the effective keys for the 3-round characteristic.

For a general function F, we need to find the intersection of the sets $\mathcal{K}_2$ and $F(\mathcal{K}_1)$, in order to identify the effective keys. Notice that the key schedules in block ciphers tend to be light and simple comparing with the round functions, the problem can be much simplified if we place certain conditions on the function F. Assume that the key schedule only involves a few nonlinear operations, i.e., the round keys satisfy $k_{i+1} = L \circ N(k_i)$, where N is a nonlinear function that only applies to a small fraction of the bits in k_i and the L function is linear. Then, we have $k_2 = L \circ N(k_1)$, or equivalently, $L^{-1}(k_2) = N(k_1)$. Since N only applies to a small fraction of the bits of k_1, $L^{-1}(k_2) = N(k_1)$ involves only a small number of non-linear equations. By deleting these non-linear equations from this system, we get a linear system $\mathbb{L}(k_1, k_2) = 0$ which could be reduced to $\mathbb{L}'(x_1, x_2) = 0$. If $\mathbb{L}'(x_1, x_2) = 0$ has no solutions, we can claim that the set of the effective keys is empty which implies that Ω is singular.

Note that the strategy for computing the effective keys of a 3-round differential characteristic can be extended to any rounds, as shown in Algorithm 1. As only the linear relations of the key schedule are utilised in the linear equation systems, the effective keys of a characteristic form a subset of the solutions of the equation system. As a result, our strategy may overlook the singularity of some characteristics. If there are only a few effective keys found by the equation system, it is possible to filter out the genuine keys for such characteristics with the key schedule.

4.2 Singular clusters

When a differential contains only singular characteristics, clearly it has no effective keys. As a more general setting, consider two or more characteristics simultaneously, it is possible

Algorithm 1 The algorithm to detect if a characteristic is singular

Input: An r-round characteristic $\Omega : \alpha_0 \xrightarrow{S} \beta_0 \xrightarrow{P} P\beta_0 = \alpha_1 \xrightarrow{S} \beta_1 \xrightarrow{P} P\beta_1 \ldots \alpha_{r-1} \xrightarrow{S} \beta_{r-1} \xrightarrow{P} P\beta_{r-1} = \alpha_r$
Output: The singularity of Ω

```
Find effective keys for every 2-consecutive-round of $\Omega$ as $\mathcal{K}_i = \{A_i x_i \oplus B_i\}, 1 \le i \le r-1$.
Build an equation system based on the key schedule: $k_{i+1} = L \circ N(k_i), 1 \le i \le r-2, k_i \in \mathcal{K}_i$.
Delete the nonlinear equations and get a linear equation system $\mathbb{L}(k_i, k_{i+1}) = 0, 1 \le i \le r-2$.
Reduce the linear equation system into $\mathbb{L}'(x_i, x_{i+1}) = 0, 1 \le i \le r-2$.
if Rank of coefficient matrix $\neq$ Rank of augment matrix then
    return The characteristic $\Omega$ is singular.
else
    return The singularity of $\Omega$ is undetermined.
```

Algorithm 2 The algorithm to detect if a pair of characteristics form a singular cluster

Input: r-round characteristics $\Omega : \alpha_0 \xrightarrow{S} \beta_0 \xrightarrow{P} P\beta_0 = \alpha_1 \xrightarrow{S} \beta_1 \xrightarrow{P} P\beta_1 \ldots \alpha_{r-1} \xrightarrow{S} \beta_{r-1} \xrightarrow{P} P\beta_{r-1} = \alpha_r$ and $\Omega^* : \alpha_0^* \xrightarrow{S} \beta_0^* \xrightarrow{P} P\beta_0^* = \alpha_1^* \xrightarrow{S} \beta_1^* \xrightarrow{P} P\beta_1^* \ldots \alpha_{r-1}^* \xrightarrow{S} \beta_{r-1}^* \xrightarrow{P} P\beta_{r-1}^* = \alpha_r^*$
Output: The singularity of the collection $\{\Omega, \Omega^*\}$

```
Find effective keys for every 2-consecutive-round of $\Omega$ and $\Omega^*$ as $\mathcal{K}_i = \{A_i x_i \oplus B_i\}$, and $\mathcal{K}_i^* = \{A_i^* x_i^* \oplus B_i^*\}$, respectively, for $1 \le i \le r-1$
Build a linear equation system with $k_i = k_i^*, 1 \le i \le r-2, k_i \in \mathcal{K}_i, k_i^* \in \mathcal{K}_i^*$.
if Rank of coefficient matrix $\neq$ Rank of augment matrix then
    return The collection is a singular cluster.
else
    return The collection is undetermined.
```

that the intersection of their effective keys turns out to be an empty set. In such case, the collection of these characteristics will have no effective key.

Definition 4 Let $\mathcal{D} = \{\Omega_0, \ldots, \Omega_{t-1}\}$ be a set of differential characteristics of a block cipher E. Then, $\mathcal{D}$ is called a singular cluster of E if the corresponding effective keys form an empty set:

$$\cap_{i=0}^{t-1} K_{\Omega_i} = \emptyset.$$

At a first glance, one can determine the effective keys of each differential characteristic separately, and then find the intersection of these effective keys. However, we have a more efficient algorithm for two characteristics:

For 2-round differential characteristics Ω and Ω^*, we write the effective keys in these characteristics as $\mathcal{K} = Ax \oplus B$ and $\mathcal{K}^* = A^*x^* \oplus B^*$, respectively. The effective keys $k \in \mathcal{K}, k^* \in \mathcal{K}^*$ satisfy $k = k^*$, which is linear with respect to x and x^*. Therefore, if this linear system has no solutions, we can declare that Ω and Ω^* form a singular cluster. To determine the singularity of the equation systems is equivalent to compare the ranks of the coefficient matrix $[A, A^*]$ and the augmented matrix $[A, A^*, B \oplus B^*]$. The strategy can be generalised for any number of r-round characteristics, for simplicity, here we show it for a pair of r-round characteristics in Algorithm 2. Notice that we only compare the effective keys of the characteristics in the same segment, however, it is also possible to find singular clusters by taking the key schedule into consideration, i.e., exploring the effective keys in different segments of two characteristics.

Remark 1 Here we note a difference between singular characteristics and singular clusters. A characteristic being singular implies that its probability is zero for all master keys. However, the probability of characteristics in a singular cluster is not necessarily zero, since the singular behaviour is that some of the characteristics in the cluster cannot exist simultaneously for every master key. Moreover, for a collection of characteristics, if there are two characteristics forming a singular cluster on their own, the collection is again a singular cluster. To estimate the congregated effect of the characteristics, it is necessary to execute the above process for all pairs of characteristics in the collection.

