Boomerang Connectivity Table Revisited

Application to SKINNY and AES

Ling Song[1,2,3], Xianrui Qin[4] and Lei Hu[3]

[1] School of Physical and Mathematical Sciences
Nanyang Technological University, Singapore

[2] Strategic Centre for Research in Privacy-Preserving Technologies and Systems
Nanyang Technological University, Singapore

[3] State Key Laboratory of Information Security,
Institute of Information Engineering, Chinese Academy of Sciences, China

[4] School of Mathematics, Shandong University, China
songling.qs@gmail.com, qxr@mail.sdu.edu.cn, hulei@iie.ac.cn

Abstract. The boomerang attack is a variant of differential cryptanalysis which regards a block cipher E as the composition of two sub-ciphers, *i.e.*, $E = E_1 \circ E_0$, and which constructs distinguishers for E with probability p^2q^2 by combining differential trails for E_0 and E_1 with probability p and q respectively. However, the validity of this attack relies on the dependency between the two differential trails. Murphy has shown cases where probabilities calculated by p^2q^2 turn out to be zero, while techniques such as boomerang switches proposed by Biryukov and Khovratovich give rise to probabilities greater than p^2q^2. To formalize such dependency to obtain a more accurate estimation of the probability of the distinguisher, Dunkelman *et al.* proposed the sandwich framework that regards E as $\tilde{E}_1 \circ E_m \circ \tilde{E}_0$, where the dependency between the two differential trails is handled by a careful analysis of the probability of the middle part E_m. Recently, Cid *et al.* proposed the Boomerang Connectivity Table (`BCT`) which unifies the previous switch techniques and incompatibility together and evaluates the probability of E_m theoretically when E_m is composed of a single S-box layer. In this paper, we revisit the `BCT` and propose a generalized framework which is able to identify the actual boundaries of E_m which contains dependency of the two differential trails and systematically evaluate the probability of E_m with any number of rounds. To demonstrate the power of this new framework, we apply it to two block ciphers `SKINNY` and `AES`. In the application to `SKINNY`, the probabilities of four boomerang distinguishers are re-evaluated. It turns out that E_m involves 5 or 6 rounds and the probabilities of the full distinguishers are much higher than previously evaluated. In the application to `AES`, the new framework is used to exclude incompatibility and find high probability distinguishers of `AES`-128 under the related-subkey setting. As a result, a 6-round distinguisher with probability $2^{-109.42}$ is constructed. Lastly, we discuss the relation between the dependency of two differential trails in boomerang distinguishers and the properties of components of the cipher.

Keywords: block cipher · boomerang attack · sandwich attack · boomerang connectivity table · SKINNY · AES

1 Introduction

Differential cryptanalysis, proposed by Biham and Shamir [BS93], is one of the most powerful approaches to assess the security of block ciphers. The basic idea is to exploit

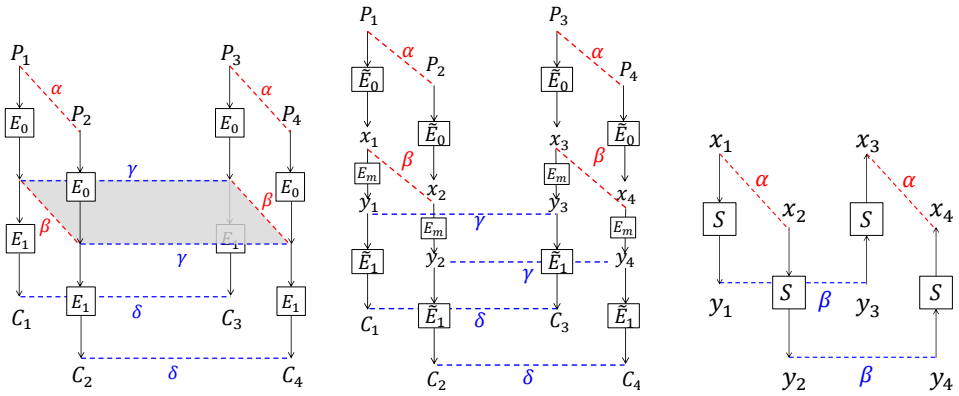

Figure 1: Basic boomerang attack (left), sandwich attack (middle) and E_m with only an S-box layer,

non-random pairs of input and output differences of the cipher, *i.e.*, high probability differentials. In many cases, it is hard or impossible to find long differentials. In such cases, the boomerang attack [Wag99] which was proposed as an extension of the differential cryptanalysis may be applied to combine short differentials with high probabilities to get a long one.

