Statistical Integral Distinguisher with Multi-structure and Its Application on AES

Tingting Cui[1,2,4], Ling Sun[1], Huaifeng Chen[1], and Meiqin Wang[1,3(✉)]

[1] Key Laboratory of Cryptologic Technology and Information Security, Ministry of Education, Shandong University, Jinan 250100, China
mqwang@sdu.edu.cn
[2] Science and Technology on Communication Security Laboratory, Chengdu 610041, China
[3] State Key Laboratory of Cryptology, P.O. Box 5159, Beijing 100878, China
[4] Nanyang Technological University, Singapore, Singapore

Abstract. Advanced Encryption Standard (AES), published by NIST, is widely used in data encryption algorithms, hash functions, authentication encryption schemes and so on. Studying distinguishing attacks on (reduced round) AES can help designers and cryptanalysts to evaluate the security of target ciphers. Since integral attack is one of the most powerful tool in the field of symmetric ciphers, in this paper, we evaluate the security of AES by integral cryptanalysis. Firstly we put forward a new statistical integral distinguisher with multiple structures on input and integral properties on output, which enables us to reduce the data complexity comparing to the traditional integral distinguishers under multiple structures. As illustrations, we propose a secret-key distinguisher on 5-round AES with secret S-box under chosen-ciphertext mode. Its data, time and memory complexities are $2^{114.32}$ chosen ciphertexts, 2^{110} encryptions and $2^{33.32}$ blocks. This is the best integral distinguisher on AES with secret S-box under secret-key setting so far. Then we present improved known-key distinguishers on 8-round and full 10-round AES-128 with reduced complexities based on Gilbert's work at ASIACRYPT'14. These distinguishers are the best ones according to the time complexity. Moreover, the proposed statistical integral model could be used to proceed known-key distinguishing attacks on other AES-like ciphers.

Keywords: Statistical integral model · Secret S-box · Secret key · Known key · AES

1 Introduction

Advanced Encryption Standard (AES) [8], published by NIST, is widely used in the field of symmetric ciphers. For instances, AES and reduced-round versions of AES are usually used as components for hash functions, authentication encryption schemes and so on. Since the goal of distinguishing attack is to distinguish

J. Pieprzyk and S. Suriadi (Eds.): ACISP 2017, Part I, LNCS 10342, pp. 402–420, 2017.
DOI: 10.1007/978-3-319-60055-0_21

a target cipher from random permutations with some special property, studying the distinguishers on AES can help designers and cryptanalysts to evaluate the security of target cipher, which is meaningful.

In secret-key distinguishing attack, adversary needs to distinguish the target cipher from random permutations without knowing the key and internal states. Such distinguisher can be used in key-recovery attack. Furthermore, reduced round AES are often utilized to design authentication encryptions such as the third-round candidates AES-OTR [20] in CAESAR competition [6]. It is necessary to research the secret-key distinguisher on AES. Beside that, the performances of block ciphers under known key settings need be considered. Block ciphers, because of their security and simplicity, are often adopted as components of hash functions by designers, such as Whirlpool [3] and Photon [13]. Since the attacker can fully control the inter behaviour of a hash function, if a block cipher is used to design hash function, its resistance to known-key attack or chosen-key attack, where the adversaries know the key or can choose the key, should be considered. The first known-key security model is proposed by Knudsen and Rijmen for block cipher in [15] where the secret key is known to the attacker and the goal is to distinguish the block cipher from a random permutation by constructing a set of plaintext/ciphertext pairs satisfying a special property. Such a property is easy to check but impossible to achieve for any random permutation with the same complexity and a non-negligible probability by using oracle accesses to this random permutation and its inverse. Since its establishment, several types of known-key distinguishers have been proposed, such as distinguishers with integral property [1,12,15,21], subspace distinguishers [17,18], (multiple) limited-birthday distinguishers [11,14], and the known-key distinguisher for PRESENT by combining meet-in-the-middle technique and truncated differential [5]. Moreover, the chosen-key distinguishing attack on the full AES-256 has been provided in [4].

Integral attack is an important cryptanalytic technique for symmetric-key ciphers, which was firstly put forward by Daemen *et al.* in [7], then unified as integral attack by Knudsen and Wagner in [16]. In an integral distinguisher, one fixes a part of plaintext bits and takes all possible values for the other plaintext bits such that the values on partial bits of ciphertext are uniformly distributed, to distinguish an actual cipher from a random permutation. If one additional linear layer is considered, the property will be that the XOR of all possible values of the specific part of ciphertext becomes zero, which is referred as zero-sum property [2]. In order to reduce the data complexity, Wang *et al.* applied statistical technique on original integral distinguisher and proposed a statistical integral distinguisher at FSE'16 [24], which consists of applying a statistical technique to the original integral distinguisher with the active property. As a result, this statistical integral distinguisher requires less data complexity than that of the original integral distinguisher. However, Wang *et al.* only considered the case that only one integral property on ciphertext, they didn't discuss the cases that there are several integral properties on ciphertext and multiple structures of data should be used at the same time. These limit the effect of integral attacks on block ciphers, especially for known-key distinguishing attacks.

In this paper, we consider the cases omitted in [24] and use our statistical integral model to improve secret-key and known-key distinguishing attacks on AES with further less data and time complexities.

1.1 Our Contributions

Statistical Integral Distinguisher with Multiple Structures. We propose a statistical integral distinguisher with multiple structures on input and integral properties on output. In some situations of integral attacks such as known-key distinguishing attack on AES, multiple structures of input have to be used where for each structure s input bits take all possible values and the corresponding b t-bit outputs are uniformly distributed respectively. The statistical integral distinguisher in [24] can reduce the data complexity from $\mathcal{O}(2^s)$ to $\mathcal{O}(2^{s-t/2})$ by using one t-bit integral property if only one structure is used. But if there are N_s structures involved, the model in [24] cannot be applied. For the sake of reducing the data requirements for the original integral distinguisher with multiple structures, we construct a new statistical integral distinguisher. In our new distinguisher, the data complexity is

$$\mathcal{O}(\sqrt{N_s/b} \cdot 2^{s-\frac{t}{2}}),$$

while the data complexity of the original distinguisher is

$$\mathcal{O}(N_s \cdot 2^s).$$

In order to verify our theoretical model, we implement the experiments for mini version of AES. It shows that the experimental results are in good accordance with the theoretic results.

Improved Secret-Key Integral Distinguisher on AES. AES is one of the most famous block ciphers. Until 2015, the best secret-key distinguishers on AES were 4 rounds, such as impossible differential, zero-correlation linear hull and integral distinguisher. Then at CRYPTO'16, Sun *et al.* proposed a 5-round distinguisher on AES with secret S-box under chosen-ciphertext mode with integral zero-correlation technique in [23]. But the data complexity of this distinguisher is up to 2^{128}. Recently, Grassi *et al.* put forward a 5-round distinguisher on AES with secret S-box by utilizing a 4-round impossible differential in [9]. The data complexity is $2^{98.2}$. Later, they proposed another one on 5-round AES in [10]. That distinguisher is independent with the details of S-box, MC operation and secret-key, and its data complexity is reduced to 2^{32}. However, it utilizes the property of AES structure and has nothing with the secret-key, this weakness limits it to be used in key recovery attacks. In this paper, we will evaluate the security of AES from the point of integral distinguishing attack. We present a secret-key distinguisher on 5-round AES with secret S-box by adopting our statistical integral model under chosen-ciphertext mode. The data and time complexities are $2^{114.32}$ chosen ciphertexts and 2^{110} encryptions respectively. Its memory requirements are $2^{33.32}$ blocks. This is the best integral distinguisher on AES with secret S-box under secret-key setting so far (Table 1).