5 Singular characteristics of PRINCE

PRINCE is a lightweight block cipher design by Borghoff et al. [8]. The S-box of PRINCE is 4-bit, and the linear layer is a composition $SR \circ M'$, where SR is a shift-row-like rotation operation and M' is a 64×64 block diagonal matrix with $(\hat{M}^{(0)}, \hat{M}^{(1)}, \hat{M}^{(1)}, \hat{M}^{(0)})$ as diagonal blocks. The matrices $\hat{M}^{(0)}$ and $\hat{M}^{(1)}$ are defined as follows.

$$\begin{pmatrix} M_0 & M_1 & M_2 & M_3 \\ M_1 & M_2 & M_3 & M_0 \\ M_2 & M_3 & M_0 & M_1 \\ M_3 & M_0 & M_1 & M_2 \end{pmatrix}, \begin{pmatrix} M_1 & M_2 & M_3 & M_0 \\ M_2 & M_3 & M_0 & M_1 \\ M_3 & M_0 & M_1 & M_2 \\ M_0 & M_1 & M_2 & M_3 \end{pmatrix}.$$

The binary matrices M_0, M_1, M_2, M_3 in PRINCE are as follows.

$$\begin{pmatrix} 0\,0\,0\,0 \\ 0\,1\,0\,0 \\ 0\,0\,1\,0 \\ 0\,0\,0\,1 \end{pmatrix}, \begin{pmatrix} 1\,0\,0\,0 \\ 0\,0\,0\,0 \\ 0\,0\,1\,0 \\ 0\,0\,0\,1 \end{pmatrix}, \begin{pmatrix} 1\,0\,0\,0 \\ 0\,1\,0\,0 \\ 0\,0\,0\,0 \\ 0\,0\,0\,1 \end{pmatrix}, \begin{pmatrix} 1\,0\,0\,0 \\ 0\,1\,0\,0 \\ 0\,0\,1\,0 \\ 0\,0\,0\,0 \end{pmatrix}.$$

A notable feature of PRINCE is the identical round keys used in the PRINCE$_{core}$ function. In such a case, it could result in a great inaccuracy by assuming independent round keys when analysing the probability of the differential characteristics.

In this section, we study differential characteristics of PRINCE with the following active patterns:

$$\left\{ \begin{pmatrix} * & 0 & * & 0 \\ 0 & 0 & 0 & 0 \\ * & 0 & * & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ * & 0 & * & 0 \\ 0 & 0 & 0 & 0 \\ * & 0 & * & 0 \end{pmatrix}, \begin{pmatrix} 0 & * & 0 & * \\ 0 & 0 & 0 & 0 \\ 0 & * & 0 & * \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & * & 0 & * \\ 0 & 0 & 0 & 0 \\ 0 & * & 0 & * \end{pmatrix} \right\},$$

where the nonzero difference is chosen from $\{\texttt{1}, \texttt{4}, \texttt{5}\}$ or $\{\texttt{2}, \texttt{8}, \texttt{a}\}$. It has been shown by Canteaut et al. that such active patterns of the characteristics can be preserved through the round function [9].

5.1 3-Round PRINCE

Given a specific characteristic, a corresponding linear equation system can be derived and checked with Algorithm 1 in Sect. 4 to verify the consistency of the keys throughout a number of rounds. For instance, the following 3-round characteristic of PRINCE can be proved singular even though its EDP is as high as 2^{-35}.

$$\Omega_1 : \begin{pmatrix} 8&0&4&0\\ 0&0&0&0\\ 4&0&8&0\\ 0&0&0&0 \end{pmatrix} \xrightarrow{S} \begin{pmatrix} 8&0&4&0\\ 0&0&0&0\\ 8&0&4&0\\ 0&0&0&0 \end{pmatrix} \xrightarrow{M'} \begin{pmatrix} 8&0&4&0\\ 0&0&0&0\\ 8&0&4&0\\ 0&0&0&0 \end{pmatrix} \xrightarrow{SR} \begin{pmatrix} 8&0&4&0\\ 0&0&0&0\\ 4&0&8&0\\ 0&0&0&0 \end{pmatrix} \xrightarrow{S} \begin{pmatrix} 8&0&5&0\\ 0&0&0&0\\ 8&0&5&0\\ 0&0&0&0 \end{pmatrix}$$

$$\xrightarrow{M'} \begin{pmatrix} 8&0&5&0\\ 0&0&0&0\\ 8&0&5&0\\ 0&0&0&0 \end{pmatrix} \xrightarrow{SR} \begin{pmatrix} 8&0&5&0\\ 0&0&0&0\\ 5&0&8&0\\ 0&0&0&0 \end{pmatrix} \xrightarrow{S} \begin{pmatrix} 2&0&5&0\\ 0&0&0&0\\ 2&0&5&0\\ 0&0&0&0 \end{pmatrix}.$$

The singularity can also be illustrated by a contradiction from the constraints on the key bits as follows. Assume that the input bit of the rth linear layer is $x_{r,i}$, $r = 0, 1, i = 0, 1, \ldots, 63$ and the output bit after the rth linear layer (which is also the input to the $r + 1$th S-box layer) is $y_{r,i}$, $r = 0, 1, i = 0, 1, \ldots, 63$. We have the following conditions on the corresponding key bits with respect to the locations of the active S-boxes when the round keys are identical,

$$\begin{cases} x_{r,4} \oplus x_{r,8} \oplus x_{r,12} \oplus c_{r,0} \oplus k_0 = y_{r,0} \\ x_{r,1} \oplus x_{r,9} \oplus x_{r,13} \oplus c_{r,1} \oplus k_1 = y_{r,1} \\ x_{r,2} \oplus x_{r,6} \oplus x_{r,14} \oplus c_{r,2} \oplus k_2 = y_{r,2} \\ x_{r,3} \oplus x_{r,7} \oplus x_{r,11} \oplus c_{r,3} \oplus k_3 = y_{r,3} \\ x_{r,32} \oplus x_{r,40} \oplus x_{r,44} \oplus c_{r,8} \oplus k_8 = y_{r,8} \\ x_{r,33} \oplus x_{r,37} \oplus x_{r,45} \oplus c_{r,9} \oplus k_9 = y_{r,9} \\ x_{r,34} \oplus x_{r,38} \oplus x_{r,42} \oplus c_{r,10} \oplus k_{10} = y_{r,10} \\ x_{r,39} \oplus x_{r,43} \oplus x_{r,47} \oplus c_{r,11} \oplus k_{11} = y_{r,11} \\ x_{r,32} \oplus x_{r,36} \oplus x_{r,40} \oplus c_{r,32} \oplus k_{32} = y_{r,32} \\ x_{r,37} \oplus x_{r,41} \oplus x_{r,45} \oplus c_{r,33} \oplus k_{33} = y_{r,33} \\ x_{r,34} \oplus x_{r,42} \oplus x_{r,46} \oplus c_{r,34} \oplus k_{34} = y_{r,34} \\ x_{r,35} \oplus x_{r,39} \oplus x_{r,47} \oplus c_{r,35} \oplus k_{35} = y_{r,35} \\ x_{r,0} \oplus x_{r,4} \oplus x_{r,12} \oplus c_{r,40} \oplus k_{40} = y_{r,40} \\ x_{r,1} \oplus x_{r,5} \oplus x_{r,9} \oplus c_{r,41} \oplus k_{41} = y_{r,41} \\ x_{r,6} \oplus x_{r,10} \oplus x_{r,14} \oplus c_{r,42} \oplus k_{42} = y_{r,42} \\ x_{r,3} \oplus x_{r,11} \oplus x_{r,15} \oplus c_{r,43} \oplus k_{43} = y_{r,43} \end{cases}, \tag{2}$$

where $c_{r,i}$ is the ith bit of the round constant in round r, $r = 0, 1$.