In boomerang attacks, a cipher E is regarded as the composition of two sub-ciphers E_0 and E_1, *i.e.*, $E = E_1 \circ E_0$. Suppose there exists a differential $\alpha \to \beta$ of E_0 with probability p and a differential $\gamma \to \delta$ of E_1 with probability q. Under the assumption that the two differentials are independent, the boomerang attack exploits the high probability of the following differential property:

$$\Pr\left[E^{-1}(E(P_1) \oplus \delta) \oplus E^{-1}(E(P_1 \oplus \alpha) \oplus \delta) = \alpha\right] = p^2q^2.$$

As illustrated in the left part of Figure 1, two plaintexts with difference α are encrypted and the resulting ciphertexts are then XORed with δ to generate two new ciphertexts. These two new ciphertexts are then decrypted to give two new plaintexts. If the difference between the two new plaintexts is also α, it is said the boomerang returns and the two pairs of plaintexts form a *right quartet*. According to [Wag99], if $(pq)^{-2} < 2^n$, where n is the block size, then E can be distinguished from an ideal cipher with a complexity corresponding to $(pq)^{-2}$ adaptive chosen plaintext/ciphertext queries.

Later, refinements on the boomerang attack were proposed. Particularly, Kelsey et al. [KKS00] developed amplified boomerangs which are pure chosen-plaintext attacks. In amplified boomerang attacks, the probability of finding a right quartet is $2^{-n}p^2q^2$ while for a random permutation the expected probability is 2^{-2n}. In [BDK01], Biham et al. proposed the rectangle attack which allows any value of β and γ to occur as long as $\beta \neq \gamma$. As a result, the probability of generating a right quartet increases to $2^{-n}\hat{p}^2\hat{q}^2$, where $\hat{p} = \sqrt{\Sigma_i \Pr^2(\alpha \to \beta_i)}$ and $\hat{q} = \sqrt{\Sigma_j \Pr^2(\gamma_j \to \delta)}$.

In applications of boomerang attacks to concrete block ciphers, such as [Wag99,BDK01, ALLW14], attackers typically aim to find differential trails with high probability and then combine them to form long boomerang distinguishers. However, the dependency between the two differential trails highly affects the probability of the boomerang distinguisher. As pointed out by Murphy in [Mur11], there exist cases where the probabilities formulated by p^2q^2 are highly inaccurate. He showed that in some cases of S-box based ciphers, two independently chosen differential trails are *incompatible*, making the boomerang never return, and in other cases, the dependency leads to a higher probability than p^2q^2.

Further, Biryukov *et al.* made an improvement on exploiting the positive dependency of boomerang distinguishers, which was named *boomerang switch* [BK09]. The idea was to optimize the transition between the differential trails of E_0 and E_1 in order to minimize the overall complexity of the boomerang distinguisher. In [BK09], three types of switches were proposed. Instead of decomposing a cipher into rounds by default, the ladder switch decomposes the cipher regarding smaller operations, like columns and bytes, which may lead to better distinguishers. The S-box switch refers to the case when both differential trails activate the same S-box with identical input and output differences, the probability of this S-box counts only once for the boomerang distinguisher. The Feistel switch, also noted in [BDK05], stands for a free middle round in the boomerang distinguisher for a Feistel cipher.

The above cases of dependency were later covered and unified in the *sandwich attack* proposed by Dunkelman *et al.* [DKS10, DKS14], which is depicted in the middle part of Figure 1. It regards E as $E = \tilde{E}_1 \circ E_m \circ \tilde{E}_0$ instead, where the middle part E_m specifically handles the dependency and contains a relatively small number of rounds. If the probability of generating a right quartet for E_m is r, then the probability of the whole boomerang distinguisher is

$$\Pr\left[E^{-1}(E(P_1) \oplus \delta) \oplus E^{-1}(E(P_1 \oplus \alpha) \oplus \delta) = \alpha\right] = \tilde{p}^2\tilde{q}^2 r,$$