Improved Known-Key Distinguishers on AES. We apply the statistical integral distinguisher with multiple structures into the known-key distinguishing attacks on AES. The first known-key distinguisher on AES was proposed by Knudsen and Rijmen in [15], where they gave an integral known-key distinguisher for 7-round AES. At ASIACRYPT'14, Gilbert provided a very important untwisted representation of AES and used this representation to distinguish 8-round AES and the full 10-round AES with the complexity 2^{64} under the known-key model in [12]. Besides the integral known-key distinguishers, the known-key distinguisher with match-in-the-middle technique for 7-round AES was presented in [19], with rebound technique for 8-round AES were provided in [11,14] whose complexities are 2^{48} and 2^{44} 8-round encryptions respectively. In this paper, we take advantage of our statistical integral model to improve known-key distinguisher on 8-round AES and full 10-round AES, whose respective time complexities are $2^{42.61}$ computations and $2^{59.60}$ computations. These distinguishers are the best known-key ones on AES according to the time complexity so far. See Table 2.

Table 1. Summary of secret-key integral distinguishers on AES

Type	Rounds	Data (CC)	Time	Memory	Source
Integral	5	2^{128}	2^{128}	-	[23]
Statistical integral	**5**	$\mathbf{2^{114.32}}$	$\mathbf{2^{110}}$	$\mathbf{2^{33.32}}$	**Section 4**

CC: Chosen-cipertext

Table 2. Summary of known-key distinguishing attacks on AES

Type	Rounds	Time	Memory	Source
Integral	7	2^{56}	–	[15]
MITM	7	2^{24}	–	[19]
Limited-birthday	8	2^{48}	2^{35} bytes	[11]
Multiple limited-birthday	8	2^{44}	2^{35} bytes	[14]
Integral	8	2^{64}	–	[12]
Statistical integral	**8**	$\mathbf{2^{42.61}}$	$\mathbf{2^{13}}$ **bytes**	**Section 5**
Integral	10	2^{64}	–	[12]
Statistical integral	**10**	$\mathbf{2^{59.60}}$	$\mathbf{2^{58.84}}$ **bytes**	**Section 5**

MITM: Match-in-the-middle

1.2 Ontline of This Paper

In Sect. 2, some preliminaries are given. Then we present a statistical integral model with multiple structures on input and integral properties on output in Sect. 3. In Sects. 4 and 5, secret-key statistical integral distinguisher and improved known-key distinguishers on AES are put forward respectively. At last, we conclude this paper in Sect. 6.

2 Preliminaries

2.1 Description of AES

AES is a byte-orient Substitution-Permutation Network (SPN). It has three versions, namely AES-128, -192 and -256. The block-size/key-size/total-rounds of these versions are 128/128/10, 128/192/12 and 128/256/14 respectively. Each round function includes 4 components:

- SubBytes (SB): A nonlinear bijective mapping $\mathbb{F}_2^8 \to \mathbb{F}_2^8$ for each byte of state;
- ShiftRows (SR): Left rotate the i-th row by i bytes, where $i = 0, 1, 2, 3$;
- MixColumns (MC): Left multiply with an MDS matrix over the field $GF(2^8)$ on each column;
- AddRoundKey (AK): XOR with a 128 bits subkey.

It is worth noting that there is a whiten key XORed with plaintext before the first round function and the MC operation is omitted in the last round. Since we do not use the key schedule in this paper, we ignore it here.

All in all, $2r$-round AES can be described as follows:

$$AES_{2r} = AK \diamond (SB \diamond SR \diamond MC \diamond AK)^{2r-1} \diamond SB \diamond SR \diamond AK \tag{1}$$

where $A \diamond B$ denotes to implement A operation firstly, then B operation.

In [12], Gilbert proposed a new representation of AES. Firstly he defined two operations T and SC as follows, then built two special byte permutations $P = SR \diamond T \diamond SR^{-1}$ and $Q = SR^{-1} \diamond T \diamond SR \diamond SC$. With these two permutations, Gilbert proposed two transformations $S = Q^{-1} \diamond SB \diamond MC \diamond AK \diamond SB \diamond P^{-1}$ and $R = P \diamond SR \diamond MC \diamond AK \diamond SR \diamond Q$, which operate on columns and rows respectively.

$$T : \begin{pmatrix} a_0 & a_4 & a_8 & a_{12} \\ a_1 & a_5 & a_9 & a_{13} \\ a_2 & a_6 & a_{10} & a_{14} \\ a_3 & a_7 & a_{11} & a_{15} \end{pmatrix} \mapsto \begin{pmatrix} a_0 & a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 & a_7 \\ a_8 & a_9 & a_{10} & a_{11} \\ a_{12} & a_{13} & a_{14} & a_{15} \end{pmatrix}$$

$$SC : \begin{pmatrix} a_0 & a_4 & a_8 & a_{12} \\ a_1 & a_5 & a_9 & a_{13} \\ a_2 & a_6 & a_{10} & a_{14} \\ a_3 & a_7 & a_{11} & a_{15} \end{pmatrix} \mapsto \begin{pmatrix} a_0 & a_{12} & a_8 & a_4 \\ a_1 & a_{13} & a_9 & a_5 \\ a_2 & a_{14} & a_{10} & a_6 \\ a_3 & a_{15} & a_{11} & a_7 \end{pmatrix}$$

As a result, $2r$-round AES has three equivalent representations:

$$AES_{2r} = AK \diamond SR \diamond Q \diamond (S \diamond R)^{r-1} \diamond S \diamond P \diamond SR \diamond AK, \tag{2}$$

$$AES_{2r} = AK \diamond P^{-1} \diamond SB \diamond R \diamond (S \diamond R)^{r-1} \diamond SB \diamond Q^{-1} \diamond AK, \tag{3}$$

$$AES_{2r} = AK \diamond SB \diamond SR \diamond MC \diamond AES_{2r-2} \diamond AK^{-1} \diamond MC \diamond AK \cdot SB \diamond SR \diamond AK. \tag{4}$$

Throughout this paper, we use $X_{(i)}$ and $X_{(i\sim j)}, i, j = 0, 1, \ldots, 15$ to denote the i-th byte and $i \sim j$-th bytes of state X respectively.

2.2 Brief Description of Known-Key Distinguishers on AES in [12]

In this subsection, we briefly recall the known-key distinguishers for 8-round and 10-round AES proposed by Gilbert at ASIACRYPT'14 [12].

In order to mount a known-key distinguisher for AES_8, Gilbert firstly proposed two integral distinguishers shown in Fig. 1, where $(A_1^j, A_2^j, A_3^j, A_4^j), j = 0, 1, \ldots, 4$, A and C denote uniform distribution on 4 bytes, uniform distribution on 1 bytes and constant respectively. Then given 2^{64} data $\mathcal{Z} = \{R(x, 0, 0, 0) \oplus (y, 0, 0, 0) | x, y \in (0, 1)^{32}\}$, this set $\mathcal{Z}$ can be divided into 2^{32} structures according to different values of x, and each structure takes all 2^{32} values on the first column and constants on other columns. So the set $\mathcal{Z}$ satisfies the first integral distinguisher in Fig. 1. Since R operation is an affine mapping, $R(\mathcal{Z}^{-1}) = \{(x, 0, 0, 0) \oplus R^{-1}(y, 0, 0, 0)\}$ can be divided into 2^{32} structures according to different values of y, thus the set $R(\mathcal{Z}^{-1})$ satisfies the second integral distinguisher in Fig. 1.

Combining with these two integral distinguishers with R operation above, a known-key distinguisher on AES_8 is built that all input and output bytes resulted from 2^{64} middle texts $\mathcal{Z}$ are uniformly distributed. However, for random permutations, the upper bound of the probability satisfying the uniformly distributed property for each byte is $\frac{1}{2^{128-1}}$ with $q \leq N = 2^{64}$ oracle queries.