Particularly, with a linear combination of two equations, we get

$$x_{r,2} \oplus y_{r,2} \oplus x_{r,10} \oplus y_{r,42} \oplus c_{r,2} \oplus c_{r,42} = k_2 \oplus k_{42}.$$

Consider the difference propagation in the first S-box layer, $RO(8, 8) = \{1111, 0111\}$, and $RO(4, 8) = \{0110, 1110\}$, therefore we have $x_{0,2} = 1, x_{0,10} = 1$. Similarly in the second S-box layer, the right inputs of the 0th and 2nd S-boxes are $RI(8, 8) = \{0001, 1001\}$ and $RI(8, 5) = \{0111, 1111\}$, we have $y_{0,2} = 0, y_{0,42} = 1$. Given the round constant being 13198a2e03707344, we have the constant bits $c_{0,2} = 0, c_{0,42} = 1$. The key bits k_2 and k_{42} satisfy the following constraint.

$$k_2 \oplus k_{42} = c_{0,2} \oplus c_{0,42} \oplus 1 = 0.$$

However, an analogous deduction on the second round key shows that

$$k_2 \oplus k_{42} = 1.$$

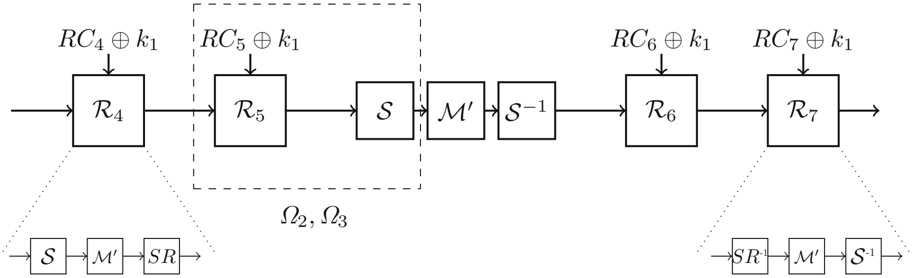

Fig. 1 PRINCE$_{core}$ reduced to 6 rounds. The dashed box shows the location of Ω_2 and Ω_3 in the 6-round characteristics adopted for multiple differential cryptanalysis

The contradiction confirms the singularity of the characteristic we obtained through the execution of Algorithm 1.

5.2 6-Round PRINCE

Notice that the multiple differential cryptanalysis applied to PRINCE [9] consist of characteristics with active patterns as Ω_1. Figure 1 illustrates the structure of the PRINCE$_{core}$ reduced to 6 rounds. Our idea is to concatenate two characteristics with no intersection in their effective key set at the two ends of the middle two rounds $S^{-1} \circ M' \circ S$.

We found the following singular cluster in 2-round PRINCE where the characteristics Ω_2 and Ω_3 have no common effective keys.

$$\Omega_2 : \begin{pmatrix} 8&0&4&0\\ 0&0&0&0\\ 4&0&8&0\\ 0&0&0&0 \end{pmatrix} \xrightarrow{S} \begin{pmatrix} 5&0&8&0\\ 0&0&0&0\\ 5&0&8&0\\ 0&0&0&0 \end{pmatrix} \xrightarrow{M'} \begin{pmatrix} 0&0&0&0\\ 5&0&8&0\\ 0&0&0&0\\ 5&0&8&0 \end{pmatrix} \xrightarrow{SR} \begin{pmatrix} 0&0&0&0\\ 0&8&0&5\\ 0&0&0&0\\ 0&5&0&8 \end{pmatrix} \xrightarrow{S} \begin{pmatrix} 0&0&0&0\\ 0&5&0&2\\ 0&0&0&0\\ 0&5&0&2 \end{pmatrix},$$

$$\Omega_3 : \begin{pmatrix} 8&0&1&0\\ 0&0&0&0\\ 1&0&8&0\\ 0&0&0&0 \end{pmatrix} \xrightarrow{S} \begin{pmatrix} 4&0&8&0\\ 0&0&0&0\\ 4&0&8&0\\ 0&0&0&0 \end{pmatrix} \xrightarrow{M'} \begin{pmatrix} 0&0&0&0\\ 4&0&8&0\\ 0&0&0&0\\ 4&0&8&0 \end{pmatrix} \xrightarrow{SR} \begin{pmatrix} 0&0&0&0\\ 0&8&0&4\\ 0&0&0&0\\ 0&4&0&8 \end{pmatrix} \xrightarrow{S} \begin{pmatrix} 0&0&0&0\\ 0&5&0&2\\ 0&0&0&0\\ 0&5&0&2 \end{pmatrix}.$$

The characteristics are located in the dashed box as depicted in Fig. 1.

The characteristics forming a singular cluster indicate that at most one of them is valid under all possible keys, such that they show no congregating effect if multiple differentials are considered. Furthermore, since round-reduced PRINCE is symmetric with respect to the middle round, the characteristics Ω_2, Ω_3 in reversed order are still within a singular cluster located in the last few rounds.

As a matter of fact, the reflection property of the PRINCE cipher can be further utilised in order to construct singular characteristics. For instance, connecting Ω_2 and Ω_3 in their tails by the middle switch $\mathcal{M}'$, we get a 4-round characteristic Ω_4 as below. Taking the constants into account, it can be verified that Ω_4 is indeed a singular characteristic covering the middle 4 rounds as in Fig. 1, which was considered as valid to build the distinguisher in the previous study [9].

$$\Omega_4 : \begin{pmatrix} 8&0&4&0\\ 0&0&0&0\\ 4&0&8&0\\ 0&0&0&0 \end{pmatrix} \xrightarrow{S} \begin{pmatrix} 5&0&8&0\\ 0&0&0&0\\ 5&0&8&0\\ 0&0&0&0 \end{pmatrix} \xrightarrow{M'} \begin{pmatrix} 0&0&0&0\\ 5&0&8&0\\ 0&0&0&0\\ 5&0&8&0 \end{pmatrix} \xrightarrow{SR} \begin{pmatrix} 0&0&0&0\\ 0&8&0&5\\ 0&0&0&0\\ 0&5&0&8 \end{pmatrix} \xrightarrow{S} \begin{pmatrix} 0&0&0&0\\ 0&5&0&2\\ 0&0&0&0\\ 0&5&0&2 \end{pmatrix} \xrightarrow{M'} \begin{pmatrix} 0&0&0&0\\ 0&5&0&2\\ 0&0&0&0\\ 0&5&0&2 \end{pmatrix}$$

$$\xrightarrow{S^{-1}} \begin{pmatrix} 0&0&0&0\\ 0&8&0&4\\ 0&0&0&0\\ 0&4&0&8 \end{pmatrix} \xrightarrow{SR^{-1}} \begin{pmatrix} 0&0&0&0\\ 4&0&8&0\\ 0&0&0&0\\ 4&0&8&0 \end{pmatrix} \xrightarrow{M'} \begin{pmatrix} 4&0&8&0\\ 0&0&0&0\\ 4&0&8&0\\ 0&0&0&0 \end{pmatrix} \xrightarrow{S^{-1}} \begin{pmatrix} 8&0&1&0\\ 0&0&0&0\\ 1&0&8&0\\ 0&0&0&0 \end{pmatrix}.$$

Moreover, it can be verified that Ω_2 and Ω_3 are actually the fragments of two 3-round characteristics with input difference

$$\begin{pmatrix} 2&0&1&0\\ 0&0&0&0\\ 2&0&1&0\\ 0&0&0&0 \end{pmatrix}.$$

So it is possible to extend Ω_4 by one round in the beginning and one round in the end, and the 6-round characteristic remains singular.

5.3 The influence of the constants

During the deduction of the singular characteristic Ω_1, we notice that the singular characteristics are sensitive to the keys as well as to the constant. It is interesting to show that the singularity of the characteristic Ω_1 depends heavily on the round number, and it can be similarly verified that Ω_1 is still singular when it covers round 1 to 3. We can further show that Ω_4 is singular when the constants RC_i and RC_{i+1} with $i = 3, 4, 5, 6, 7$ are XORed to the 4-round characteristic.

In fact, in a key-alternating cipher with the constants XORed to the state, the constants can be equivalently regarded as part of the key schedule. As a result, different round constants lead to a change in the constraints on the effective keys in consecutive rounds, which alter the singularity of a characteristic.