where $\tilde{p}$ (resp. $\tilde{q}$) is the probability of the differential of $\tilde{E}_0$ (resp. $\tilde{E}_1$). Let (x_1,x_2,x_3,x_4) and (y_1,y_2,y_3,y_4) be input and output quartet values for E_m, where $y_i = E_m(x_i)$. Suppose the differential trail for $\tilde{E}_0$ (resp. $\tilde{E}_1$) ends (resp. starts) with difference β (resp. γ), *i.e.*, $x_1 \oplus x_2 = x_3 \oplus x_4 = \beta$ and $y_1 \oplus y_3 = y_2 \oplus y_4 = \gamma$. Then, r was formally defined as:

$$r = \Pr\Big[(x_3 \oplus x_4) = \beta | (x_1 \oplus x_2 = \beta) \wedge (y_1 \oplus y_3 = \gamma) \wedge (y_2 \oplus y_4 = \gamma)\Big]$$

In [DKS10, DKS14], the probability r of E_m was evaluated by experiments.

Recently in [CHP+18], the issue of dependency in boomerang distinguishers was revisited, and a tool named *Boomerang Connectivity Table* (`BCT`) was proposed, which calculates r theoretically when E_m is composed of a single S-box layer, as shown in the right part of Figure 1. More importantly, the previous observations on the S-box including the ladder switch and the S-box switch as well as the incompatibility can be well explained by `BCT`, which gives new insights into boomerang attacks and provides a new point of view for designing a good S-box. As a follow-up, Boura and Canteaut [BC18] gave a thorough analysis of `BCT` properties of some important families of S-boxes.

Although the introductory paper of `BCT` [CHP+18] well handles the dependency of two differential trails in boomerang distinguishers when E_m is of one S-box layer, the following questions may be asked naturally.

- How to decide the actual boundaries of E_m which contains dependency of the two differential trails in boomerang distinguishers?
- How to calculate r when E_m contains multiple rounds?

Answers to these questions would be of great importance on evaluating the probability of the boomerang distinguishers. Only when the probability of the boomerang distinguisher is accurately computed can we evaluate the exact resistance of a cipher against boomerang attacks.

Our contributions. This paper gives the first solution to the above questions by proposing a generalized framework of `BCT`. Specifically, our new framework is able to not only find the actual boundaries of E_m which contains dependency of two differential trails in the setting of boomerang attacks, but also systematically calculate the probability r of

E_m with any number of rounds. With the issues of E_m settled, the probability of the full distinguisher of $E = \tilde{E}_1 \circ E_m \circ \tilde{E}_0$ can then be closely modeled by $\tilde{p}^2\tilde{q}^2 r$.

To achieve this, we start with the basic formula of BCT and then extend it to general cases. Specifically, new formulas are developed for all possible cases with the help of a new concept named *crossing difference* which refers to the difference propagated from the other differential trail of the boomerang distinguisher. With the crossing difference, the middle part E_m can be well described. First, the boundaries of E_m are delineated by the round where the crossing differences turn into random. Second, the probability r of E_m depends on the distribution of the crossing difference. Finally, the case considered in [CHP+18] where E_m is of one S-box layer maps to the case here where the crossing differences are fixed.

To demonstrate the power of our generalized framework, we apply it to SKINNY [BJK+16] and AES [DR02], which are two typical block ciphers using weak and strong round functions respectively. In the case of SKINNY, we re-evaluate the probabilities of the four boomerang distinguisher proposed in [LGS17] and the results are summarized in Table 1. As shown in Table 1, the lengths of E_m for these distinguishers are 5 or 6 rounds. The corresponding r probabilities are computed and confirmed by experiments. Adjacent to E_m there are some passive rounds for all these distinguishers, so the probability remains r with these passive rounds included. The increased numbers of rounds by adding these passive rounds are displayed in parentheses. The probabilities of the full boomerang distinguishers are then computed accordingly with $\tilde{p}^2\tilde{q}^2 r$ which turn out to be much higher than the probabilities given in [LGS17] by $\hat{p}^2\hat{q}^2$. In the case of AES, we propose a 6-round related-subkey boomerang distinguisher of probability $2^{-109.42}$ by combining two 3-round differential trails. In this case, E_m is of two rounds. Our framework is then used to exclude incompatibility and optimize $\tilde{p}^2\tilde{q}^2 r$ by selecting a good combination.