Furthermore, with the representation of Eq. (3), Gilbert mounted a known-key distinguisher for AES_{10}. This distinguisher is implemented by extending one round on each side based on the distinguisher for AES_8. The same 2^{64} middle texts $\mathcal{Z}$ as for the known-key distinguisher on AES_8 are used. For the corresponding input-output pairs $(p_i, c_i), i = 1, \ldots, 2^{64}$, the adversary can find at least one value (Δ, Γ), where $\Delta, \Gamma \in (0, 1)^{128}$, to make each byte of $R \circ SB(P^{-1}(p_i) \oplus \Delta)$ and $R^{-1} \circ SB^{-1}(Q(c_i) \oplus \Gamma)$ be uniform distribution within time complexity 2^{64}. However, for a random permutation, the upper bound of the probability satisfying the uniformly distributed property for each byte is $2^{-16.5}$ with $q \leq N = 2^{64}$ oracle queries.

Since Gilbert's work is based on the integral distinguisher and uses the active property[1], if we can improve the statistical integral model proposed by Wang *et al.* in [24], we can further improve Gilbert's work and widely utilize the new method to all AES-like ciphers. With the improved known-key distinguishers, 10-round AES-like ciphers cannot be regarded as ideal random permutations, and the time complexities of new distinguishers are less than previous ones.

[1] Active property means that the values on target bits are uniform distributed.

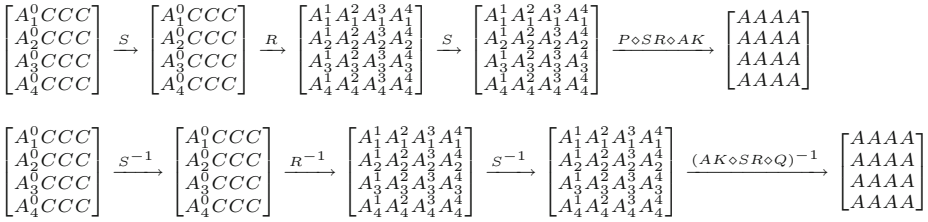

Fig. 1. Two integral distinguishers under the new representation of AES in [12]

2.3 Statistical Integral Distinguisher

In this subsection, we recall the statistical integral distinguisher proposed by Wang *et al.* in [24].

Assume that $H : \mathbb{F}_2^n \to \mathbb{F}_2^n$ is a part of a block cipher, its input and output both can be split into two parts as follows:

$$H : \mathbb{F}_2^r \times \mathbb{F}_2^s \to \mathbb{F}_2^t \times \mathbb{F}_2^u, H(x, y) = \begin{pmatrix} H_1(x, y) \\ H_2(x, y) \end{pmatrix}.$$

If the first r bits of input are fixed as a constant λ and only the first t bits of output are considered, then the function H can be denoted as T_λ:

$$T_\lambda : \mathbb{F}_2^s \to \mathbb{F}_2^t, T_\lambda(y) = H_1(\lambda, y).$$

When y takes over all possible values, the outputs $T_\lambda(y)$ are uniformly distributed, then an integral distinguisher is constructed.

If the adversary only takes $N < 2^s$ different y, sets a counter $V[T_\lambda(y)]$ and initializes this counter as zero, a statistical integral distinguisher can be constructed by investigating the distribution of the statistic as follows:

$$T = \sum_{i=0}^{2^t-1} \frac{(V[T_\lambda(y)] - N \cdot 2^{-t})^2}{N \cdot 2^{-t}} \tag{5}$$

For the right key guess (the target cipher), the statistic T follows a χ^2 distribution with mean $\mu_0 = (2^t - 1)\frac{2^s - N}{2^s - 1}$ and variance $\sigma^2 = 2(2^t - 1)(\frac{2^s - N}{2^s - 1})^2$, but for the wrong key guess (a random permutation), it follows a χ^2 distribution with mean $\mu_0 = (2^t - 1)$ and variance $\sigma^2 = 2(2^t - 1)$. The relation of data complexity, type-I error probability α_0 and type-II error probability α_1 is as follows

$$N = \frac{(2^s - 1)(q_{1-\alpha_0} + q_{1-\alpha_1})}{\sqrt{(2^t - 1)/2} + q_{1-\alpha_1}} + 1, \tag{6}$$

3 Statistical Integral Distinguisher with Multiple Structures on Input and Integral Properties on Output

In some integral distinguishers, there are b groups of t output bits with the active property. If we can utilize all properties at the same time, the data complexity can be further reduced. What's more, in some attack settings, N_s structures, i.e. that N_s different λ, should be used together. For these special settings, we construct the new statistical integral distinguisher in this section.

Firstly, we split the input into two parts and output into $b+1$ parts.

$$H : \mathbb{F}_2^r \times \mathbb{F}_2^s \to \mathbb{F}_2^t \times \mathbb{F}_2^t \times \ldots \times \mathbb{F}_2^t \times \mathbb{F}_2^u,\ H(x,y) = \begin{pmatrix} H_1(x,y) \\ H_2(x,y) \\ \ldots \\ H_{b+1}(x,y) \end{pmatrix}.$$

Then we use T_λ^i to denote the function H_i where the first r bits of its input are fixed to the value λ and b outputs $H_i, 1 \le i \le b$, are considered:

$$T_\lambda^i : \mathbb{F}_2^s \to \mathbb{F}_2^t,\ T_\lambda^i(y) = H_i(\lambda, y), i = 1, 2, \ldots, b.$$

For a special integral distinguisher, when y iterates all possible values of $\mathbb{F}_2^s$, $T_\lambda^i(y), i = 1, 2, \ldots, b$ are all uniformly distributed with probability one. Further more, if we take N_s values for λ, *i.e.* N_s structures and in each structure y iterates all possible values of $\mathbb{F}_2^s$, the integral properties on output are satisfied as well.

Now assume we need $N < 2^s$ values of y under each structure and we use N_s structures which are independent. $T_\lambda^i(y) \in \mathbb{F}_2^t, i = 1, 2, \ldots, b$ are computed for each y and we allocate a counter vector $V_i[T_\lambda^i(y)]$ to store the occurrences of $T_\lambda^i(y)$. Then we investigate the distribution of the following statistic:

$$C = \sum_{\lambda=1}^{N_s} \sum_{i=1}^{b} \sum_{T_\lambda^i(y)=0}^{2^t-1} \frac{(V_i[T_\lambda^i(y)] - N \cdot 2^{-t})^2}{N \cdot 2^{-t}}. \tag{7}$$

The statistic C follows different distributions determined by whether we are dealing with an actual cipher or a random permutation.