6 Singular characteristics of the AES

Since its publication, the security of the AES has been studied intensively. Meanwhile, thanks to the well-studied properties of the AES, other primitives such as hash functions also adopt a similar strategy. For example, the SHA3 candidate Grøstl [15] adopts two permutations which are similar to the AES.

Clearly, the singularity of differential characteristics is sensitive to the key schedule. For two ciphers with the same round function but different key schedules, it is likely to find characteristics which are valid for one but singular for another. Here, by encrypting a pair of messages with a 128-bit random master key through the AES, and tracking the difference propagation for 3 rounds (from round 3 to 5), we get a valid 3-round characteristic of the AES-128[1]:

[1] Although characteristics with fewer active S-boxes are often preferred by attackers, it is difficult to confirm that such characteristics are not singular. So we construct the example in such a way that the characteristic is guaranteed to possess at least two right inputs. Here, we conjecture that in general a valid characteristic would probably turn into a singular one when the key schedule is modified.

$$\Omega_5 : \begin{pmatrix} \texttt{c} & \texttt{ae} & \texttt{21} & \texttt{17} \\ \texttt{8} & \texttt{57} & \texttt{21} & \texttt{39} \\ \texttt{4} & \texttt{57} & \texttt{63} & \texttt{2e} \\ \texttt{4} & \texttt{f9} & \texttt{42} & \texttt{17} \end{pmatrix} \xrightarrow{S} \begin{pmatrix} \texttt{d9} & \texttt{fd} & \texttt{94} & \texttt{7e} \\ \texttt{15} & \texttt{d8} & \texttt{51} & \texttt{f2} \\ \texttt{f3} & \texttt{ee} & \texttt{14} & \texttt{7c} \\ \texttt{ec} & \texttt{eb} & \texttt{8b} & \texttt{b7} \end{pmatrix} \xrightarrow{P} \begin{pmatrix} \texttt{79} & \texttt{82} & \texttt{26} & \texttt{a6} \\ \texttt{f9} & \texttt{37} & \texttt{8e} & \texttt{f6} \\ \texttt{eb} & \texttt{7b} & \texttt{bd} & \texttt{2a} \\ \texttt{c9} & \texttt{f2} & \texttt{6b} & \texttt{74} \end{pmatrix} \xrightarrow{S} \begin{pmatrix} \texttt{5} & \texttt{24} & \texttt{b0} & \texttt{94} \\ \texttt{7b} & \texttt{f4} & \texttt{fc} & \texttt{e8} \\ \texttt{8} & \texttt{c} & \texttt{3e} & \texttt{a3} \\ \texttt{8d} & \texttt{fe} & \texttt{9c} & \texttt{c3} \end{pmatrix}$$

$$\xrightarrow{P} \begin{pmatrix} \texttt{f0} & \texttt{79} & \texttt{ae} & \texttt{2e} \\ \texttt{77} & \texttt{b4} & \texttt{9d} & \texttt{ea} \\ \texttt{d3} & \texttt{9} & \texttt{51} & \texttt{48} \\ \texttt{58} & \texttt{32} & \texttt{cc} & \texttt{f3} \end{pmatrix} \xrightarrow{S} \begin{pmatrix} \texttt{99} & \texttt{5e} & \texttt{74} & \texttt{5} \\ \texttt{3e} & \texttt{23} & \texttt{af} & \texttt{88} \\ \texttt{5f} & \texttt{dd} & \texttt{49} & \texttt{7e} \\ \texttt{19} & \texttt{60} & \texttt{95} & \texttt{aa} \end{pmatrix}.$$

However, it can be verified that Ω_5 is actually a singular characteristic of AES-192 (from round 3 to 5), which implies a valid attack based on differential characteristics on AES-128 might be invalid on AES-192.

The existence of singular characteristics in the AES-like designs may undermine the validity of some cryptanalyses. To be specific, if a differential characteristic in use is a singular one, the table constructed in the meet-in-the-middle technique may be invalid. Thus, to show the effectiveness of meet-in-the-middle attack against round-reduced AES, one has to prove that the differential characteristic is real.

6.1 A 3-round singular characteristic for AES-128

The following 3-round differential characteristic for AES-128 is singular for the first 3 rounds of the AES-128:

$$\Omega_6 : \begin{pmatrix} \texttt{1} & \texttt{3} & \texttt{0} & \texttt{3} \\ \texttt{1} & \texttt{1} & \texttt{0} & \texttt{3} \\ \texttt{0} & \texttt{1} & \texttt{1} & \texttt{6} \\ \texttt{0} & \texttt{6} & \texttt{3} & \texttt{1} \end{pmatrix} \xrightarrow{S} \begin{pmatrix} \texttt{1} & \texttt{2} & \texttt{0} & \texttt{2} \\ \texttt{1} & \texttt{1} & \texttt{0} & \texttt{2} \\ \texttt{0} & \texttt{1} & \texttt{1} & \texttt{1} \\ \texttt{0} & \texttt{1} & \texttt{2} & \texttt{1} \end{pmatrix} \xrightarrow{P} \begin{pmatrix} \texttt{1} & \texttt{5} & \texttt{7} & \texttt{4} \\ \texttt{1} & \texttt{1} & \texttt{5} & \texttt{1} \\ \texttt{1} & \texttt{0} & \texttt{1} & \texttt{7} \\ \texttt{1} & \texttt{7} & \texttt{0} & \texttt{2} \end{pmatrix} \xrightarrow{S} \begin{pmatrix} \texttt{1} & \texttt{4} & \texttt{1} & \texttt{9} \\ \texttt{1} & \texttt{1} & \texttt{4} & \texttt{80} \\ \texttt{1} & \texttt{0} & \texttt{1} & \texttt{1} \\ \texttt{1} & \texttt{1} & \texttt{0} & \texttt{3} \end{pmatrix} \xrightarrow{P} \begin{pmatrix} \texttt{3} & \texttt{4} & \texttt{99} & \texttt{11} \\ \texttt{3} & \texttt{E} & \texttt{18} & \texttt{B} \\ \texttt{7} & \texttt{1} & \texttt{80} & \texttt{B} \\ \texttt{5} & \texttt{B} & \texttt{80} & \texttt{1A} \end{pmatrix}$$

$$\xrightarrow{S} \begin{pmatrix} \texttt{2} & \texttt{9} & \texttt{3} & \texttt{1} \\ \texttt{2} & \texttt{2} & \texttt{1} & \texttt{1} \\ \texttt{1} & \texttt{1} & \texttt{2} & \texttt{2} \\ \texttt{4} & \texttt{1} & \texttt{2} & \texttt{1} \end{pmatrix}.$$

6.2 Singular cluster in the AES

Singular clusters can also be found in the AES. As an example, we find that the following 2-round differential characteristics of the AES-128 form a singular cluster:

$$\Omega_7 : \begin{pmatrix} \texttt{1} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{0} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{0} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{0} & \texttt{0} & \texttt{0} & \texttt{0} \end{pmatrix} \xrightarrow{S} \begin{pmatrix} \texttt{18} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{0} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{0} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{0} & \texttt{0} & \texttt{0} & \texttt{0} \end{pmatrix} \xrightarrow{P} \begin{pmatrix} \texttt{30} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{18} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{18} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{28} & \texttt{0} & \texttt{0} & \texttt{0} \end{pmatrix} \xrightarrow{S} \begin{pmatrix} \texttt{3} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{1} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{1} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{2} & \texttt{0} & \texttt{0} & \texttt{0} \end{pmatrix},$$

$$\Omega_8 : \begin{pmatrix} \texttt{1} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{0} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{0} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{0} & \texttt{0} & \texttt{0} & \texttt{0} \end{pmatrix} \xrightarrow{S} \begin{pmatrix} \texttt{14} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{0} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{0} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{0} & \texttt{0} & \texttt{0} & \texttt{0} \end{pmatrix} \xrightarrow{P} \begin{pmatrix} \texttt{28} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{14} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{14} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{3c} & \texttt{0} & \texttt{0} & \texttt{0} \end{pmatrix} \xrightarrow{S} \begin{pmatrix} \texttt{3} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{1} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{1} & \texttt{0} & \texttt{0} & \texttt{0} \\ \texttt{2} & \texttt{0} & \texttt{0} & \texttt{0} \end{pmatrix}.$$

