Des. Codes Cryptogr. (2018) 86:1541–1552
https://doi.org/10.1007/s10623-017-0411-z

Upper bound of the length of truncated impossible differentials for AES

Qian Wang[1] · **Chenhui Jin**[1]

Received: 25 January 2017 / Revised: 29 August 2017 / Accepted: 30 August 2017 /
Published online: 7 September 2017

Abstract On the provable security of a block cipher against impossible differential cryptanalysis, the maximal length of impossible differentials is an essential aspect. Most previous work on finding impossible differentials for AES, omits the non-linear component (S-box), which is important for the security. In EUROCRYPT 2016, Sun et al. showed how to bound the length of impossible differentials of a SPN "structure" using the primitive index of its linear layer. They proved that there do not exist impossible differentials longer than four rounds for the AES "structure", instead of the AES cipher. Since they do not consider the details of the S-box, their bound is not feasible for a concrete cipher. With their result, the upper bound of the length of impossible differentials for AES, is still unknown. We fill this gap in our paper. By revealing some important properties of the AES S-box, we further prove that even though the details of the S-box are considered, there do not exist truncated impossible differentials covering more than four rounds for AES, under the assumption that round keys are independent and uniformly random. Specially, even though the details of the S-box and key schedule are both considered, there do not exist truncated impossible differentials covering more than four rounds for AES-256.

Keywords Truncated impossible differential · AES · S-box · Linear space

Mathematics Subject Classification 94A60

1 Introduction

Since Rijndael was selected as the Advanced Encryption Standard (AES) [8], its security has been widely studied by many cryptanalysts. Differential cryptanalysis [5] and linear crypt-

Communicated by C. Carlet.

✉ Qian Wang
e_alpha@163.com

[1] Information Science and Technology Institute, Zhengzhou, China

analysis [19] are among the most powerful cryptanalysis tools, against which the AES adopts "wide trail strategy [9]" to provide elegant provable security. Differential attack makes use of differential characteristics with high probability to sieve right keys. By counting the number of differential active S-boxes (S-box with non-zero input difference) in a differential characteristic covering several rounds, we could give upper bound on the differential probability of this differential characteristic. Things are similar for the linear attack.

Unlike differential attack, impossible differential attack [3,15] utilizes impossible differentials, namely differentials with probability 0, to discard wrong keys. Among these impossible differentials, truncated impossible differentials [13,16] attract much attention. For AES, there are many four-round truncated impossible differentials being constructed [1,4,6,22]. In [1,4,6,22], some four-round truncated impossible differentials, which are the longest truncated impossible differentials found so far for AES, are used to mount impossible differential attack on (seven-round) AES-128. Therefore, the maximal length of impossible differentials for a cipher, are estimated to evaluate the resistance of this cipher against impossible differential cryptanalysis.

Although many approaches (like U-method [14], *UID*-method [18], *WW*-method [26] and *Cui*-method [7]) have been proposed to find impossible differentials for SPN block ciphers, they do not take advantage of the properties of the non-linear substitution layer (e.g., S-box), which might result in tighter bound on the length of impossible differentials.

In EUROCRYPT 2016, Sun et al. [24] utilized the "primitive index" [7,24] of the linear layer of SPN "structure", to bound the length of impossible differentials. They proved that there do not exist impossible differentials covering more than four rounds for AES "structure". Their target is not the concrete cipher, but the "structure" [24,25]:

Definition 1 [24] Let $E : F_2^n \to F_2^n$ be a block cipher with bijective S-boxes as the basic non-linear components.

1. A structure ε^E on F_2^n is defined as a set of block ciphers E' which is exactly the same as E except that the S-boxes can take all possible bijective transformations on the corresponding domains.
2. Let $a, b \in F_2^n$. If for any $E' \in \varepsilon^E$, $a \to b$ is an impossible differential of E', $a \to b$ is called an impossible differential of ε^E.

With their method, we are able to get the upper bound of the length of impossible differentials for a "structure". However, the concept of "structure" is far from practical cipher, since S-boxes are often fixed, or varying within a much small subspace of the space of all bijective S-boxes (when viewing AddRoundKey transformation and fixed S-boxes into variable S-boxes controlled by keys) in a concrete cipher. Their method cannot prove that there do not exist impossible differentials covering more than four rounds for AES.

These existing methods do not exploit the properties of the non-linear S-box transformation. Therefore, finding a method, that exploits the properties of the S-boxes and gives practical maximal length of impossible differentials for SPN ciphers (such as AES and AES-like ciphers), is a problem worth further investigating.

Relation to the five-round impossible differential distinguisher proposed by Grassi et al. [11]. As we mentioned before, four-round "key-independent" impossible differentials are used to mount impossible differential attack on round-reduced AES (typically on seven-round AES-128). Such a four-round key-independent impossible differentials can be used to mount key-recovery attacks on five-round AES, namely, five-round AES can be distinguished from a random permutation on knowing some key bytes of the first or last round. In [23], a five-round "key-dependent" zero-correlation linear hull (it can also be seen as an impossible

differential) was derived from a four-round "key-independent" zero-correlation linear hull, on knowing the difference of two key bytes. This "key-dependent" distinguishing property (in other words, five-round key-recovery attack) was further turn into a five-round integral distinguisher on AES by Sun et al. [23]. Later, such five-round "key-dependent" distinguisher property was further explained with "subspace trails" notation and greatly improved by Grassi et al. [11]. In [11], Grassi et al. set up an impossible differential attack on five rounds of AES to recover the difference of two key bytes, based on a four-round "key-independent" impossible differential. Furthermore, they turned this five-round key recovery attack into a five-round (subspace) impossible differential distinguisher for AES. Therefore, both the five-round zero-correlation linear hull in [23] and five-round impossible differential in [11] are much different from the usual "key-independent" zero-correlation linear hulls or impossible differentials that are used in the zero-correlation linear attack or the impossible differential attack. In our paper, we consider the maximal length of the core "key-independent" impossible differentials, that are used in the impossible differential attacks and their [11,23] five-round distinguisher on AES.