Proposition 1. *For sufficiently large N, and t, the statistic $\frac{2^s-1}{2^s-N} C_{cipher}$ (C_{cipher} is the statistic C for cipher) follows a χ^2-distribution with degree of freedom $b \cdot N_s \cdot (2^t - 1)$, which means that C_{cipher} approximately follows a normal distribution with mean and variance*

$$\mu_0 = Exp(C_{cipher}) = b \cdot N_s \cdot (2^t - 1)\frac{2^s - N}{2^s - 1},\ \sigma_0^2 = Var(C_{cipher}) = 2b \cdot N_s \cdot (2^t - 1)(\frac{2^s - N}{2^s - 1})^2.$$

The statistic C_{random} (C_{random} is the statistic C for randomly drawn permutation) follows a χ^2-distribution with degree of freedom $b \cdot N_s \cdot (2^t - 1)$, which means that C_{random} approximately follows a normal distribution with mean and variance

$$\mu_1 = Exp(C_{random}) = b \cdot N_s \cdot (2^t - 1) \text{ and } \sigma_1^2 = Var(C_{random}) = 2b \cdot N_s \cdot (2^t - 1).$$

Proof. Deduced from Proposition 1 in [24], for a randomly drawn permutation, the statistic $\sum_{T_\lambda^i(y)=0}^{2^t-1} \frac{(V_i[T_\lambda^i(y)]-N\cdot 2^{-t})^2}{N\cdot 2^{-t}}$ follows a χ^2-distribution with degree of freedom 2^t-1 for any λ and i. Then the statistic C'_{random} for the randomly drawn permutation

$$C_{random} = \sum_{\lambda=1}^{N_s}\sum_{i=1}^{b}\sum_{T_\lambda^i(y)=0}^{2^t-1} \frac{(V_i[T_\lambda^i(y)] - N\cdot 2^{-t})^2}{N\cdot 2^{-t}}$$

is the sum of $N_s\cdot b$ independent χ^2 statistics with degree of freedom 2^t-1, so the statistic C_{random} follows a χ^2-distribution with degree of freedom $b\cdot N_s\cdot(2^t-1)$. Then for sufficiently large N and t, C_{random} approximately follows a normal distribution with the expected value and variance:

$$Exp(C_{random}) = b\cdot N_s\cdot(2^t-1) \text{ and } Var(C_{random}) = 2b\cdot N_s\cdot(2^t-1).$$

Since the statistic for the cipher $\frac{2^s-1}{2^s-N}\sum_{T_\lambda^i(y)=0}^{2^t-1} \frac{(V_i[T_\lambda^i(y)]-N\cdot 2^{-t})^2}{N\cdot 2^{-t}}$, for any λ and i, follows a χ^2-distribution with degree of freedom 2^t-1 deduced from [24]. Then the statistic $\frac{2^s-1}{2^s-N}C'_{cipher}$ for the cipher

$$\frac{2^s-1}{2^s-N}C_{cipher} = \sum_{\lambda=1}^{N_s}\sum_{i=1}^{b}\frac{2^s-1}{2^s-N}\sum_{T_\lambda^i(y)=0}^{2^t-1} \frac{(V_i[T_\lambda^i(y)] - N\cdot 2^{-t})^2}{N\cdot 2^{-t}}$$

is the sum of $N_s\cdot b$ independent χ^2 statistics with degree of freedom 2^t-1, so the statistic $\frac{2^s-1}{2^s-N}C_{cipher}$ follows a χ^2-distribution with degree of freedom $b\cdot N_s\cdot(2^t-1)$. Then for sufficiently large N and t, C_{cipher} approximately follows a normal distribution with the expected value and variance:

$$Exp(C_{cipher}) = b\cdot N_s\cdot(2^t-1)\cdot\frac{2^s-1}{2^s-N} \text{ and } Var(C_{cipher}) = 2b\cdot N_s\cdot(2^t-1)\cdot(\frac{2^s-1}{2^s-N})^2.$$

□

Corollary 1. *Under the assumption of Proposition 1, for type-I error probability α_0 (the probability to wrongfully discard the cipher), and type-II error probability α_1 (the probability to wrongfully accept a randomly chosen permutation as the cipher), to distinguish a cipher and a random permutation based on b independent t-bit outputs when randomly choosing N_s values for r-bit inputs and N values for s-bit inputs, then the following equation holds.*

$$N = \frac{(2^s-1)(q_{1-\alpha_0}+q_{1-\alpha_1})}{\sqrt{(b\cdot N_s\cdot(2^t-1))/2}+q_{1-\alpha_0}} + 1, \tag{8}$$

where $q_{1-\alpha_0}$ and $q_{1-\alpha_1}$ are the respective quantiles of the standard normal distribution.

Corollary 1 is obtained from the equation about the decision threshold $\tau = \mu_0+\sigma_0 q_{1-\alpha_0} = \mu_1-\sigma_1 q_{1-\alpha_1}$. And the statistic test is also based on the decision

threshold τ: if $C \leq \tau$, the test outputs 'cipher'; Otherwise, if the statistic $C > \tau$, the test outputs 'random'. Note that in this statistical method the success probability $Ps = 1 - \alpha_0$, and the relation between α_1 and the advantage of the attack a is $\alpha_1 = 2^{-a}$.

In order to verify the theoretical model in Corollary 1, we implement the experiments for mini version of AES in Appendix A.1. It shows that the experimental results are in good accordance with the theoretic results.

From Eq. (8), we know that the data complexity for the statistical distinguisher is $N \cdot N_s$. For the given values of $n, s, t, \alpha_0, \alpha_1$, the ratio of the data complexity with N_s structures to that with one structure is $\sqrt{N_s}$. It means that more structures will result in high data complexity, so we should avoid to utilize more structures. However, for the known-key integral distinguisher for AES etc., we have to use enough structures to make the plaintexts and the ciphertexts satisfying the desired properties simultaneously. Moreover, if b is increased, the data complexity can be reduced, but as b increases, the time complexity in some situations will be increased accordingly. Thus, we should take the proper value for b according to the time-data tradeoff.

4 Secret-Key Statistical Integral Distinguisher on Reduced 5-Round AES

In this section, we propose a secret-key distinguisher on 5-round AES with our statistical integral model based on the work of Sun *et al.* in [23]. In this distinguisher, the S-box used in AES is secret.

Firstly, we slightly modify the zero-correlation linear hull for 5-round decryption of AES under chosen-ciphertext mode proposed by Sun *et al.* in [23] (Lemma 3). Let $V = \{(x_{(i)}) \in F_{2^8}^{16} | x_{(0)} \oplus x_{(13)} = (k_5)_{(0)} \oplus (k_5)_{(13)}\}$, and assume that the input mask $\Gamma_I = (a_{(i)})_{0 \leq i \leq 15}$ and output mask $\Gamma_O^0 = (\beta_{(i)})_{0 \leq i \leq 15}$ satisfy:

$$a_{(i)} = \begin{cases} a, i = 0, 13, \\ 0, otherwise. \end{cases} \qquad \beta_{(j)} = \begin{cases} nonzero, j = \{0, 5, 10, 15\} \\ 0, otherwise. \end{cases}$$

Then the correlation for $\Gamma_I \rightarrow \Gamma_O^0$ on V is always 0. Note that there are three other zero-correlation linear hulls as well, when $j = \{1, 6, 11, 12\}$, $\{2, 7, 8, 13\}$, $\{3, 4, 9, 14\}$. The corresponding output masks are Γ_O^1, Γ_O^2 and Γ_O^3 respectively. One of the four cases is shown in Fig. 2.

With the technique proposed by Sun *et al.* in [22], these four zero-correlation linear hulls can be transformed into integral ones. Taking the linear hull $\Gamma_I \rightarrow \Gamma_O^0$ as an example, the corresponding integral distinguisher is that if the adversary takes over 2^{120} different values of ciphertexts c satisfying $c_{(0)} \oplus c_{(13)} = (k_5)_{(0)} \oplus (k_5)_{(13)}$, then the values on 4 bytes of plaintext $(p_{(0)}, p_{(5)}, p_{(10)}, p_{(15)})$ are uniformly distributed.