The set of effective keys $K_{\Omega_7,\Omega_8} = K_{\Omega_7} \cap K_{\Omega_8} = \emptyset$ means that they could not exist at the same time for any key; thus, they show no congregating effect when considered in a differen-

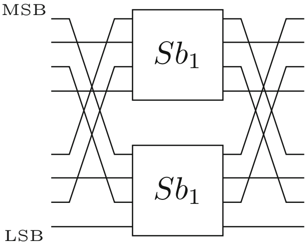

Fig. 2 The S-box SSb_0 of Midori-128

tial. Moreover, even if they are extended to cover more rounds, the resulting characteristics still form a singular cluster as long as they are located at the same fragment.

Remark 2 Singular clusters are not rare. Recall that the singularity can be identified by a linear equation system, we expect a similar behaviour as the characteristics being singular, that is, the linear equation system derived for a set of characteristics is more likely to be singular if differences in the characteristics are of higher Hamming weights. The singular characteristics and singular clusters can be detected for many lightweight designs, especially when they adopt an AES-like structure. This presents a concern for the attackers about the effectiveness of cryptanalysis based on differential characteristics.

Remark 3 The observations on the singular characteristics and singular clusters in the AES indicate that even though the provable bound derived from minimum number of active S-boxes guarantees the difficulty in finding an effective differential distinguisher for the AES-like primitives, vulnerability to differential cryptanalysis might still exist.

7 Singular characteristics of Midori-128

Midori-128 is an SPN block cipher designed by Banik et al. with optimised features for low energy [1]. The 8-bit S-box is constructed with the parallel connection of two 4-bit S-box Sb_1, sandwiched by two bit-shuffling layers to keep the involution property. For instance, the S-box SSb_0 is shown as Fig. 2,

where

$$Sb_1 = \{\texttt{1}, \texttt{0}, \texttt{5}, \texttt{3}, \texttt{e}, \texttt{2}, \texttt{f}, \texttt{7}, \texttt{d}, \texttt{a}, \texttt{9}, \texttt{b}, \texttt{c}, \texttt{8}, \texttt{4}, \texttt{6}\}.$$

The detail of the 8-bit Sbox $SSb_i, i = 1, 2, 3$ are similar to SSb_0 and can be found in the original design [1]. The MixColumns matrix is a 4×4 binary one with branch number 4:

$$M = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

Identical round keys are XORed to the state in each round. Meanwhile, the round constants only add to the least significant bit of each cell.

Lemma 1 *The MixColumns function M has fixed points with the active form* $(a, 0, a, 0)$ *and* $(0, a, 0, a)$, *for* $a \in \mathbb{F}_{2^8}$.

Lemma 2 *Let $aa = (a_7, a_6, \ldots, a_0), bb = (b_7, b_6, \ldots, b_0)$ of SSb_0 represent the concatenation of two identical 4-bit numbers, i.e. $a_0 = a_4, a_1 = a_5, a_2 = a_6, a_3 = a_7$, and $b_0 = b_4, b_1 = b_5, b_2 = b_6, b_3 = b_7$. If (aa, bb) is a possible difference propagation of the S-box SSb_i, $i = 0, 1, 2, 3$, then the set of right input is*

$$RI(aa, bb) = \{zz, ww, P(zw), P(wz)\}$$

or

$$\begin{aligned} RI(aa, bb) = \{&zz, ww, mm, nn, P(zw), P(wz), P(mn), P(nm),\\ &P(zm), P(mz), P(wm), P(mw), P(zn), P(nz), P(wn), P(nw)\}, \end{aligned}$$

where z, w, m, n are 4-bit.

Proof For simplicity, we take SSb_0 as an example, and the cases for SSb_i, $0 < i \leq 3$ can be proved similarly. Analogously, the set of right outputs $RO(aa, bb)$ can be characterised. As it can be seen, the input differences for the two 4-bit S-box Sb_1's are equal, and so are the output differences. Notice that Sb_1 is a 4-uniform S-box, suppose that $RI_{sb_1}(a, b) = \{z, w\}$, we have $RI_{SSb_0}(aa, bb) = \{zz, ww, P(zw), P(wz)\}$. Similarly, when $RI_{sb_1}(a, b) = \{z, w, m, n\}$, we have

$$\begin{aligned} RI_{SSb_0}(aa, bb) = \{&zz, ww, mm, nn, P(zw), P(wz), P(mn), P(nm),\\ &P(zm), P(mz), P(wm), P(mw), P(zn), P(nz), P(wn), P(nw)\}. \end{aligned}$$

We construct the following differential characteristic that is singular for the first 3 rounds of Midori-128:

$$\begin{aligned} \Omega_9 : &\begin{pmatrix} 22 & 22 & 0 & 0 \\ 22 & 22 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{S} \begin{pmatrix} 88 & 88 & 0 & 0 \\ 88 & 88 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{Sf} \begin{pmatrix} 88 & 0 & 0 & 0 \\ 0 & 88 & 0 & 0 \\ 88 & 0 & 0 & 0 \\ 0 & 88 & 0 & 0 \end{pmatrix} \xrightarrow{MC} \begin{pmatrix} 88 & 0 & 0 & 0 \\ 0 & 88 & 0 & 0 \\ 88 & 0 & 0 & 0 \\ 0 & 88 & 0 & 0 \end{pmatrix} \\ \xrightarrow{S} &\begin{pmatrix} 22 & 0 & 0 & 0 \\ 0 & 22 & 0 & 0 \\ 22 & 0 & 0 & 0 \\ 0 & 22 & 0 & 0 \end{pmatrix} \xrightarrow{Sf} \begin{pmatrix} 22 & 0 & 0 & 22 \\ 0 & 0 & 0 & 0 \\ 22 & 0 & 0 & 22 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{MC} \begin{pmatrix} 22 & 0 & 0 & 22 \\ 0 & 0 & 0 & 0 \\ 22 & 0 & 0 & 22 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{S} \begin{pmatrix} 33 & 0 & 0 & 88 \\ 0 & 0 & 0 & 0 \\ 88 & 0 & 0 & 88 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

□

8 Effective vectors of linear functions

An important observation in previous sections is that the differential characteristics have a much closer connection with the keys than the statement in the hypothesis of stochastic equivalence. The existence of singular characteristics and singular clusters shows that the information of the keys and the key schedules should be taken into consideration, not merely for related-key models and other open-key models as we previously believed.

Consider the following illustrative example. Assume that Ω is a valid 10-round differential characteristic with at least one pair of right inputs in the AES-128. We can find such a characteristic by encrypting a pair of plaintexts under a random key, and the solutions to the linear equation system are candidates for the effective keys. In most experiments we carried out, the only candidate from the solution of the equation system is exactly the master key used in the encryption.

What we believe critical is that the hypothesis of stochastic equivalence and Markov model need to be interpreted from two different perspectives: the designer and the attacker.