Our contributions By revealing some important properties of the AES S-box, we prove that there do not exist truncated impossible differentials covering more than four rounds for AES with the details of S-boxes considered (stronger than the "structure" as [24]), under practical assumption that round keys are independent and uniformly random. Specially, for AES-256, even though the S-boxes and the key schedule are both taken into consideration, there do not exist truncated impossible differentials covering more than four rounds.

Outline In Sect. 2, we give some preliminaries to make our representation clear and understandable. In Sect. 3, we reveal an important property of the AES S-box, and then attain the longest rounds of truncated impossible differentials for AES. Whole paper is concluded in Sect. 4.

2 Preliminaries

2.1 Notations

$\oplus$: bitwise XOR (it is coincided with the addition in the finite field F_{2^8} in our representation).
$a \vee b$: bitwise OR of binary vectors a and b.
$|A|$ or $\#\{A\}$: the number of elements in the set A.
X : matrix defined over a finite field, $X = (x_{i,j})_{4\times 4}$, $x_{i,j} \in F_{2^8}$.
$x_{\bullet,j}$: the jth column of the state matrix X.
ΔX : the difference of X.
$\Delta \tilde{X}$: the truncated difference of X.
$\circ$: composition of transformations.
f, f^{-1} : transformation f and its inverse.

2.2 Some definitions

AES is an SPN block cipher with block size 128 bits. These 128 bits are organized as a 4×4 matrix of bytes, denoted by $X = (x_{i,j})_{4\times 4}$, $x_{i,j} \in F_{2^8}$, as Fig. 1.

The round function of AES is composed of four transformations [8,20]:

Fig. 1 AES state

$x_{0,0}$	$x_{0,1}$	$x_{0,2}$	$x_{0,3}$
$x_{1,0}$	$x_{1,1}$	$x_{1,2}$	$x_{1,3}$
$x_{2,0}$	$x_{2,1}$	$x_{2,2}$	$x_{2,3}$
$x_{3,0}$	$x_{3,1}$	$x_{3,2}$	$x_{3,3}$

1. SubBytes (*SB*): apply the same 8-bit invertible S-box to 16 bytes in parallel. The S-box is composed of the multiplicative inverse function in F_{2^8} and an invertible affine transformation, denoted by $s(x) = A \cdot x^{-1} \oplus b$.
2. ShiftRows (*SR*): shift the ith row by i bytes to the left circularly.
3. MixColumns (*MC*): left multiplication by a 4×4 MDS matrix with elements in F_{2^8}. This matrix and its inverse are defined as:

$$C = \begin{bmatrix} 2 & 3 & 1 & 1 \\ 1 & 2 & 3 & 1 \\ 1 & 1 & 2 & 3 \\ 3 & 1 & 1 & 2 \end{bmatrix}, C^{-1} = \begin{bmatrix} e & b & d & 9 \\ 9 & e & b & d \\ d & 9 & e & b \\ b & d & 9 & e \end{bmatrix}.$$

4. AddRoundKey (AK_{K^i}): bitwise exclusive-or with 128-bit round key K^i. K^i denotes the subkey of the ith round, and K^0 denotes the whiten key before the first round.

The ith full round transformation of AES can be written as $AK_{K^i} \circ MC \circ SR \circ SB$. Since *SR* and *MC* are linear, we could change the order of *SR*, *MC* and AK_{K^i} by XORing equivalent subkeys before *MC* or *SR*, denoted by $MC \circ AK_{MC^{-1}(K^i)} \circ SR \circ SB$ and $MC \circ SR \circ AK_{SR^{-1} \circ MC^{-1}(K^i)} \circ SB$ respectively. Here, $MC^{-1}(K^i)$ and $SR^{-1} \circ MC^{-1}(K^i)$ denote the equivalent subkeys at differential places respectively.

We wouldn't introduce the details of the key schedule of AES, since our main result is under the assumption that its round keys are independent and uniformly random.

Definition 2 For function $f : F_2^n \mapsto F_2^n$, its differential probability for the differential $\Delta x \to \Delta y$ is defined as:

$$P(\Delta x \xrightarrow{f} \Delta y) = \frac{\#\{x | f(x \oplus \Delta x) \oplus f(x) = \Delta y\}}{2^n}.$$

Definition 3 For a keyed function $f_k : F_2^n \mapsto F_2^n, k \in K$, its differential probability for differential $\Delta x \to \Delta y$ is defined as:

$$P(\Delta x \to \Delta y) = \frac{1}{|K|} \sum_{k \in K} P\left(\Delta x \xrightarrow{f_k} \Delta y\right).$$

Definition 4 (*Pattern*) Let $X = (x_1, x_2, \ldots, x_n) \in F_{2^m}^n$, then the pattern of X is defined as:

$$\chi(X) = (y_1, y_2, \ldots, y_n) \in F_2^n,$$

where $y_i = 0$ if $x_i = 0$, and $y_i = 1$ otherwise.

To characterize the diffusion of differences through the MixColumns transformation, we define the operation of *MC* on the difference patterns (truncated differences) as:

Definition 5 For the MixColumns transformation MC of AES, its operation on the state pattern (or truncated difference) $\tilde{X} = (\tilde{x}_{i,j})_{4\times 4}$, is defined as: let state pattern $\tilde{Y} = MC(\tilde{X})$, then the $j^{th}(0 \le j \le 3)$ column of $\tilde{Y}$ is calculated as $\tilde{y}_{\bullet,j} = MC(\tilde{x}_{\bullet,j}) = (\tilde{x}_{0,j}\tilde{c}_{\bullet,0}) \vee (\tilde{x}_{1,j}\tilde{c}_{\bullet,1}) \vee (\tilde{x}_{2,j}\tilde{c}_{\bullet,2}) \vee (\tilde{x}_{3,j}\tilde{c}_{\bullet,3})$, where $\tilde{c}_{\bullet,i} = \chi(c_{\bullet,i})$ and $c_{\bullet,i}$ denotes the $i^{th}(0 \le i \le 3)$ column of the MixColumns matrix C. Similar for MC^{-1}.