Based on these integral distinguishers, we can implement a statistical integral distinguisher for each candidate $\Delta = (k_5)_{(0)} \oplus (k_5)_{(13)}$, where $s = 120$ and $t = 32$. In order to have the success probability $(1 - \alpha_0)^{2^8} = (1 - \alpha_1)^{2^8} = 95\%$, we set

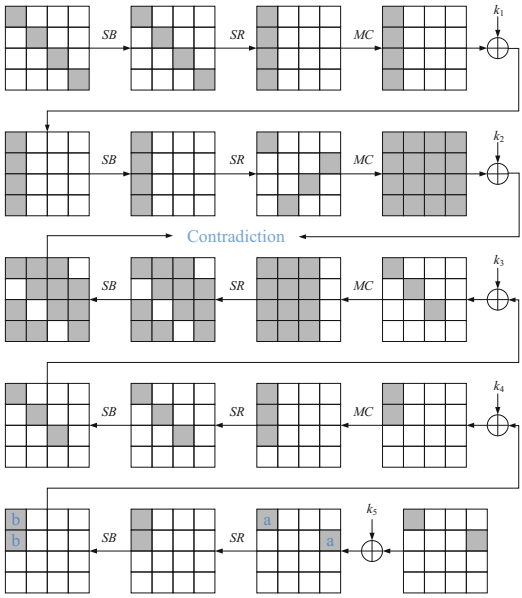

Fig. 2. Zero-correlation linear hull on 5-round AES with secret S-box under secret-key setting. Gray and white cells denote nonzero and zero masks respectively. The two cells with a or b are exactly the same mask.

Algorithm 1. Secret-key statistical integral distinguisher on 5-round AES with secret S-box

for 2^8 *candidates of* Δ **do**
  Set a counter $V[4][2^{32}]$ and initialize it to zero;
  for N *chosen ciphertext/plaintext pairs* (c, p) **do**
    `// Consider those four integrals together.`
    for $i \leftarrow 0 \sim 3$ **do**
      Increment counter $V[i][c^i_{part}]$ by one according to the related 4 bytes $c^i_{part} \in (0,1)^{32}$ of ciphertext c;
  Calculate the statistic $T_\Delta = \sum_{b=0}^{3} \sum_{z=0}^{2^{32}-1} \frac{(V[b][z]-N\cdot 2^{-32})^2}{N\cdot 2^{-32}}$;
if *Only one* Δ *such that* $T_\Delta < \tau$ **then**
  return AES;
return random permution;

$\alpha_0 = \alpha_1 = 0.0002$, then $q_{1-\alpha_0} = q_{1-\alpha_1} \approx 3.54$. Meanwhile, we can use these four integral distinguishers above together within one structure, so $b = 4$ and $N_s = 1$. Thus by Eq. (8), $N = 2^{106.32}$ chosen ciphertexts. The decision threshold is about $\tau \approx 17179212992.15$. As there are 2^8 different values of Δ, the total data complexity of this distinguisher is $N' = 2^{106.32} \times 2^8 = 2^{114.32}$ chosen ciphertexts.

What's more, we can see from Algorithm 1, the main time complexity happens on Step 5, which is about $2^8 \times 2^{106.32} \times 4 \times 1/16 \times 1/5 \approx 2^{110}$ encryptions, if we regard one simple operation as $\frac{1}{16}$ one round encryption. Beside that, memory requirements are about $4 \times 2^{32} \times 10 \approx 2^{37.32}$ bytes $= 2^{33.32}$ blocks.

As far as we know, this distinguisher is the best secret-key integral one on 5-round AES with secret S-box.

5 Improved Known-Key Distinguishers on AES

In this section, we will use our new statistical integral model to reduce the complexities of known-key distinguishers on AES proposed by Gilbert at ASIACRYPT'14 in Sects. 5.1 and 5.2.[2] The time complexity is reduced to $2^{42.61}$ in the known-key distinguisher on 8-round AES. For the 10-round AES, the time complexity is reduced to $2^{59.60}$. Compared to all the public known-key distinguishers for 8-round AES, our distinguisher is the best one according to both time and memory complexities. Moreover, our known-key distinguisher on 10-round AES is the best one according to the time complexity.

5.1 Improved Known-Key Distinguisher on 8-Round AES

As described in Subsect. 2.2, the known-key distinguisher for AES_8 is based on the uniformly distributed integral property with 2^{32} structures and each structure takes 2^{32} texts. This integral property can be transformed to a statistical integral property by using Proposition 1. So in our known-key distinguisher on AES_8, we utilize the statistical integral properties on each byte of input and output to distinguish the actual cipher and random permutations. In this way, the required number of structures and texts of one structure can be reduced. The process to distinguish the actual cipher AES_8 from the random permutation is described in Algorithm 2.

Since in the middle of the distinguisher, the numbers of structures before and after R operation should be the same, i.e. that $N = N_s$. By applying Proposition 1 in above case, we have $s = 32$, $t = 8$, $b = 16$ and $N = N_s$. If we set the error probabilities $\alpha_0 = 2^{-128}$ and $\alpha_1 = 2^{-128}$ (the values of α_0 and α_1 can be different and take any suitable values), then $q_{1-\alpha_0} = q_{1-\alpha_1} \approx 13.06$. According to Eq. (8), $N = N_s \approx 2^{20.81}$ and the threshold value $\tau \approx 7478730631.39$.

For the case of AES_8, as $\alpha_0 = 2^{-128}$, the probability to wrongly regard AES_8 as a random permutation is $\alpha_0 + (1-\alpha_0)\alpha_0 \approx 2^{-127}$, which means that the success probability to correctly identify AES cipher is about $(1-\alpha_0)^2 \approx 1 - 2^{-127}$.

While for the case of random permutation, the adversary can implement encryption and decryption oracle queries to the cipher and random permutation. But statistical integral property (exploiting χ^2 distribution) is different

[2] These improved known-key distinguishers on AES in this paper follow the idea in Gilbert' work at ASIACRYPT'14, but we adopt statistical integral method instead of integral method and more delicate processes to reduce the data and time complexities.

Algorithm 2. Improved known-key distinguisher on AES_8

Initialize the statistic C' and C'' as zero;
for *all* N *values of* $x \in (0,1)^{32}$ **do**
  Initialize the counter vector $V[16][2^8]$ to zero;
  for *all* N *values of* $y \in (0,1)^{32}$ **do**
    Compute 16 bytes of input $p_{(l)}, l = 0, \ldots, 15$ from $Z = (x, 0, 0, 0) \oplus R(y, 0, 0, 0)$;
    Increment the corresponding counter $V[l][p_{(l)}]$ by one;
  $C' = C' + \sum_{l=0}^{15} \sum_{p_{(l)}=0}^{2^8-1} [\frac{(V[l][p_{(l)}] - N \times 2^{-8})^2}{N \times 2^{-8}}]$;
if $C' > \tau$ **then**
  return $\perp$; `// The distinguishing attack is failed.`
for *all* N *values of* $y \in (0,1)^{32}$ **do**
  Initialize the counter vector $V[16][2^8]$ to zero;
  for *all* N *values of* $x \in (0,1)^{32}$ **do**
    Compute 16 bytes of output $c_{(l)}, l = 0, \ldots, 15$ from $Z = (x, 0, 0, 0) \oplus R(y, 0, 0, 0)$;
    Increment the corresponding counter $V[l][c_{(l)}]$ by one;
  $C'' = C'' + \sum_{l=0}^{15} \sum_{c_{(l)}=0}^{2^8-1} [\frac{(V[l][c_{(l)}] - N \times 2^{-8})^2}{N \times 2^{-8}}]$;
For AES_8, $C'' \leq \tau$;
For any random permutation, $C'' > \tau$.

from traditional integral property (utilizing uniform distribution). At the best of times the adversary chooses the data which automatically satisfy the statistical property on the input, but to satisfy the statistical property on the output, the probability is $\alpha_1 = 2^{-128}$. In order to satisfy the statistical properties both on the input and output, the probability to wrongly regard this random permutation as AES cipher is $1 \times \alpha_1 = 2^{-128}$.

To summarize, the advantage to distinguish AES cipher from random permutation is not negligible. The total time complexity of this known-key distinguisher is about $2 \times 2^{41.61} = 2^{42.61}$ computations. The memory requirements are about $16 \times 2^8 \times 2 = 2^{13}$ bytes used for storing the counter vector $V[16][2^8]$.