Designer's Perspective. When designing a block cipher or a permutation, the model based on the hypothesis of stochastic equivalence provides an approximate view of an overall behaviour possessed by general characteristics under all independently chosen random keys. Under such a scenario, it makes sense that the designers take the security bound as a main factor of consideration rather than a particular characteristic having probability 0 or a marginal nonzero probability.

Attacker's Perspective. For the attackers whose main target is to identify one specific non-randomness in the primitives, the hypothesis of stochastic equivalence for a designer's perspective might lead to a distorted image.[2] One of the underestimated factors is the role of the key schedule as we show for the singular characteristics, which is unfortunately often ignored in many practical analyses.

To fully characterise the effective keys mathematically is not a trivial task. However, we know from the discussions in Sect. 4 that the conflicts of effective keys in 2-round singular characteristics are already illustrative for detecting singular characteristics with more rounds.

In this section, we will focus on the property of 2-round characteristics. From a practical view point, the characteristics with effective keys being the whole space are of particular interests. Therefore, a first step is to locate those characteristics which are independent with the keys. For a two-round SPN cipher with one key addition layer, if the assumption of round independence holds, the probability of a 2-round characteristic is exactly the product of that of each round, and the probability will be independent with the round key.

8.1 Effective vectors of a diffusion function

For a vector $\alpha = (\alpha_0, \ldots, \alpha_{n-1})^T \in \mathbb{F}_{2^b}^n$, denote by $\mathrm{supp}(\alpha)$ the support of α, i.e., the set of indexes i such that $\alpha_i \neq 0$:

$$\mathrm{supp}(\alpha) = \{i | \alpha_i \neq 0\}.$$

Suppose I and O are two subsets of $\{0, 1, \ldots, n-1\}$. The sub-matrix $M_{n\times n}^{(I,O)}$ of $M_{n\times n}$ is selected by taking the rows and the columns indexed by the set I and O, respectively.

Definition 5 Let $P \in \mathbb{F}_{2^b}^{n\times n}$ be an invertible matrix. $0 \neq \beta \in \mathbb{F}_{2^b}^n$ is called an effective vector of P if

$$rank\left(P^{(\mathrm{supp}(P\beta),\overline{\mathrm{supp}(\beta)})}\right) = \#\mathrm{supp}(P\beta) = w(P\beta),$$

and we use $EV(P)$ to denote the set of all effective vectors of P.

Since $(1, 1, \ldots, 1)^T \notin EV(P)$, we always have $EV(P) \subsetneq \mathbb{F}_{2^b}^n$. The rank of

$$P^{(\mathrm{supp}(P\beta),\overline{\mathrm{supp}(\beta)})}$$

being

$$\#\mathrm{supp}(P\beta)$$

[2] Some experiments show that the estimation under the hypothesis is rather close to reality, see for instance, [21]. It is noteworthy that the irregularity is what the attackers have to pay extra attention to, which is supported by a number of studies such as those we have previously referred to in this paper.

means that

$$\#\overline{\mathrm{supp}(\beta)} \geq \#\mathrm{supp}(P\beta),$$

which is equivalent to

$$n \geq \#\mathrm{supp}(P\beta) + \#\mathrm{supp}(\beta).$$

Therefore, we have:

Lemma 3 *Let $P \in \mathbb{F}_{2^b}^{n\times n}$ be an invertible matrix. Then $EV(P) \neq \emptyset$ implies the branch number $\mathcal{B}(P) \leq n$, and $\mathcal{B}(P) = n+1$ implies $EV(P) = \emptyset$.*

8.2 Characters of 2-round characteristics

Despite the existence of singular characteristics and effective keys, we can construct 2-round characteristics such that the fixed-key differential probability equals the EDP. Firstly, from previous discussion, we have the following lemma:

Lemma 4 *Let E be 2-round SPN cipher. For a differential characteristic $\Omega : \alpha_0 \overset{S}{\to} \beta_0 \overset{P}{\to} \alpha_1 \overset{S}{\to} \beta_1$ of E, the character has the following property:*

$$\chi(\Omega) = \begin{cases} \chi(\alpha_0 \overset{S}{\to} \beta_0)\chi(\alpha_1 \overset{S}{\to} \beta_1) & K \textit{ is effective}, \\ 0 & \textit{Otherwise}. \end{cases}$$

With the information of the effective vectors, one can also compute the exact value of $\Pr(\Omega)$ for some characteristics.

Theorem 2 *Let E be a 2-round SPN cipher which uses a matrix in $\mathbb{F}_{2^b}^{n\times n}$ as the diffusion layer. For any 2-round characteristic $\Omega : \alpha_0 \overset{S}{\to} \beta_0 \overset{P}{\to} \alpha_1 \overset{S}{\to} \beta_1$ where $\beta_0 \in EV(P)$, we always have $K_\Omega = \mathbb{F}_{2^b}^{n}$ and*

$$\Pr(\Omega) = \Pr\left(\alpha_0 \overset{S}{\to} \beta_0\right)\Pr\left(\alpha_1 \overset{S}{\to} \beta_1\right).$$

Proof Firstly, we have

$$\Pr\left(\alpha_0 \overset{S}{\to} \beta_0 \overset{P}{\to} \alpha_1 \overset{S}{\to} \beta_1\right) = \Pr\left(\alpha_1 \overset{S}{\to} \beta_1 | \alpha_0 \overset{S}{\to} \beta_0\right)\Pr\left(\alpha_0 \overset{S}{\to} \beta_0\right).$$

Without loss of generality, suppose the first i elements of α_0 are non-zero and the other $n-i$ elements are 0. Then, each output of the right pair of $\alpha_0 \to \beta_0$ can be written as

$$(c_0, \ldots, c_{i-1}, x_i, \ldots, x_{n-1})^T \triangleq \begin{pmatrix} C \\ X \end{pmatrix},$$

$$\left(c_0 \oplus \beta_0^0, \ldots, c_{i-1} \oplus \beta_0^{i-1}, x_i, \ldots, x_{n-1}\right)^T \triangleq \begin{pmatrix} C \oplus \beta_0 \\ X \end{pmatrix},$$

where $(c_t, c_t \oplus \beta_0^t)$ is the right output of the corresponding active S-box.

Similarly, without loss of generality, let the first j elements of α_1 be non-zero and the other $n-j$ elements be 0. Denote by

$$P = \begin{pmatrix} A_{j\times i} & B_{j\times(n-i)} \\ M_{(n-j)\times i} & N_{(n-j)\times(n-i)} \end{pmatrix} \triangleq \begin{pmatrix} A & B \\ M & N \end{pmatrix}.$$

Then for each $C = (c_0, \ldots, c_{i-1})^T$, we have:

$$P\begin{pmatrix} C \\ X \end{pmatrix} = \begin{pmatrix} A & B \\ M & N \end{pmatrix}\begin{pmatrix} C \\ X \end{pmatrix} = \begin{pmatrix} AC \oplus BX \\ MC \oplus NX \end{pmatrix}.$$

Since the last $n - j$ elements of α_1 are 0, $MC \oplus NX$ should be a constant for all X. Otherwise, there may be some non-zero elements in the last $n - j$ elements. On the other hand, rank$B = j$ implies that $\{AC \oplus BX\}$ is equal to the vector space $\mathbb{F}_{2^b}^j$ which is independent of the specific value of C. Therefore, we have

$$\Pr\left(\alpha_1 \overset{S}{\to} \beta_1 | \alpha_0 \overset{S}{\to} \beta_0\right) = \Pr\left(\alpha_1 \overset{S}{\to} \beta_1\right).$$

Thus, $\beta_0 \in EV(P)$ indicates $\Pr(\Omega_E) = \Pr(\alpha_0 \overset{S}{\to} \beta_0) \Pr(\alpha_1 \overset{S}{\to} \beta_1)$. □

The above theorem shows that while $\beta_0 \in EV(P)$, one can use the expected differential probability to compute the fixed-key probability of a 2-round characteristic. Notice that $EV(P) = \emptyset$ for an MDS matrix P, the fixed-key differential probability of a 2-round characteristic in the AES or any AES-like ciphers with an MDS matrix is unlikely to equal to the EDP, which matches the observation in the plateau characteristics.