Definition 6 [13] For function $f : F_{2^m}^n \mapsto F_{2^m}^n$, its truncated differential probability for truncated differential $\Delta\tilde{X} \to \Delta\tilde{Y}$ is defined as:

$$P(\Delta\tilde{X} \xrightarrow{f} \Delta\tilde{Y}) = \frac{\sum_{\chi(\Delta X)=\Delta\tilde{X}, \chi(\Delta Y)=\Delta\tilde{Y}} P(\Delta X \xrightarrow{f} \Delta Y)}{\#\left\{\Delta X | \chi(\Delta X) = \Delta\tilde{X}\right\}}.$$

Definition 7 For a keyed function $f_k : F_{2^m}^n \mapsto F_{2^m}^n, k \in K$, its truncated differential probability for truncated differential $\Delta\tilde{X} \to \Delta\tilde{Y}$ is defined as:

$$P(\Delta\tilde{X} \xrightarrow{f_k,K} \Delta\tilde{Y}) = \frac{1}{|K|}\sum_{k\in K} P(\Delta\tilde{X} \xrightarrow{f_k} \Delta\tilde{Y}).$$

Then for a cipher E_k with key space K, a truncated differential $\Delta\tilde{X} \to \Delta\tilde{Y}$ is impossible if $P(\Delta\tilde{X} \xrightarrow{E_k,K} \Delta\tilde{Y}) = 0$, which means that for any $k \in K$ and any differential $\Delta X \to \Delta Y$ such that $\chi(\Delta X) = \Delta\tilde{X}$ and $\chi(\Delta Y) = \Delta\tilde{Y}$, we have $P(\Delta X \xrightarrow{E_k} \Delta Y) = 0$. However, $\Delta\tilde{X} \to \Delta\tilde{Y}$ is possible if for some $k \in K$ and differential $\Delta X \to \Delta Y$ such that $\chi(\Delta X) = \Delta\tilde{X}$ and $\chi(\Delta Y) = \Delta\tilde{Y}$, we have $P(\Delta X \xrightarrow{E_k} \Delta Y) \neq 0$.

3 Our claims on AES

To prove that the upper bound of the length of truncated impossible differentials for AES is four rounds, we would show that any nontrivial truncated differential $\Delta\tilde{X} \to \Delta\tilde{Y}$ is possible for five-round AES (without MixColumns in the last round). As stated at the end of the Section 2.1, we aim to construct a concrete differential $\Delta X \to \Delta Y$ such that $\chi(\Delta X) = \Delta\tilde{X}$ and $\chi(\Delta Y) = \Delta\tilde{Y}$, as well as a series of round keys (we make the assumption that round keys are independent and uniformly random), which make that $P(\Delta X \xrightarrow{5\_round\_AES} \Delta Y) \neq 0$.

Firstly, we need to investigate some details about the AES S-box in Sect. 3.1.

3.1 Some properties of the AES S-box

In this section, we investigate the properties of the AES S-box. The AES S-box is composed of two functions: the multiplicative inverse function in F_{2^8}, denoted by x^{-1}, and an invertible affine transformation, denoted by $L(x) = A \cdot x \oplus b$. Then, the AES S-box is described as:

$$s(x) = L(x^{-1}) = A \cdot x^{-1} \oplus b, \ with\ 0^{-1} = 0,$$

here,

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}, b = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

We mainly focus on the properties of the multiplicative inverse function in F_{2^8}. Then, a nontrivial differential $\alpha \to \beta$ is possible for this function if for some x,

$$(x \oplus \alpha)^{-1} \oplus x^{-1} = \beta. \tag{1}$$

For Eq. (1), when the output difference β is fixed, the set of all the possible input difference α possesses a special property, deduced by Daemen et al. [10]. This property can be expressed as the following lemma:

Lemma 1 *For any nonzero output difference β of the multiplicative inverse function in F_{2^8}, the set of the multiplicative inverse elements of all the possible input differences α (add 0 element) forms a 7-dimensional linear space over F_2. Namely, $V_\beta = \{\alpha^{-1} | (x \oplus \alpha)^{-1} \oplus (x)^{-1} = \beta, x \in F_{2^8}\} \bigcup \{0\}$ is a 7-dimensional linear space over F_2.*

Proof This fact has been proved by Daemen et al. [10]. □

Based on the linear property of the multiplicative inverse function in Lemma 1, we can get the following essential property of the AES S-box, which plays an important role in the construction of our differential trails.

Theorem 1 *For all $(\beta_0, \beta_1, \beta_2, \beta_3) \in [F_{2^8}^*]^4$, $(c_0, c_1, c_2, c_3) \in [F_{2^8}^*]^4$, there exist $(k_0, k_1, k_2, k_3) \in F_{2^8}^4$ and $\alpha \in F_{2^8}^*$, satisfying the following system of equations:*

$$\begin{cases} \beta_0 = s(c_0\alpha \oplus k_0) \oplus s(k_0) \\ \beta_1 = s(c_1\alpha \oplus k_1) \oplus s(k_1) \\ \beta_2 = s(c_2\alpha \oplus k_2) \oplus s(k_2) \\ \beta_3 = s(c_3\alpha \oplus k_3) \oplus s(k_3) \end{cases}, \tag{*}$$

where s denotes the AES S-box, $s(x) = A \cdot x^{-1} \oplus b$ and $0^{-1} = 0$.

Proof By peeling off the bijective affine transformation in AES S-box s, we get that the system (*) is equivalent to the following one

$$\begin{cases} c_0(A^{-1}\beta_0) = (\alpha \oplus c_0^{-1}k_0)^{-1} \oplus (c_0^{-1}k_0)^{-1} \\ c_1(A^{-1}\beta_1) = (\alpha \oplus c_1^{-1}k_1)^{-1} \oplus (c_1^{-1}k_0)^{-1} \\ c_2(A^{-1}\beta_2) = (\alpha \oplus c_2^{-1}k_2)^{-1} \oplus (c_2^{-1}k_0)^{-1} \\ c_3(A^{-1}\beta_3) = (\alpha \oplus c_3^{-1}k_3)^{-1} \oplus (c_3^{-1}k_0)^{-1} \end{cases} \tag{**}$$