5.2 Improved Known-Key Distinguisher on 10-Round AES

The statistical integral distinguisher on AES_{10} is based on the distinguishing property of AES_{10} in [12], which is represented according to Eq. (4), see Fig. 3.

Along with the idea within the distinguisher on AES_{10} in [12], in our known-key distinguisher on AES_{10}, we use $N_s < 2^{32}$ structures, each of which takes $N = N_s$ middle texts, to obtain N^2 input/output pairs. For AES cipher, there is one value for (Δ, Γ) to let each byte of $R \circ SB(R^{-1}(input \oplus \Delta))$ and $R^{-1} \circ SB^{-1}(Q(output \oplus \Gamma))$ satisfy the statistical integral property with a high probability. But for any random permutation, the probability to have one solution for (Δ, Γ) to obtain the same property is very low.

However, in above way, the distinguisher has high time complexity. In order to reduce the time complexity, we implement the distinguisher in the following way. As N_s structures are used, we divide them into N_s/n_s groups and each group has n_s structures. Then we compute the statistic value for each group. There is one value (Δ, Γ) to make all the statistics for N_s/n_s groups on both states $Input' = MC \circ SR \circ SB(input \oplus \Delta)$ and $Output' = MC^{-1} \circ SB^{-1} \circ SR^{-1}(output \oplus \Gamma)$ less than the given threshold τ for AES_{10}. However, for the random permutation, even if the attacker can carefully choose the inputs to find one value of Δ to satisfy the statistical property on the state $Input'$ with probability one, the probability to find one value Γ to satisfy the statistical property on the state $Output'$ is very low.

In order to further reduce the time complexity, we focus on statistics on 8-byte states – $Input'_{(0\sim3)}$ and $Output'_{(0\sim3)}$. So we only need to find two 32-bit values for $\Delta' = (\Delta_{(0)}, \Delta_{(5)}, \Delta_{(10)}, \Delta_{(15)})$ and $\Gamma' = (\Gamma_{(0)}, \Gamma_{(7)}, \Gamma_{(10)}, \Gamma_{(13)})$. The detailed process for this known-key distinguisher on AES_{10} is described in Algorithm 3.

In this setting, by applying Proposition 1, $s = 32$, $t = 8$, $b = 1$ and $n_s = 2^8$. If we set the error probabilities $\alpha_0 = 2^{-50}$ and $\alpha_1 = 2^{-10.51}$, then $N = 2^{27.92}$ and $\tau = 64123.53$ according to Eq. (8).

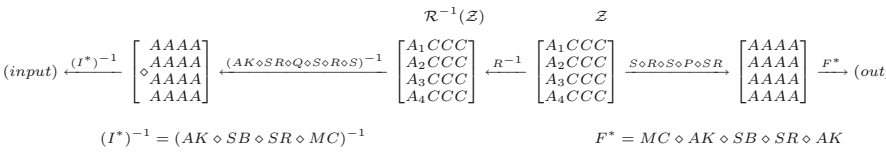

Fig. 3. Known-key distinguisher for AES_{10}. (A_1, A_2, A_3, A_4) and A denote uniform distribution on 4 bytes and 1 byte respectively. C denotes constant byte.

In Algorithm 3, we filter out the wrong values for $\Delta' = (\Delta_{(0)}, \Delta_{(5)}, \Delta_{(10)}, \Delta_{(15)})$ with the statistics on $Input'_{(0\sim3)}$ one by one. At last, the probability that one wrong Δ' is remained after all $2^{27.92-8}$ filtering processes is about $(2^{32}-1) \cdot \alpha_1^{4\times 2^{19.92}} \approx 0$, while the probability that the right candidate Δ cannot pass the filtering process is $1-(1-\alpha_0)^{4\times 2^{19.92}} \approx 2^{-28.08}$.

In the similar way, we filter out the wrong values for $\Gamma' = (\Gamma_{(0)}, \Gamma_{(7)}, \Gamma_{(10)}, \Gamma_{(13)})$ with the statistics for $Output'_{(0\sim3)}$ one by one. Finally, the probability that one wrong Γ' can pass the filtering process is also about 0, while the probability that the right Γ' cannot pass the filtering process is also $2^{-28.08}$. Therefore, for the case of AES_{10}, the probability to correctly identify the AES_{10} cipher is about $(1-2^{-28.08})^2 \approx 1-2^{-27.08}$.

While for the case of random permutation, at the best of the times the adversary can choose the inputs that there is always at least one value of Δ' remaining after the filtering process, but the probability that there is at least one Γ' surviving after the filtering process is about 0.

Algorithm 3. Improved known-key distinguisher on AES_{10}

Allocate vectors $V[N][N]$, $V'[N][N]$;
for *all* N^2 *values of* $(y_i, x_j), 0 \le i, j < N$ **do**
  Calculate input p and output c from $Z = (x_j, 0, 0, 0) \oplus R(y_i, 0, 0, 0)$ and let $V[j][i] = (p_{(0)}, p_{(5)}, p_{(10)}, p_{(15)})$, $V'[i][j] = (c_{(0)}, c_{(7)}, c_{(10)}, c_{(13)})$;
// Steps $4 \sim 30$ proceed the first group with n_s structures.
for *all* 2^{16} *values of* $(\Delta_{(0)}, \Delta_{(5)})$ **do**
  Allocate vectors $V_1[n_s][2^{24}]$;
  for *all* n_s *values of* j *and* N *values of* i **do**
    Get $(p_{(0)}, p_{(5)}, p_{(10)}, p_{(15)})$ from V[j][i];
    Compute $W_0 = 2 \cdot SB(p_{(0)} \oplus \Delta_{(0)}) \oplus 3 \cdot SB(p_{(5)} \oplus \Delta_{(5)})$; // $\cdot$ operate on F_2^8.
    Let $V_1[j][W_0, p_{(10)}, p_{(15)}]$ increase one;
  for *all* 2^8 *values of* $\Delta_{(10)}$ **do**
    Allocate a counter vectors $V_2[n_s][2^{16}]$, and initialize to zero;
    for *all* n_s *values of* j *and all* 2^{24} *values of* $W_0 \| p_{(10)} \| p_{(15)}$ **do**
      Compute $W_1 = W_0 \oplus (SB(p_{(10)} \oplus \Delta_{(10)}))$;
      Let $V_2[j][W_1, p_{(15)}]$ += $V_1[j][W_0, p_{(10)}, p_{(15)}]$;
    for *all* 2^8 *values of* $\Delta_{(15)}$ **do**
      Allocate counter vectors $V_3[n_s][2^8]$, and initialize to zero;
      for *all* n_s *values of* j *and all* 2^{16} *values of* $(W_1, p_{(15)})$ **do**
        $W' = W_1 \oplus (SB(p_{(15)} \oplus \Delta_{(15)}))$, let $V_3[j][W']$ += $V_2[j][W_1, p_{(15)}]$;
      $C_1 = \sum_{j=0}^{n_s-1} \sum_{W'=0}^{2^8-1} \frac{(V_3[j][W'] - N \times 2^{-8})^2}{N \times 2^{-8}}$;
      if $C_1 \le \tau$ **then**
        Put $\Delta' = (\Delta_{(0)}, \Delta_{(5)}, \Delta_{(10)}, \Delta_{(15)})$ into V_k. // About remain $2^{32} \cdot \alpha_1$ values.
for *all values of* $\Delta' \in V_k$ **do**
  Allocate counter vectors $V_4[n_s][2^8]$, and initialize to zero;
  for *all* n_s *values of* j *and* N *values of* i **do**
    Get $(p_{(0)}, p_{(5)}, p_{(10)}, p_{(15)})$ from V[j][i] and compute $Input'_1 = SB(p_{(0)} \oplus \Delta_{(0)}) \oplus 2 \cdot SB(p_{(5)} \oplus \Delta_{(5)}) \oplus 3 \cdot SB(p_{(10)} \oplus \Delta_{(10)}) \oplus SB(p_{(15)} \oplus \Delta_{(15)})$;
    Increment $V_4[j][Input'_1]$ by one;
  $C_2 = \sum_{j=0}^{n_s-1} \sum_{W=0}^{2^8-1} \frac{(V_4[j][W] - N \times 2^{-8})^2}{N \times 2^{-8}}$;
  if $C_2 \le \tau$ **then**
    Put Δ' into $V_{k_1}[\cdot]$. // About $2^{32} \cdot \alpha_1^2$ values are remained.
Proceed the similar steps as 22-29 for the other 2 bytes $Input'_{(2\sim3)}$. // About 1 value is remained.
Check if this Δ' satisfies the other $N/n_s - 1$ groups of n_s structures;
if *there is no solution for* Δ' *remained* **then**
  return $\perp$. // The distinguishing attack is failed.
Proceed Steps $4 \sim 31$ with $V'[N][N]$ to compute the distributions on $Output'_{(0\sim3)}$ by guessing $\Gamma' = (\Gamma_{(0)}, \Gamma_{(7)}, \Gamma_{(10)}, \Gamma_{(13)})$;
For AES_{10}, there exists one solution for Γ';
For any random permutation, there is no solution for Γ'.