Table 1: Probabilities of the boomerang distinguishers of SKINNY where $|E_m|$ denotes the number of rounds E_m contains

Version	n	E_m		$E = \tilde{E}_1 \circ E_m \circ \tilde{E}_0$			Trails
		$\lvert E_m\rvert$	r	$\lvert E\rvert$	$\tilde{p}^2\tilde{q}^2 r$	Pr. [LGS17]	
n-2n	64	6 (13)	$2^{-12.96}$	17	$\mathbf{2^{-29.78}}$	$2^{-48.72}$	Table 7
	128	5 (12)	$2^{-11.45}$	18	$\mathbf{2^{-77.83}}$	$2^{-103.84}$	Table 4
n-3n	64	5 (17)	$2^{-10.50}$	22	$\mathbf{2^{-42.98}}$	$2^{-54.94}$	Table 8
	128	5 (17)	$2^{-9.88}$	22	$\mathbf{2^{-48.30}}$	$2^{-76.84}$	Table 8

Lastly, we discuss the relation between the dependency of two differential trails in boomerang distinguishers and the properties of the round function. It is deduced from our generalized framework that the length of E_m is mainly determined by the diffusion effect of the linear layer, and the probability r is strongly affected by differential properties of the non-linear layer.

Organization. The rest of the paper is organized as follows. Section 2 provides preliminaries of boomerang attacks and previous works on BCT. Our generalized framework of BCT is presented in Section 3. Section 4 applies the new framework to SKINNY. Section 5 extends the application to AES. We then discuss in Section 6 the relation between the dependency occurring in boomerang distinguishers and the properties of the cipher. Finally, Section 7 concludes the paper.

2 Preliminaries

This section gives a clearer picture of the boomerang attack and reviews the previous works on the boomerang connectivity table. In addition, notations used throughout this paper are also introduced.

2.1 Framework of Boomerang Attacks

The boomerang attack, proposed by David Wagner [Wag99], treats a block cipher E as the composition of two sub-ciphers E_0 and E_1, for which there exist short differentials $\alpha \to \beta$ and $\gamma \to \delta$ of probabilities p and q respectively. The two differentials are then combined in a chosen plaintext and ciphertext attack setting to construct a long boomerang distinguisher, as shown in Figure 1. Later, the basic boomerang attack was extended to the related-key setting and was formulated in [BDK05] by using four related-key oracles.

Let $E_K(P)$ and $E_K^{-1}(C)$ denote the encryption of P and the decryption of C under a key K, respectively. Suppose $\Delta K, \nabla K$ are the master key differences of the differentials. Then the boomerang framework in the related-key setting works as follows.

1. $K_1 \leftarrow K, K_2 \leftarrow K_1 \oplus \Delta K, K_3 \leftarrow K_1 \oplus \nabla K, K_4 \leftarrow K_1 \oplus \Delta K \oplus \nabla K$.
2. Repeat the following steps many times.
   (a) $P_1 \leftarrow random()$ and $P_2 \leftarrow P_1 \oplus \alpha$.
   (b) $C_1 \leftarrow E_{K_1}(P_1)$ and $C_2 \leftarrow E_{K_2}(P_2)$.
   (c) $C_3 \leftarrow C_1 \oplus \delta$ and $C_4 \leftarrow C_2 \oplus \delta$.
   (d) $P_3 \leftarrow E_{K_3}^{-1}(C_3)$ and $P_4 \leftarrow E_{K_4}^{-1}(C_4)$.
   (e) Check if $P_3 \oplus P_4 = \alpha$.

In step 2(e), if $P_3 \oplus P_4 = \alpha$ holds, then a *right quartet* (P_1, P_2, P_3, P_4) is found such that $P_1 \oplus P_2 = P_3 \oplus P_4 = \alpha$ and $C_1 \oplus C_3 = C_2 \oplus C_4 = \delta$. This happens with probability p^2q^2 under the assumption that the two differentials are independent.

2.2 Boomerang Connectivity Table

We introduce here the definitions and propositions related to the boomerang connectivity table of S-box $S : \mathbb{F}_2^n \to \mathbb{F}_2^n$.