8.3 An insightful example

Example 2 We define a toy cipher as follows. The 20-bit input is divided into five 4-bit words:

$$x = (x_0, x_1, x_2, x_3, x_4)^T \in \mathbb{F}_2^{4\times 5}.$$

In each round, the state x first passes through 5 parallel 4×4 S-boxes:

$$y = S(x) = (S_0(x_0), S_1(x_1), S_2(x_2), S_3(x_3), S_4(x_4))^T,$$

then a linear layer is applied to y:

$$\begin{pmatrix} z_0 \\ z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}.$$

For $\beta = (a, 0, a, 0, 0)^T$, we have $P\beta = \beta$. Therefore the sub-matrix is

$$P^{(\{0,2\},\{1,3,4\})} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Thus $\beta \notin EV(P)$, which means the probability of the differential characteristic is not necessarily the product of two rounds.

We check the probability of the differential characteristic by experiments. For all 2^{16} combinations of input differences and output differences with the form $(*, 0, *, 0, 0)$, about 2^{10} show inaccuracy between the theoretical probability EDP and the experimental fixed-key probability DP_k, about 2^9 of them are impossible characteristics in reality but mistaken as possible by theoretical estimation.

For real ciphers, an experiment verification taking a sampling space of the whole key space reflects the reality only when the effective key set of a characteristic is (close to) the whole space. When constructing differential characteristics, we often require the characteristic involving as less S-boxes as possible. Since for the differential 0 → 0 of an S-box, the set of right inputs is exactly $\mathbb{F}_{2^b}$, it is possible that K_Ω is the whole space or approximately the whole space if Ω has a low hamming weight. Our observations on the effective vectors of the permutation layers implies that the singularity of characteristics differs from ciphers. Therefore, it requires more inspections on the properties of the ciphers to find singular characteristic, as shown in previous sections.

9 Concluding remarks

The key schedule of a block cipher is rarely exploited in differential cryptanalysis. The hypothesis of stochastic equivalence and the Markov model are the foundations to evaluate the security of block ciphers against differential cryptanalysis and serve as a guideline for designing cryptographic primitives. In this paper, we show that, applying characteristic-based differential cryptanalysis to a real cipher may lead to incorrect results.

To show our viewpoint, we propose the concept of singular characteristics by studying their effective keys, i.e., characteristics with nonzero EDP but with probability 0 for all master keys. In addition, we study the congregating effect of characteristics and propose singular clusters to find multiple differential characteristics with no effective keys. We found examples of a 3-round singular characteristic in the AES and Midori-128, by investigating the property of its key schedule. We also construct a valid differential characteristic of AES-128 while it is proved to be singular in AES-192 to demonstrate that the existence of a differential characteristic is sensitive to the key schedule. In addition, we give a sufficient condition for the EDP to be equal to fixed-key differential probability for all keys, with the study on effective vector of linear functions.

It is also interesting to note that the opposite of singular characteristics may also be useful for attackers. If a characteristic has a high probability for most of the genuine keys, it is in fact a good distinguisher even though the EDP of the characteristic might be marginal. A similar phenomenon has been observed through the analysis of PRINTcipher in multidimensional linear cryptanalysis [25, Sect. 4.3].

Acknowledgements This paper is supported by National Natural Science Foundation of China under Grants 61902414, 61672530, 61772545, 61672330, and Onderzoeksraad, KU Leuven (C16/15/058 and OT/13/071).

References

1. Banik S., Bogdanov A., Isobe T., Shibutani K., Hiwatari H., Akishita T., Regazzoni F.: Midori: a block cipher for low energy. In: Advances in Cryptology—ASIACRYPT 2015, pp. 411–436. Springer (2015).
2. Biham E., Shamir A.: Differential cryptanalysis of DES-like cryptosystems. In: Advances in Cryptology—CRYPTO '90, 10th Annual International Cryptology Conference, Santa Barbara, California, USA, 11–15 August 1990. Proceedings, pp. 2–21 (1990).
3. Biham E., Shamir A.: Differential cryptanalysis of Snefru, Khafre, REDOC-II, LOKI and Lucifer. In: Advances in Cryptology—CRYPTO '91, 11th Annual International Cryptology Conference, Santa Barbara, California, USA, 11–15 August 1991. Proceedings, pp. 156–171 (1991).
4. Biham E., Biryukov A., Shamir A.: Cryptanalysis of Skipjack reduced to 31 rounds using impossible differentials. In: Advances in Cryptology—EUROCRYPT '99, International Conference on the Theory and Application of Cryptographic Techniques, Prague, Czech Republic, 2–6 May 1999. Proceeding, pp. 12–23 (1999).