So the success probability of this distinguisher is about $1 - 2^{-27.08}$. The advantage to distinguish AES_{10} from random permutation is not negligible. The time complexity of Steps $2 \sim 3$ is $N \times N = 2^{55.84}$ full round encryptions. Then the time complexity of Steps $4 \sim 9$ is about $2^{16} \times n_s \times N = 2^{51.92}$ memory accesses (MA). Steps $10 \sim 14$ take $2^{16} \times 2^8 \times n_s \times 2^{24} = 2^{56}$ MA, and Steps $15 \sim 21$ require about $2^{32} \times n_s \times 2^{16} = 2^{56}$ MA. Since $\alpha_1 = 2^{-10.51}$, Steps $22 \sim 29$ take $2^{32} \times \alpha_1 \times n_s \times N = 2^{57.41}$ MA and Step 30 needs about $(2^{32} \times \alpha_1^2 + 2^{32} \times \alpha_1^3) \times n_s \times N \approx 2^{46.91}$ MA. After one filter process, the number of candidates for Δ is about 1. Consequently by filtering with other $N/n_s - 1$ groups of structures, the time complexity of Step 31 is $(N/n_s - 1) \times n_s \times N \approx 2^{55.84}$ MA. Then if we roughly set one access to a table is equivalent to one full round encryption, the total complexity from Step $4 \sim 31$ is about $2^{51.92} + 2^{56} + 2^{56} + 2^{57.41} + 2^{46.91} + 2^{55.84} \approx 2^{58.49}$ encryptions. Since Step 34 also takes $2^{58.49}$ encryptions, the total time complexity of the whole attack is about $2^{55.84} + 2 \times 2^{58.49} \approx 2^{59.60}$ full round encryptions. In addition, the dominant memory requirements happen on $V[N][N]$ and $V'[N][N]$, which need about $2 \times 4 \times N \times N = 2^{58.84}$ bytes.

6 Conclusion

In this paper, we propose a statistical integral distinguisher with multiple structures on input and multiple integral properties on output based the work of Wang *et al.* at FSE'16. With this distinguisher, we give the known-key distinguishing attack on 8-round and full round AES-128 based on the Gilbert's work at ASIACRYPT'14, which are the best known-key distinguishers for AES so far according to the time complexity. Beside that, we present a secret-key statistical integral distinguisher on 5-round AES with secret S-box under chosen-ciphertext mode. This is the best integral distinguisher on AES with secret S-box under secret-key setting. As a future work, we try to apply more statistical techniques into the field of symmetric ciphers and find improved attack on AES and AES-like ciphers.

Acknowledgement. This work has been supported by 973 Program (No. 2013CB834205), NSFC Projects (No. 61133013, No. 61572293), Program for New Century Excellent Talents in University of China (NCET-13-0350), Program from Science and Technology on Communication Security Laboratory of China (No. 9140c110207150c11050).

A Appendix

A.1 Experiment Results

In order to verify the theoretical model of statistical integral distinguisher in Sect. 3, we implement the distinguishing attack in Sect. 5 on a mini variant of AES with the block size 64-bit denoted as AES* here. The round function of AES* is similar to that of AES, including four operations, *i.e.,* SB, SR, MC and

AK. 64-bit block is partitioned into 16 nibbles and SB uses S-box S_0 in LBlock. SR is same as that of AES, and the matrix used in MC is

$$M = \begin{pmatrix} 1\,1\,4\,9 \\ 9\,1\,1\,4 \\ 4\,9\,1\,1 \\ 1\,4\,9\,1 \end{pmatrix},$$

which is defined over $GF(2^4)$. For the multiplication, each nibble and value in M are considered as a polynomial over $GF(2)$ and then the nibble is multiplied modulo $x^4 + x + 1$ by the value in M. The addition is simply XOR operation. The subkeys are XORed with the nibbles in AK operation.

There is similar known-key integral distinguisher for 8-round AES* since its similarity to AES, see Fig. 1. Given a set of data $\mathcal{Z} = \{(x, 0, 0, 0) \oplus R(y, 0, 0, 0) | x \in (0, 1)^{16}\}$ for fixed y, *i.e.,* the first column of $\mathcal{Z}$ takes all 2^{16} possible values and other columns are fixed to some constants, after $S \diamond R \diamond S$ operation, each column of output u is active, i.e. that 2^{16} values are uniformly distributed on each column of output. Since $R^{-1}(\mathcal{Z}) = \{R^{-1}((x, 0, 0, 0) \oplus (y, 0, 0, 0))\}$ has 2^{16} structures that each one takes all 2^{16} possible values on the first columns and constants on other columns, after $(S \diamond R \diamond S)^{-1}$ operation, each column of output u is active.

In our experiment, we consider the distributions of four 8-bit values in v including the first and second nibble in each column of v. Here $s = 16, t = 8$ and $b = 4$. If we set $\alpha_0 = 0.2$ and take different values for N and N_s, α_1 and τ can be computed using Eq. (8). By randomly choosing N_s values for y and N values for x, we proceed the experiment to compute the statistics C' for AES* and random permutations. With 2000 times of experiments, we can obtain the empirical error probabilities $\widehat{\alpha_0}$ and $\widehat{\alpha_1}$. The experimental results for $\widehat{\alpha_0}$ and $\widehat{\alpha_1}$ are compared with the theoretical values α_0 and α_1 in Fig. 4.

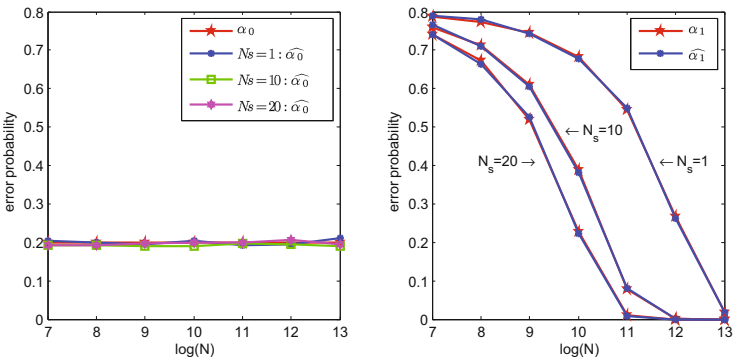

Fig. 4. Experimental results for AES* considering four input bytes. In detail, set the value of α_0 and change the values of N and N_s, the theoretical and empirical α_0 are shown in the left part of figure, corresponding α_1 calculated and tested by Eq. (5) are shown in the right part of figure.