By Lemma 1, we know that $V_{c_i(A^{-1}\beta_i)} (0 \le i \le 3)$ are all 7-dimensional linear space over F_2. Since F_{2^8} is an 8-dimensional linear space over F_2, we have $\left|V_{c_0(A^{-1}\beta_0)} \bigcap V_{c_1(A^{-1}\beta_1)} \bigcap V_{c_2(A^{-1}\beta_2)} \bigcap V_{c_3(A^{-1}\beta_3)}\right| \ge 2^4$. Take nonzero $\alpha^{-1} \in V_{c_0(A^{-1}\beta_0)}$

$\bigcap V_{c_1(A^{-1}\beta_1)} \bigcap V_{c_2(A^{-1}\beta_2)} \bigcap V_{c_3(A^{-1}\beta_3)}$, then there exists corresponding $x_i (0 \le i \le 3)$ such that $c_i(A^{-1}\beta_i) = (\alpha \oplus x_i)^{-1} \oplus (x_i)^{-1} (0 \le i \le 3)$. Take $k_i = c_i x_i (0 \le i \le 3)$. Thus, resulting $\alpha \in F_{2^8}^*$ and $(k_0, k_1, k_2, k_3) \in F_{2^8}^4$ satisfy the system (**), also the system (*). □

3.2 Bound the length of truncated impossible differentials for AES

To prove that any nontrivial truncated differential $\Delta\tilde{X}^0 \to \Delta\tilde{X}^5$ is possible for five-round AES (without MixColumns in the last round), the main idea of our proof is: for any nontrivial truncated differential $\Delta\tilde{X}^0 \to \Delta\tilde{X}^5$, construct a five-round truncated differential trail, $\Delta\tilde{X}^0 \to \Delta\tilde{X}^1 \to \Delta\tilde{X}^2 \to \Delta\tilde{X}^3 \to \Delta\tilde{X}^4 \to \Delta\tilde{X}^5$, following from Sun's idea "meet-in-the-middle" in [24] (this procedure is shown with a special example in Fig. 2); along with the construction of this truncated differential trail, a concrete differential trail, $\Delta X^0 \to \Delta X^1 \to \Delta X^2 \to \Delta X^3 \to \Delta X^4 \to \Delta X^5$ such that $\chi(\Delta X^i) = \Delta\tilde{X}^i$, is constructed, whose existence heavily depends on the algebraic property of the AES S-box in Theorem 1.

To construct the differential trail as we want, we need to have a good knowledge of the difference propagation of the AES round function. Combining the properties of the AES S-box and its linear layer, we get the following lemmas on the differential properties of the AES round function.

Lemma 2 *For any nonzero truncated difference $\Delta\tilde{X}$ and concrete difference ΔY of the AES state such that $\chi(\Delta Y) = MC(\Delta\tilde{X})$, there exist ΔX and X, such that $\chi(\Delta X) = \Delta\tilde{X}$, satisfying $SB \circ MC(X \oplus \Delta X) \oplus SB \circ MC(X) = \Delta Y$.*

Proof Let $\Delta Y^1 = MC(\Delta X)$ and $Y^1 = MC(X)$. Then the $j^{th} (0 \le j \le 3)$ column of ΔY^1 is represented as

$$\begin{cases} \Delta y^1_{0,j} = 2\Delta x_{0,j} \oplus 3\Delta x_{1,j} \oplus \Delta x_{2,j} \oplus \Delta x_{3,j} \\ \Delta y^1_{1,j} = \Delta x_{0,j} \oplus 2\Delta x_{1,j} \oplus 3\Delta x_{2,j} \oplus \Delta x_{3,j} \\ \Delta y^1_{2,j} = \Delta x_{0,j} \oplus \Delta x_{1,j} \oplus 2\Delta x_{2,j} \oplus 3\Delta x_{3,j} \\ \Delta y^1_{3,j} = 3\Delta x_{0,j} \oplus \Delta x_{1,j} \oplus \Delta x_{2,j} \oplus 2\Delta x_{3,j} \end{cases}$$

If $\Delta\tilde{x}_{\cdot,j} = 0$, then $\Delta x_{\cdot,j} = 0$ and $\Delta y_{\cdot,j} = MC(\Delta x_{\cdot,j}) = 0$. Naturally, $x_{\cdot,j}$ can be arbitrary.

Therefore, we only need to consider those nonzero columns of $\Delta\tilde{x}$, i.e. those $\Delta\tilde{x}_{\cdot,j} \neq 0$. Suppose that $\Delta x_{i,j} = c_{i,j}\alpha_j (0 \le i \le 3)$. Here, $c_{i,j} (0 \le i \le 3)$ are taken as the following steps:

1. If $\Delta\tilde{x}_{0,j} \neq 0$, take nonzero $c_{0,j}$ then $(2c_{0,j}, c_{0,j}, c_{0,j}, 3c_{0,j})$ is all-nonzero, else take $c_{0,j} = 0$;
2. If $\Delta\tilde{x}_{1,j} \neq 0$, take nonzero $c_{1,j}$ such that $(2c_{0,j} \oplus 3c_{1,j}, c_{0,j} \oplus 2c_{1,j}, c_{0,j} \oplus c_{1,j}, 3c_{0,j} \oplus c_{1,j})$ is all-nonzero, else take $c_{1,j} = 0$;
3. If $\Delta\tilde{x}_{2,j} \neq 0$, take nonzero $c_{2,j}$ such that $(2c_{0,j} \oplus 3c_{1,j} \oplus c_{2,j}, c_{0,j} \oplus 2c_{1,j} \oplus 3c_{2,j}, c_{0,j} \oplus c_{1,j} \oplus 2c_{2,j}, 3c_{0,j} \oplus c_{1,j} \oplus c_{2,j})$ is all-nonzero, else take $c_{2,j} = 0$;
4. If $\Delta\tilde{x}_{3,j} \neq 0$, take nonzero $c_{3,j}$ such that $(2c_{0,j} \oplus 3c_{1,j} \oplus c_{2,j} \oplus c_{3,j}, c_{0,j} \oplus 2c_{1,j} \oplus 3c_{2,j} \oplus c_{3,j}, c_{0,j} \oplus c_{1,j} \oplus 2c_{2,j} \oplus 3c_{3,j}, 3c_{0,j} \oplus c_{1,j} \oplus c_{2,j} \oplus 2c_{3,j})$ is all-nonzero, else take $c_{3,j} = 0$.