5. Biham E., Dunkelman O., Keller N.: New results on boomerang and rectangle attacks. In: Fast Software Encryption, 9th International Workshop, FSE 2002, Leuven, Belgium, 4–6 February 2002. Revised Papers, pp. 1–16 (2002).
6. Biryukov A., Khovratovich D.: Related-key cryptanalysis of the full AES-192 and AES-256. In: Advances in Cryptology—ASIACRYPT 2009, 15th International Conference on the Theory and Application of Cryptology and Information Security, Tokyo, Japan, 6–10 December 2009. Proceedings, pp. 1–18 (2009).
7. Blondeau C., Gérard B.: Multiple differential cryptanalysis: theory and practice. In: Fast Software Encryption—18th International Workshop, FSE 2011, Lyngby, Denmark, 13–16 February 2011. Revised Selected Papers, pp. 35–54 (2011).
8. Borghoff J., Canteaut A., Güneysu T., Kavun E.B., Knezevic M., Knudsen L.R., Leander G., Nikov V., Paar C., Rechberger C., Rombouts P., Thomsen S.S., Yalçin T.: PRINCE—a low-latency block cipher for pervasive computing applications—extended abstract. In: Advances in Cryptology—ASIACRYPT 2012—18th International Conference on the Theory and Application of Cryptology and Information Security, Beijing, China, 2–6 December 2012. Proceedings, pp. 208–225 (2012).
9. Canteaut A., Fuhr T., Gilbert H., Naya-Plasencia M., Reinhard J.: Multiple differential cryptanalysis of round-reduced PRINCE. In: Fast Software Encryption—21st International Workshop, FSE 2014, London, UK, 3–5 March 2014. Revised Selected Papers, pp. 591–610 (2014).
10. Canteaut A., Lambooij E., Neves S., Rasoolzadeh S., Sasaki Y., Stevens M.: Refined probability of differential characteristics including dependency between multiple rounds. IACR Trans. Symmetric Cryptol. **2017**(2), 203–227 (2017).
11. Daemen J., Rijmen V.: AES and the wide trail design strategy. In: EUROCRYPT 2002, pp. 108–109 (2002).
12. Daemen J., Rijmen V.: The Design of Rijndael: AES-The Advanced Encryption Standard. Information Security and CryptographySpringer, Berlin (2002).
13. Daemen J., Rijmen V.: Plateau characteristics. IET Inf. Secur. **1**(1), 11–17 (2007).
14. Derbez P., Fouque P., Jean J.: Improved key recovery attacks on reduced-round AES in the single-key setting. In: Advances in Cryptology—EUROCRYPT 2013, 32nd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Athens, Greece, 26–30 May 2013. Proceedings, pp. 371–387 (2013).
15. Gauravaram P., Knudsen L.R., Matusiewicz K., Mendel F., Rechberger C., Schläffer M., Thomsen S.S.: Grøstl-a SHA-3 candidate. In: Dagstuhl Seminar Proceedings. Schloss Dagstuhl-Leibniz-Zentrum für Informatik (2009).
16. Hall C., Kelsey J., Rijmen V., Schneier B., Wagner D.: Cryptanalysis of SPEED. In: Selected Areas in Cryptography ’98, SAC’98, Kingston, Ontario, Canada, 17–18 August 1998. Proceedings, pp. 319–338 (1998).
17. Karpman P., Peyrin T., Stevens M.: Practical free-start collision attacks on 76-step SHA-1. In: Advances in Cryptology—CRYPTO 2015—35th Annual Cryptology Conference, Santa Barbara, CA, USA, 16–20 August 2015. Proceedings, Part I, pp. 623–642 (2015).
18. Khovratovich D., Nikolic I., Pieprzyk J., Sokolowski P., Steinfeld R.: Rotational cryptanalysis of ARX revisited. In: Fast Software Encryption—22nd International Workshop, FSE 2015, Istanbul, Turkey, 8–11 March 2015. Revised Selected Papers, pp. 519–536 (2015).
19. Knudsen L.R.: Iterative characteristics of DES and s^2-DES. In: Advances in Cryptology—CRYPTO ’92, 12th Annual International Cryptology Conference, Santa Barbara, California, USA, 16–20 August 1992. Proceedings, pp. 497–511 (1992).
20. Knudsen L.R.: Truncated and higher order differentials. In: Fast Software Encryption: Second International Workshop, Leuven, Belgium, 14–16 December 1994. Proceedings, pp. 196–211 (1994).
21. Kölbl S., Leander G., Tiessen T.: Observations on the SIMON block cipher family. In: Advances in Cryptology—CRYPTO 2015—35th Annual Cryptology Conference, Santa Barbara, CA, USA, 16–20 August 2015. Proceedings, Part I, pp. 161–185. Springer (2015).
22. Lai X.: Higher order derivatives and differential cryptanalysis. Commun. Cryptogr. **276**, 227–233 (1994).
23. Lai X., Massey J.L., Murphy S.: Markov ciphers and differential cryptanalysis. In: Advances in Cryptology—EUROCRYPT ’91, Workshop on the Theory and Application of Cryptographic Techniques, Brighton, UK, 8–11 April 1991, Proceedings, pp. 17–38 (1991).
24. Lallemand V., Naya-Plasencia M.: Cryptanalysis of KLEIN. In: Fast Software Encryption—21st International Workshop, FSE 2014, London, UK, 3–5 March 2014. Revised Selected Papers, pp. 451–470 (2014).
25. Leander G., Abdelraheem M., AlKhzaimi H., Zenner E.: A cryptanalysis of PRINTcipher: the invariant subspace attack. In: Advances in Cryptology—CRYPTO 2011—31st Annual Cryptology Conference, Santa Barbara, CA, USA, 14–18 August 2011. Proceedings, pp. 206–221. Springer (2011).

26. Leurent G.: Analysis of differential attacks in ARX constructions. In: Advances in Cryptology—ASIACRYPT 2012—18th International Conference on the Theory and Application of Cryptology and Information Security, Beijing, China, 2–6 December 2012. Proceedings, pp. 226–243 (2012).
27. Mendel F., Rechberger C., Schläffer M., Thomsen S.S.: The rebound attack: cryptanalysis of reduced whirlpool and Grøstl. In: Fast Software Encryption, 16th International Workshop, FSE 2009, Leuven, Belgium, 22–25 February 2009. Revised Selected Papers, pp. 260–276 (2009).
28. National Bureau of Standards: Data Encryption Standard. US Department of Commerce, FIPS Publication 46 (1977).
29. Stevens M., Bursztein E., Karpman P., Albertini A., Markov Y.: The first collision for full SHA-1. In: Advances in Cryptology—CRYPTO 2017—37th Annual International Cryptology Conference, Santa Barbara, CA, USA, 20–24 August 2017. Proceedings, Part I, pp. 570–596 (2017).
30. Sun B., Liu Z., Rijmen V., Li R., Cheng L., Wang Q., AlKhzaimi H., Li C.: Links among impossible differential, integral and zero correlation linear cryptanalysis. In: Advances in Cryptology—CRYPTO 2015—35th Annual Cryptology Conference, Santa Barbara, CA, USA, 16–20 August 2015. Proceedings, Part I, pp. 95–115 (2015).
31. Sun B., Liu M., Guo J., Rijmen V., Li R.: Provable security evaluation of structures against impossible differential and zero correlation linear cryptanalysis. In: Advances in Cryptology—EUROCRYPT 2016—35th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Vienna, Austria, 8–12 May 2016. Proceedings, Part I, pp. 196–213 (2016).
32. Sun S., Gerault D., Lafourcade P., Yang Q., Todo Y., Qiao K., Hu L.: Analysis of AES, skinny, and others with constraint programming. IACR Trans. Symmetric Cryptol. **2017**(1), 281–306 (2017).
33. Sun L., Wang W., Wang M.: More accurate differential properties of LED64 and Midori64. IACR Trans. Symmetric Cryptol. **2018**(3), 93–123 (2018).
34. Tolba M., Abdelkhalek A., Youssef A.M.: Truncated and multiple differential cryptanalysis of reduced round Midori128. In: Information Security—19th International Conference, ISC 2016, Honolulu, HI, USA, 3–6 September 2016. Proceedings, pp. 3–17 (2016).
35. Wagner D.: The boomerang attack. In: Fast Software Encryption, 6th International Workshop, FSE '99, Rome, Italy, 24–26 March 1999. Proceedings, pp. 156–170 (1999).
36. Wang X., Yu H.: How to break MD5 and other hash functions. In: Advances in Cryptology—EUROCRYPT 2005, 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Aarhus, Denmark, 22–26 May 2005. Proceedings, pp. 19–35 (2005).
37. Wang X., Yin Y.L., Yu H.: Finding collisions in the full SHA-1. In: Advances in Cryptology—CRYPTO 2005: 25th Annual International Cryptology Conference, Santa Barbara, California, USA, 14–18 August 2005. Proceedings, pp. 17–36 (2005).
38. Wang G., Keller N., Dunkelman O.: The delicate issues of addition with respect to XOR differences. In: Selected Areas in Cryptography, 14th International Workshop, SAC 2007, Ottawa, Canada, 16–17 August 2007. Revised Selected Papers, pp. 212–231 (2007).
39. Wang M., Sun Y., Tischhauser E., Preneel B.: A model for structure attacks, with applications to PRESENT and Serpent. In: Fast Software Encryption—19th International Workshop, FSE 2012, Washington, DC, USA, 19–21 March 2012. Revised Selected Papers, pp. 49–68 (2012).

Affiliations

Yunwen Liu[1,2] · Wenying Zhang[4] · Bing Sun[1] · Vincent Rijmen[2,3] · Guoqiang Liu[1] · Chao Li[1] · Shaojing Fu[5] · Meichun Cao[4]

[1] Department of Mathematics, National University of Defense Technology, Changsha, China

[2] imec-COSIC KU Leuven, Leuven, Belgium

[3] University of Bergen, Bergen, Norway

[4] School of Information Science and Engineering, Shandong Normal University, Jinan, China

[5] College of Computer, National University of Defense Technology, Changsha, China