Definition 1 (Difference Distribution Table)**.** Let S be a function from $\mathbb{F}_2^n$ to $\mathbb{F}_2^n$. The difference distribution table (`DDT`) is a two-dimensional table defined by

$$\texttt{DDT}(\alpha, \beta) = \#\{x \in \mathbb{F}_2^n : S(x) \oplus S(x \oplus \alpha) = \beta\}, \text{ where } \alpha, \beta \in \mathbb{F}_2^n.$$

The differential uniformity of S is the highest value in the `DDT` except for the first row and the first column.

Definition 2 (Boomerang Connectivity Table [CHP+18])**.** Let S be a permutation of $\mathbb{F}_2^n$. The boomerang connectivity table (`BCT`) of S is a two-dimensional table defined by

$$\texttt{BCT}(\alpha, \beta) = \#\{x \in \mathbb{F}_2^n : S^{-1}(S(x) \oplus \beta) \oplus S^{-1}(S(x \oplus \alpha) \oplus \beta) = \alpha\}, \text{ where } \alpha, \beta \in \mathbb{F}_2^n.$$

The boomerang uniformity of S is the highest value in the `BCT` except for the first row and the first column.

Example 1. *The* `SKINNY`*'s 4-bit S-box takes the following mapping. Its* `DDT` *and* `BCT` *are provided in Table 2 and 3 respectively. It can be noted that its differential and boomerang uniformities are 4 and 16 respectively*

x	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
$S[x]$	c	6	9	0	1	a	2	b	3	8	5	d	4	e	7	f

Table 2: `DDT` of `SKINNY`'s 4-bit S-box

	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	4	4	4	4	0	0	0	0
2	0	4	0	4	0	4	4	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	2	2	2	2	2	2	2	2
4	0	0	4	0	0	0	2	2	0	0	0	4	2	2	0	0
5	0	0	4	0	0	0	2	2	0	0	4	0	2	2	0	0
6	0	2	0	2	2	0	0	2	2	0	2	0	0	2	2	0
7	0	2	0	2	2	0	0	2	0	2	0	2	2	0	0	2
8	0	0	0	0	4	4	0	0	0	0	0	0	2	2	2	2
9	0	0	0	0	4	4	0	0	0	0	0	0	2	2	2	2
a	0	0	0	0	0	4	4	0	2	2	2	2	0	0	0	0
b	0	4	0	4	0	0	0	0	0	0	0	0	2	2	2	2
c	0	0	4	0	0	0	2	2	4	0	0	0	0	0	2	2
d	0	0	4	0	0	0	2	2	0	4	0	0	0	0	2	2
e	0	2	0	2	2	0	0	2	0	2	0	2	0	2	2	0
f	0	2	0	2	2	0	0	2	2	0	2	0	2	0	0	2

Table 3: `BCT` of `SKINNY`'s 4-bit S-box

	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
0	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
1	16	0	16	0	0	0	0	0	8	8	8	8	0	0	0	0
2	16	8	0	8	8	16	8	0	0	0	0	0	0	0	0	0
3	16	0	0	0	0	0	0	0	2	2	2	2	2	2	2	2
4	16	0	8	0	0	0	2	2	4	4	4	4	2	2	0	0
5	16	0	8	0	0	0	2	2	4	4	4	4	2	2	0	0
6	16	2	0	2	2	0	0	2	2	0	2	0	0	2	2	0
7	16	2	0	2	2	0	0	2	0	2	0	2	2	0	0	2
8	16	4	0	4	4	8	4	0	0	0	0	0	2	2	2	2
9	16	4	0	4	4	8	4	0	0	0	0	0	2	2	2	2
a	16	4	0	4	4	8	4	0	2	2	2	2	0	0	0	0
b	16	4	0	4	4	8	4	0	0	0	0	0	2	2	2	2
c	16	0	8	0	0	0	2	2	4	4	4	4	0	0	2	2
d	16	0	8	0	0	0	2	2	4	4	4	4	0	0	2	2
e	16	2	0	2	2	0	0	2	0	2	0	2	0	2	2	0
f	16	2	0	2	2	0	0	2	2	0	2	0	2	0	0	2

Following [CLN+17], we introduce the notation:

$$\mathcal{X}_{\mathtt{DDT}}(\alpha,\beta) \triangleq \{x \in \mathbb{F}_2^n : S(x) \oplus S(x \oplus \alpha) = \beta\},$$
$$\mathcal{Y}_{\mathtt{DDT}}(\alpha,\beta) \triangleq \{S(x) \in \mathbb{F}_2^n : x \in \mathbb{F}_2^n, S(x) \oplus S(x \oplus \alpha) = \beta\}.$$

Based on the above notation, the proposition from [BC18] that illuminates the relation between `BCT` and `DDT` can be written as follow.