With above values of $c_{0,j}, c_{1,j}, c_{2,j}$ and $c_{3,j}$, we can get

$$\begin{cases} \Delta y^1_{0,j} = (2c_{0,j} \oplus 3c_{1,j} \oplus c_{2,j} \oplus c_{3,j})\alpha_j = c'_{0,j}\alpha_j \\ \Delta y^1_{1,j} = (c_{0,j} \oplus 2c_{1,j} \oplus 3c_{2,j} \oplus c_{3,j})\alpha_j = c'_{1,j}\alpha_j \\ \Delta y^1_{2,j} = (c_{0,j} \oplus c_{1,j} \oplus 2c_{2,j} \oplus 3c_{3,j})\alpha_j = c'_{2,j}\alpha_j \\ \Delta y^1_{3,j} = (3c_{0,j} \oplus c_{1,j} \oplus c_{2,j} \oplus 2c_{3,j})\alpha_j = c'_{3,j}\alpha_j \end{cases}, \ where\ c'_{i,j} \neq 0 (0 \leq i \leq 3).$$

Since $\Delta Y = SB(\Delta Y^1) \oplus SB(\Delta Y^1 \oplus Y^1)$,

$$\begin{cases} \Delta y_{0,j} = s\left(c'_{0,j}\alpha_j \oplus y^1_{0,j}\right) \oplus s\left(y^1_{0,j}\right) \\ \Delta y_{1,j} = s\left(c'_{1,j}\alpha_j \oplus y^1_{1,j}\right) \oplus s\left(y^1_{1,j}\right) \\ \Delta y_{2,j} = s\left(c'_{2,j}\alpha_j \oplus y^1_{2,j}\right) \oplus s\left(y^1_{2,j}\right) \\ \Delta y_{3,j} = s\left(c'_{3,j}\alpha_j \oplus y^1_{3,j}\right) \oplus s\left(y^1_{3,j}\right) \end{cases}.$$

By Theorem 1, we know that for all $(\Delta y_{0,0}, \Delta y_{1,0}, \Delta y_{2,0}, \Delta y_{3,0}) \in [F^*_{2^8}]^4$, above system has solutions on variables α_j and $y^1_{:,j}$. We get the jth column of ΔX by taking $\Delta x_{0,j} = c_{0,j}\alpha_j$, $\Delta x_{1,j} = c_{1,j}\alpha_j$, $\Delta x_{2,j} = c_{2,j}\alpha_j$, $\Delta x_{3,j} = c_{3,j}\alpha_j$, and get the jth column of X by $X = MC^{-1}(Y^1)$. Other nonzero columns of ΔX and corresponding columns of X are similarly calculated.

Then, resulting ΔX and X satisfies $SB \circ MC(X \oplus \Delta X) \oplus SB \circ MC(X) = \Delta Y$. □

Lemma 3 *For any nonzero concrete difference ΔX and truncated difference $\Delta\tilde{Y}$ of the AES state such that $\Delta\tilde{Y} = MC^{-1}(\chi(\Delta X))$, there exist ΔY and X, such that $\chi(\Delta Y) = \Delta\tilde{Y}$, satisfying $MC^{-1} \circ SB^{-1}(\Delta X) \oplus MC^{-1} \circ SB^{-1}(\Delta X \oplus X) = \Delta Y$.*

Proof Let $\Delta X^1 = SB^{-1}(X) \oplus SB^{-1}(X \oplus \Delta X)$. Suppose that $\Delta x_{:,j} \neq 0$, then the jth column of ΔY is represented as:

$$\begin{cases} \Delta y_{0,j} = e\Delta x^1_{0,j} \oplus b\Delta x^1_{1,j} \oplus d\Delta x^1_{2,j} \oplus 9\Delta x^1_{3,j} \\ \Delta y_{1,j} = 9\Delta x^1_{0,j} \oplus e\Delta x^1_{1,j} \oplus b\Delta x^1_{2,j} \oplus d\Delta x^1_{3,j} \\ \Delta y_{2,j} = d\Delta x^1_{0,j} \oplus 9\Delta x^1_{1,j} \oplus e\Delta x^1_{2,j} \oplus b\Delta x^1_{3,j} \\ \Delta y_{3,j} = b\Delta x^1_{0,j} \oplus d\Delta x^1_{1,j} \oplus 9\Delta x^1_{2,j} \oplus e\Delta x^1_{3,j} \end{cases},$$

where $\Delta x^1_{i,j} = s^{-1}(x_{i,j}) \oplus s^{-1}(x_{i,j} \oplus \Delta x_{i,j})$, $0 \leq i \leq 3$.

For nonzero $\Delta x_{i,j}$, resulting $\Delta x^1_{i,j}$ can take $2^7 - 1$ different nonzero values when $x_{i,j}$ runs through F_{2^8} [10]. For $\Delta x_{i,j}$ equal to 0, resulting $\Delta x^1_{i,j} = 0$ obviously. Then, we are able to construct $(\Delta x^1_{0,j}, \Delta x^1_{1,j}, \Delta x^1_{2,j}, \Delta x^1_{3,j})$ such that $(\Delta y_{0,j}, \Delta y_{1,j}, \Delta y_{2,j}, \Delta y_{3,j})$ is all-nonzero, by changing $(x_{0,j}, x_{1,j}, x_{2,j}, x_{3,j})$ as the following steps:

1. If $\Delta x_{0,j} \neq 0$, take arbitrary $x_{0,j}$ then $\Delta x^1_{0,j} \neq 0$ and $(e\Delta x^1_{0,j}, 9\Delta x^1_{0,j}, d\Delta x^1_{0,j}, b\Delta x^1_{0,j})$ is all-nonzero. If $\Delta x_{0,j} = 0$, then $\Delta x^1_{0,j} = 0$ and take $x_{0,j}$ arbitrarily.
2. If $\Delta x_{1,j} \neq 0$, there are at most 4 values of $\Delta x^1_{1,j}$ which make that there is zero in $(e\Delta x^1_{0,j} \oplus b\Delta x^1_{1,j}, 9\Delta x^1_{0,j} \oplus e\Delta x^1_{1,j}, d\Delta x^1_{0,j} \oplus 9\Delta x^1_{1,j}, b\Delta x^1_{0,j} \oplus d\Delta x^1_{1,j})$. Since $\Delta x^1_{1,j}$ can take $2^7 - 1$ different nonzero values, we can easily avoid those 4 values of $\Delta x^1_{1,j}$ by changing $x_{1,j}$. Take $x_{1,j}$ such that $(e\Delta x^1_{0,j} \oplus b\Delta x^1_{1,j}, 9\Delta x^1_{0,j} \oplus e\Delta x^1_{1,j}, d\Delta x^1_{0,j} \oplus 9\Delta x^1_{1,j}, b\Delta x^1_{0,j} \oplus d\Delta x^1_{1,j})$ is all-nonzero. If $\Delta x_{1,j} = 0$, then $\Delta x^1_{1,j} = 0$ and take $x_{1,j}$ arbitrarily.

   .....

3. Step by step. Things are similar for $x_{2,j}$ and $x_{3,j}$. Eventually, we construct $(x_{0,j}, x_{1,j}, x_{2,j}, x_{3,j})$ such that $(\Delta y_{0,j}, \Delta y_{1,j}, \Delta y_{2,j}, \Delta y_{3,j})$ is all-nonzero.

For other nonzero columns of ΔX, corresponding columns of X and ΔY are similarly calculated.

For zero columns of ΔX, corresponding columns of X are arbitrary and ΔY are all-zero.

Then, resulting X and ΔY satisfying $MC^{-1} \circ SB^{-1}(\Delta X) \oplus MC^{-1} \circ SB^{-1}(\Delta X \oplus X) = \Delta Y$. □

Lemma 4 *Let $R_1 = SB \circ MC \circ SR \circ AK_K \circ SB \circ MC$. For any nonzero truncated difference $\Delta\tilde{X}$ and all-nonzero concrete difference ΔY of the AES state, there exist ΔX, K and X, such that $\chi(\Delta X) = \Delta\tilde{X}$, satisfying $R_1(X \oplus \Delta X) \oplus R_1(X) = \Delta Y$.*

Proof Let $Y^1 = SB \circ MC(X)$, $\Delta Y^1 = SB \circ MC(X) \oplus SB \circ MC(\Delta X \oplus X)$ and $Y^2 = SR \circ AK_K(Y^1)$.

Take $\Delta\tilde{Y}^1 = MC(\Delta\tilde{X})$, then $\Delta\tilde{Y}^1$ has at least one all-nonzero column by the definition of MC on a truncated difference (see Definition 5), and every column of $\Delta\tilde{Y}^2 = SR(\Delta\tilde{Y}^1)$ is nonzero.

For $\Delta\tilde{Y}^2 \xrightarrow{SB \circ MC} \Delta Y$, since every column of $\Delta\tilde{Y}^2$ is nonzero, we know that $MC(\Delta\tilde{Y}^2)$ is all-nonzero (namely, $MC(\Delta\tilde{Y}^2) = \chi(\Delta Y)$). By Lemma 2, there exist corresponding ΔY^2 and Y^2, such that $\chi(\Delta Y^2) = \Delta\tilde{Y}^2$, satisfying $\Delta Y = SB \circ MC(Y^2) \oplus SB \circ MC(Y^2 \oplus \Delta Y^2)$. Then $\Delta Y^1 = SR^{-1}(\Delta Y^2)$ and $\chi(\Delta Y^1) = \Delta\tilde{Y}^1$.

For $\Delta\tilde{X} \xrightarrow{SB \circ MC} \Delta Y^1$, since $\chi(\Delta Y^1) = \Delta\tilde{Y}^1 = MC(\Delta\tilde{X})$ as we take, by Lemma 2, there exist corresponding ΔX and X, such that $\chi(\Delta X) = \Delta\tilde{X}$, satisfying $\Delta Y^1 = SB \circ MC(X) \oplus SB \circ MC(\Delta X \oplus X)$. Take $K = Y^1 \oplus SR^{-1}(Y^2)$.

Therefore, resulting ΔX, K and X satisfying $R_1(X \oplus \Delta X) \oplus R_1(X) = \Delta Y$. □

Lemma 5 *Let $R_2 = MC^{-1} \circ SB^{-1} \circ SR^{-1} \circ MC^{-1}$. For any nonzero truncated difference $\Delta\tilde{X}$ and all-nonzero truncated difference $\Delta\tilde{Y}$ of the AES state, there exist ΔX, ΔY and X, such that $\chi(\Delta X) = \Delta\tilde{X}$ and $\chi(\Delta Y) = \Delta\tilde{Y}$, satisfying $R_2(X \oplus \Delta X) \oplus R_2(X) = \Delta Y$.*

Proof According to the proof of Lemma 3, we could take nonzero bytes of ΔX such that corresponding columns of the $MC^{-1}(\Delta X)$ are all-nonzero. Then, every column of the $\Delta X^1 = SR^{-1} \circ MC^{-1}(\Delta X)$ is nonzero.

For $\Delta X^1 \xrightarrow{MC^{-1} \circ SB^{-1}} \Delta\tilde{Y}$, since every column of the ΔX^1 is nonzero, we know that $MC^{-1}(\chi(\Delta X^1))$ is all-nonzero (namely, $MC^{-1}(\chi(\Delta X^1)) = \Delta\tilde{Y}$). By Lemma 3, there exist corresponding ΔY and X^1, such that $\chi(\Delta Y) = \Delta\tilde{Y}$, satisfying $\Delta Y = MC^{-1} \circ SB^{-1}(X^1) \oplus MC^{-1} \circ SB^{-1}(X^1 \oplus \Delta X^1)$. Take $X = MC \circ SR(X^1)$.

Therefore, resulting ΔX, ΔY and X satisfying $R_2(X \oplus \Delta X) \oplus R_2(X) = \Delta Y$. □

Based on above preparation, we are able to give our claims on AES. Note that Sun et al. [24] didn't consider the S-box, therefore, their bound is not suitable for AES (cipher). Our Theorem 2 is stronger than it, since our bound is built on the details of the AES S-box.