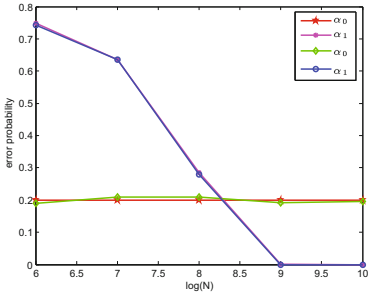

Fig. 5. Experimental results for AES* considering two input and output bytes. In detail, set the theoretical $\alpha_0 = 0.2$ and change the values of N, then the corresponding theoretical α_1 and empirical α_0 and α_1 are calculated and tested by Eq. (5) in this figure

Moreover, we implement the second experiment where we set $b = 4$ including two bytes of u and two bytes of v. We set $\alpha_0 = 0.2$ and let $N = N_s$, the empirical error probabilities are obtained from 1000 times of experiments. The experimental results for $\widehat{\alpha_0}$ and $\widehat{\alpha_1}$ are compared with the theoretical values α_0 and α_1 in Fig. 5.

Figures 4 and 5 show that the test results for the error probabilities are in good accordance with those for theoretical model.

References

1. Aoki, K.: A middletext distinguisher for full CLEFIA-128. In: 2012 International Symposium on Information Theory and its Applications (ISITA), pp. 521–525. IEEE (2012)
2. Aumasson, J., Meier, W.: Zero-sum distinguishers for reduced keccak-f and for the core functions of luffa and hamsi, 2009. Presented at the rump session of Cryptographic Hardware and Embedded Systems- CHES (2009)
3. Barreto, P.S.L.M., Rijmen, V.: Whirlpool. In: van Tilborg, H.C.A., Jajodia, S. (eds.) Encyclopedia of Cryptography and Security, 2nd edn, pp. 1384–1385. Springer, New York (2011)
4. Biryukov, A., Khovratovich, D., Nikolić, I.: Distinguisher and related-key attack on the full AES-256. In: Halevi, S. (ed.) CRYPTO 2009. LNCS, vol. 5677, pp. 231–249. Springer, Heidelberg (2009). doi:10.1007/978-3-642-03356-8_14
5. Blondeau, C., Peyrin, T., Wang, L.: Known-key distinguisher on full PRESENT. In: Gennaro, R., Robshaw, M. (eds.) CRYPTO 2015. LNCS, vol. 9215, pp. 455–474. Springer, Heidelberg (2015). doi:10.1007/978-3-662-47989-6_22
6. CAESAR: Competition for Authenticated Encryption: Security, Applicability, and Robustness. https://competitions.cr.yp.to/caesar.html
7. Daemen, J., Knudsen, L., Rijmen, V.: The block cipher square. In: Biham, E. (ed.) FSE 1997. LNCS, vol. 1267, pp. 149–165. Springer, Heidelberg (1997). doi:10.1007/BFb0052343
8. FIPS 197. Advanced Encryption Standard. Federal Information Processing Standards Publication 197, U.S. Department of Commerce/N.I.S.T (2001)

9. Grassi, L., Rechberger, C., Rønjom, S.: Subspace trail cryptanalysis and its applications to AES - extended version. https://eprint.iacr.org/2016/592
10. Grassi, L., Rechberger, C., Rønjom, S.: A new structural-differential property of 5-round AES. https://eprint.iacr.org/2017/118.pdf
11. Gilbert, H., Peyrin, T.: Super-Sbox cryptanalysis: improved attacks for AES-like permutations. In: Hong, S., Iwata, T. (eds.) FSE 2010. LNCS, vol. 6147, pp. 365–383. Springer, Heidelberg (2010). doi:10.1007/978-3-642-13858-4_21
12. Gilbert, H.: A simplified representation of AES. In: Sarkar, P., Iwata, T. (eds.) ASIACRYPT 2014. LNCS, vol. 8873, pp. 200–222. Springer, Heidelberg (2014). doi:10.1007/978-3-662-45611-8_11
13. Guo, J., Peyrin, T., Poschmann, A.: The PHOTON family of lightweight hash functions. In: Rogaway, P. (ed.) CRYPTO 2011. LNCS, vol. 6841, pp. 222–239. Springer, Heidelberg (2011). doi:10.1007/978-3-642-22792-9_13
14. Jean, J., Naya-Plasencia, M., Peyrin, T.: Multiple limited-birthday distinguishers and applications. In: Lange, T., Lauter, K., Lisoněk, P. (eds.) SAC 2013. LNCS, vol. 8282, pp. 533–550. Springer, Heidelberg (2014). doi:10.1007/978-3-662-43414-7_27
15. Knudsen, L., Rijmen, V.: Known-key distinguishers for some block ciphers. In: Kurosawa, K. (ed.) ASIACRYPT 2007. LNCS, vol. 4833, pp. 315–324. Springer, Heidelberg (2007). doi:10.1007/978-3-540-76900-2_19
16. Knudsen, L., Wagner, D.: Integral cryptanalysis. In: Daemen, J., Rijmen, V. (eds.) FSE 2002. LNCS, vol. 2365, pp. 112–127. Springer, Heidelberg (2002). doi:10.1007/3-540-45661-9_9
17. Lamberger, M., Mendel, F., Rechberger, C., Rijmen, V., Schläffer, M.: Rebound distinguishers: results on the full whirlpool compression function. In: Matsui, M. (ed.) ASIACRYPT 2009. LNCS, vol. 5912, pp. 126–143. Springer, Heidelberg (2009). doi:10.1007/978-3-642-10366-7_8
18. Lamberger, M., Mendel, F., Rechberger, C., Rijmen, V., Schläffer, M.: The rebound attack and subspace distinguishers: application to whirlpool. Cryptology ePrint Archive, Report 2010/198 (2010)
19. Mendel, F., Peyrin, T., Rechberger, C., Schläffer, M.: Improved cryptanalysis of the reduced Grøstl compression function, ECHO permutation and AES block cipher. In: Jacobson, M.J., Rijmen, V., Safavi-Naini, R. (eds.) SAC 2009. LNCS, vol. 5867, pp. 16–35. Springer, Heidelberg (2009). doi:10.1007/978-3-642-05445-7_2
20. Minematsu, K.: AES-OTR (v3.1). https://competitions.cr.yp.to/round3/aesotrv31.pdf
21. Minier, M., Phan, R.C.-W., Pousse, B.: Distinguishers for ciphers and known key attack against Rijndael with large blocks. In: Preneel, B. (ed.) AFRICACRYPT 2009. LNCS, vol. 5580, pp. 60–76. Springer, Heidelberg (2009). doi:10.1007/978-3-642-02384-2_5
22. 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: Gennaro, R., Robshaw, M. (eds.) CRYPTO 2015. LNCS, vol. 9215, pp. 95–115. Springer, Heidelberg (2015). doi:10.1007/978-3-662-47989-6_5
23. Sun, B., Liu, M., Guo, J., Qu, L., Rijmen, V.: New insights on AES-like SPN ciphers. In: Robshaw, M., Katz, J. (eds.) CRYPTO 2016. LNCS, vol. 9814, pp. 605–624. Springer, Heidelberg (2016). doi:10.1007/978-3-662-53018-4_22
24. Wang, M., Cui, T., Chen, H., Sun, L., Wen, L., Bogdanov, A.: Integrals go statistical: cryptanalysis of full Skipjack variants. In: Peyrin, T. (ed.) FSE 2016. LNCS, vol. 9783, pp. 399–415. Springer, Heidelberg (2016). doi:10.1007/978-3-662-52993-5_20