Proposition 1 ([BC18]). *For any permutation S of $\mathbb{F}_2^n$, for all $\alpha, \beta \in \mathbb{F}_2^n$, we have*

$$\mathtt{BCT}(\alpha,\beta) = \mathtt{DDT}(\alpha,\beta) + \sum_{\gamma \neq 0,\beta} \#(\mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma) \cap (\mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma) \oplus \beta)). \tag{1}$$

Note that, due to symmetry, Eq. 1 is equivalent to

$$\mathtt{BCT}(\alpha,\beta) = \mathtt{DDT}(\alpha,\beta) + \sum_{\gamma \neq 0,\alpha} \#(\mathcal{X}_{\mathtt{DDT}}(\gamma,\beta) \cap (\mathcal{X}_{\mathtt{DDT}}(\gamma,\beta) \oplus \alpha)).$$

2.3 Notations

We treat the block cipher as $E = E_1 \circ E_0$ where there exist differential trails of E_0 and E_1 with probabilities p and q respectively. Let E_m denote the middle part of the cipher which contains dependency of the two differential trails. The probability of E_m of generating a right quartet is denoted by r. We let $\tilde{E_0} \leftarrow E_0 \backslash E_m$, *i.e.*, the front rounds of E_0 that do not contain dependency, and $\tilde{E_1} \leftarrow E_1 \backslash E_m$, *i.e.*, the rear rounds of E_1 that do not contain dependency. With E_m clearly defined, we treat $E = \tilde{E_1} \circ E_m \circ \tilde{E_0}$ so that the probability of generating a right quartet (also called the probability of the boomerang distinguisher) can be computed precisely as $\tilde{p}^2\tilde{q}^2 r$ where $\tilde{p}$ (resp. $\tilde{q}$) is the probability of the differential trail of $\tilde{E_0}$ (resp. $\tilde{E_1}$). We denote the number of rounds in E by $|E|$.

3 Generalized Framework of BCT

In this section, through a new explanation of BCT, we extend the previous analysis in [CHP+18] for E_m with only one S-box layer to the case where E_m contains multiple rounds and we also show how to decide the boundaries of E_m.

In the beginning, we treat the block cipher as $E = E_1 \circ E_0$, then identify the middle part E_m that contains dependency. Once E_m is identified, we let $\tilde{E_0} \leftarrow E_0 \backslash E_m$ and $\tilde{E_1} \leftarrow E_1 \backslash E_m$ so that $E = \tilde{E_1} \circ E_m \circ \tilde{E_0}$.

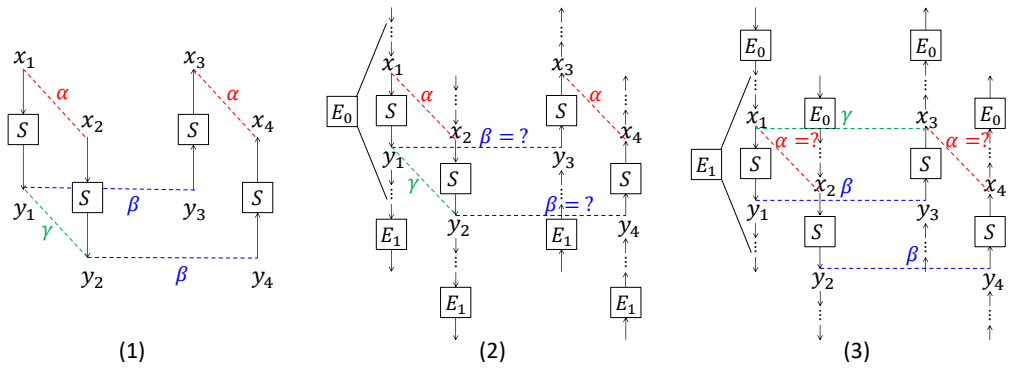

Figure 2: (1) S-box at the connecting point, (2) S-box in E_0 and (3) S-box in E_1