Theorem 2 *There do not exist truncated impossible differentials covering more than four rounds for AES with S-box considered and under the assumption that round keys are independent and uniformly random. Longer truncated impossible differentials might exist only if the key schedule is considered.*

Proof We will prove this by showing that any nontrivial truncated differential is possible for five-round AES (without MixColumns transformation in the last round, as shown in Fig. 2).

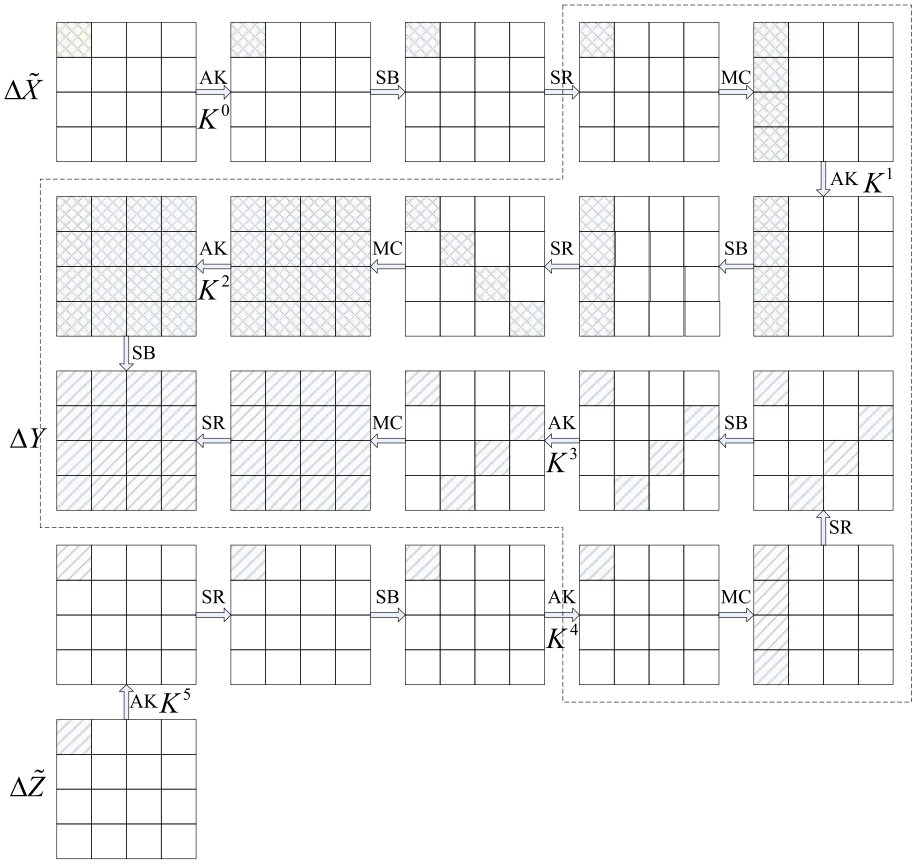

Fig. 2 Five-round AES (Note that this figure only shows the "meet-in-the-middle" of a special truncated differential which is only nonzero at one byte of the input difference and one byte of the output difference, but our proof considers all possible truncated differentials)

Write the five-round AES transformation (without MixColumns transformation in the last round) as: $5\_round\_AES = AK_{K^5} \circ SR \circ SB \circ AK_{K^4} \circ R \circ AK_{MC^{-1}(K^1)} \circ SR \circ SB \circ AK_{K^0}$, here $R = MC \circ SR \circ SB \circ MC \circ SR \circ AK_{SR^{-1} \circ MC^{-1}(K^3)} \circ SB \circ MC \circ AK_{MC^{-1}(K^2)} \circ SR \circ SB \circ MC$.

Since transformations $AK_{MC^{-1}(K^1)} \circ SR \circ SB \circ AK_{K^0}$ before the R and $AK_{K^5} \circ SR \circ SB \circ AK_{K^4}$ after the R, are "transparent"[1] with truncated differences, we only need to prove that for any nontrivial truncated differential $\Delta\tilde{X} \to \Delta\tilde{Z}$, $P(\Delta\tilde{X} \xrightarrow{R} \Delta\tilde{Z}) \neq 0$.

We prove this by showing that there exists concrete differential $\Delta X \to \Delta Z$, such that $\chi(\Delta X) = \Delta\tilde{X}$ and $\chi(\Delta Z) = \Delta\tilde{Z}$, satisfying $P(\Delta X \xrightarrow{R} \Delta Z) \neq 0$. That is to say, there exist corresponding ΔX, ΔZ, X, K^2 and K^3 satisfying $R(X) \oplus R(X \oplus \Delta X) = \Delta Z$.

We write the R as: $R = R_2 \circ AK_{MC^{-1}(K^3)} \circ SR \circ R_1$, here $R_1 = SB \circ MC \circ SR \circ AK_{SR^{-1} \circ MC^{-1}(K^2)} \circ SB \circ MC$, $R_2 = MC \circ SR \circ SB \circ MC$. Note that R_1 is key-dependent, and R_2 is key-independent.

[1] "Transparent" means that the transformations do not mix the input state pattern.

For any nontrivial truncated differential $\Delta\tilde{X} \rightarrow \Delta\tilde{Z}$, we construct one corresponding concrete differential following from "meet-in-the-middle" idea, as following steps:

1. By Lemma 5, for any nonzero $\Delta\tilde{Z}$, there exist ΔZ and Z, such that $\chi(\Delta Z) = \Delta\tilde{Z}$, satisfying $\Delta Y = R_2^{-1}(Z) \oplus R_2^{-1}(Z \oplus \Delta Z)$ is all-nonzero. Then $SR^{-1}(\Delta Y)$ is all-nonzero obviously.
2. By Lemma 4, for $SR^{-1}(\Delta Y)$ and any nonzero $\Delta\tilde{X}$, there exist ΔX, X and K^2, such that $\chi(\Delta X) = \Delta\tilde{X}$, satisfying $SR^{-1}(\Delta Y) = R_1(X) \oplus R_1(X \oplus \Delta X)$.
3. Take $K^3 = MC(SR \circ R_1(X) \oplus R_2^{-1}(Z))$.

Therefore, resulting ΔX, ΔZ, X, K^2 and K^3 satisfying $R(X \oplus \Delta X) \oplus R(X) = \Delta Z$. □

3.3 Special case for AES-256

Notice following property of the AES-256 key schedule:

Property 1 [8] *On knowing 256-bit round keys of any consecutive two rounds for AES-256, the key schedule of AES-256 would allow to uniquely regenerate all the other round keys.*

For AES-256, the assumption that round keys are independent and uniformly random can be removed. We have following theorem:

Theorem 3 *Even though the key schedule is considered, there do not exist truncated impossible differentials covering more than four rounds for AES-256 with S-box considered.*

Proof In the proof of Theorem 2, we prove that any truncated differential $\Delta\tilde{X} \rightarrow \Delta\tilde{Z}$ for the R is possible, by elaborately choosing consecutive two-round keys K^2 and K^3, input X and output Z of the R. By Property 1, "choosing consecutive two-round keys: K^2 and K^3" is feasible in the key schedule of AES-256. Therefore, proof of Theorem 2 is also suitable for AES-256 with key schedule. □

4 Conclusion

In this paper, we investigate the provable security of AES against impossible differential cryptanalysis, by giving the maximal length of truncated impossible differentials. From a more practical point than previous work, we dive into the non-linear S-box, which is always "structured [24]" or "idealized [7]" by other researchers, but significant for the practical maximal length of impossible differentials of SPN block ciphers. After revealing several essential properties of the AES S-box, we prove that there do not exist truncated impossible differentials longer than four rounds for AES with S-box considered and under the assumption that round keys are independent and uniformly random. Specially, for AES-256, even though the S-box and the key schedule are both considered, there do not exist truncated impossible differentials longer than four rounds.

On the security of the AES, impossible differentials covering more than four rounds might be constructed only if the key schedule is considered (exclude AES-256), or the differential is concrete (not truncated differential). These investigation may shed light on how to design block ciphers with provable security against impossible differential attack.

Acknowledgements We would like to thank editors and anonymous reviewers for their patience and constructive suggestions. This work was supported by National Natural Science Foundation of China (Grant Nos. 61272488, 61402523 and 61772547).

References

1. Bahrak B., Aref M.R.: Impossible differential attack on seven-round AES-128. IET Inf. Secur. **2**(2), 28–32 (2008).
2. Beierle C., Jovanovic P., Lauridsen M. M., Leander G., Rechberger C.: Analyzing permutations for AES-like ciphers: understanding ShiftRows. In: CT-RSA, pp. 37–58 (2015).
3. Biham E., Biryukov A., Shamir A.: Cryptanalysis of Skipjack reduced to 31 rounds using impossible differentials. J. Cryptol. **18**(4), 291–311 (2005).
4. Biham E., Keller N.: Cryptanalysis of reduced variants of Rijndael. In: The 3rd AES Conference (2000).
5. Biham E., Shamir A.: Differential cryptanalysis of DES-like cryptosystems. J. Cryptol. **4**(1), 3–72 (1991).
6. Cheon J.H., Kim M., Kim K., Lee J.-Y., Kang S.: Improved impossible differential cryptanalysis of Rijndael and Crypton. In: ICISC, pp. 39–49 (2001).
7. Cui T., Jin C., Zhang B., Chen Z.: Searching all truncated impossible differentials in SPN. IET Inf. Secur. doi:10.1049/iet-ifs.2015.0052.
8. Daemen J., Rijmen V.: AES proposal: Rijndael. http://csrc.nist.gov/archive/aes/rijndael/Rijndael-ammended.pdf.
9. Daemen J., Rijmen V.: The wide trail design strategy. In: IMA International Conference pp. 222–238 (2001).
10. Daemen, J., Rijmen, V.: Understanding two-round differentials in AES. In: SCN, pp. 78–94 (2006).
11. Grassi L., Rechberger C., Rnjom S.: Subspace trail cryptanalysis and its applications to AES. IACR Trans. Symmetric Cryptol. (2), 192–225 (2016).
12. Hungerford T.W.: Algebra. Springer, New York (1974).
13. Kanda M., Matsumoto T.: Security of Camellia against truncated differential cryptanalysis. In: FSE, pp. 286–299 (2001).
14. Kim J., Hong S., Lim J.: Impossible differential cryptanalysis using matrix method. Discret. Math. **310**(5), 988–1002 (2010).
15. Knudsen L.R.: DEAL-A 128-bit block cipher. Technical Report, Department of Informatics, University of Bergen, Norway (1998).
16. Knudsen L.R.: Truncated and higher order differentials. In: FSE, pp. 196–211 (1994).
17. Lidl R., Niederreiter H.: Finite Fields. Cambridge University Press, Cambridge (1996).
18. Luo Y., Lai X., Wu Z., Gong G.: A unified method for finding impossible differentials of block cipher structures. Inf. Sci. **263**, 211–220 (2014).
19. Matsui M.: Linear cryptanalysis method for DES cipher. In: Eurocrypt, pp. 386–397 (1993).
20. NIST. FIPS 197: announcing the advanced encryption standard (AES). Technical Report, National Institute of Standards and Technology (NIST) (2001). http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf.
21. Nyberg K.: Differentially uniform mappings for cryptography. In: Eurocrypt, pp. 55–64 (1993).
22. Phan C.W.: Impossible differential cryptanalysis of 7-round Advanced Encryption Standard (AES). Inf. Process. Lett. **91**(1), 33–38 (2004).
23. Sun B., Liu M., Guo J., Qu L., Rijmen V.: New insights on AES-like SPN ciphers. In: Crypto (1), pp. 605–624 (2016).
24. Sun B., Liu M., Guo J., Rijmen V., Li R.: Provable security evaluation of structures against impossible differential and zero correlation linear cryptanalysis. In: Eurocrypt (1), pp. 196–213 (2016).
25. 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: Crypto (1), pp. 95–115 (2015).
26. Wu S., Wang M.: Automatic search of truncated impossible differentials for word-oriented block ciphers. In: Indocrypt, pp. 283–302 (2012).