3.1 New Explanation

We first consider the E_m with only one S-box layer at the connecting point of E_0 and E_1, as shown in Figure 2(1). For such E_m, the differences α, β are specified by the differential trails. By re-expressing Eq. 1 as

$$\mathtt{BCT}(\alpha,\beta) = \sum_{\gamma} \#(\mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma) \cap (\mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma) \oplus \beta)), \tag{2}$$

it is known that only those γs are possible when $\mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma) \cap (\mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma) \oplus \beta)$ is not empty, *i.e.*, there exists $y_1 \in \mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma)$ such that $y_1 \oplus \beta$ also belongs to $\mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma)$, as depicted in Figure 2(1). If $y_1 \oplus \beta \in \mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma)$, we will always have $x_3 \oplus x_4 = \alpha$; otherwise, the boomerang never returns. Therefore, the probability of E_m for generating a right quartet is

$$r = \frac{\mathtt{BCT}(\alpha,\beta)}{2^n} = \sum_{\gamma} \frac{\mathtt{DDT}(\alpha,\gamma)}{2^n} \cdot \frac{\#\{y \in \mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma) : y \oplus \beta \in \mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma)\}}{\#\mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma)}. \tag{3}$$

Similar results can be obtained as follows with $\mathcal{X}_{\mathtt{DDT}}$ due to symmetry.

$$\begin{aligned} \mathtt{BCT}(\alpha,\beta) &= \sum_{\gamma'} \#(\mathcal{X}_{\mathtt{DDT}}(\gamma',\beta) \cap (\mathcal{X}_{\mathtt{DDT}}(\gamma',\beta) \oplus \alpha)), \\ r = \frac{\mathtt{BCT}(\alpha,\beta)}{2^n} &= \sum_{\gamma'} \frac{\mathtt{DDT}(\gamma',\beta)}{2^n} \cdot \frac{\#\{x \in \mathcal{X}_{\mathtt{DDT}}(\gamma',\beta) : x \oplus \alpha \in \mathcal{X}_{\mathtt{DDT}}(\gamma',\beta)\}}{\#\mathcal{X}_{\mathtt{DDT}}(\gamma',\beta)}. \end{aligned} \tag{4}$$

Even though Eq. 3 and 4 look more complex than Eq. 1, they are helpful when we consider E_m of multiple rounds. In fact, the dependency of two differential trails may penetrate into multiple rounds. Next, we are to extend the analysis to E_m with multiple layers of S-boxes around the connecting point of E_0 and E_1, and find the boundaries of E_m, as well as evaluate the probability of it.

3.2 Generalization

Now, we consider S-boxes in general cases which are not necessarily located at the connecting point of E_0 and E_1. We observe that for S-boxes away from the connecting point, the differences α in Eq. 4 or β in Eq. 3 (to be defined as crossing differences) may not be fixed but follow some distributions. Our intuition is to take into account the distributions which turn out to be a key factor for evaluating the dependency of two differential trails in a boomerang distinguisher.

Suppose the input and output differences of S-boxes in the two differential trails are given. We use $\bar{r}$ to denote the probability of getting a right quartet that follows exact differential trails. In fact, the differences in between may have many choices. The actual probability r is composed of the probabilities $\bar{r}$ corresponding to all possible intermediate differences and hence r is usually greater than or equal to any single $\bar{r}$. As an analogy to the clustering effect of differentials, we call this the *clustering effect of* E_m.

Active S-boxes in E_0. Let us consider an active S-box in the upper differential trail of E_0. This situation is illustrated in Figure 2(2) and 3(1). Suppose the input and output differences α, γ of this S-box are specified by the upper differential trail. β is the difference propagated from the lower differential trail and called the **lower crossing difference**, as depicted in Figure 2(2). The value of β may not be fixed. From Eq. 2, it can be seen that only when $\mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma) \cap (\mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma) \oplus \beta)$ is not empty will the boomerang return. That is, both y_1 and $y_3 = y_1 \oplus \beta$ should belong to $\mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma)$ (see Figure 2(2)). If the distribution of the lower crossing difference β is independent of the upper differential trail, *i.e.*, the value of β is not affected by the upper differential trail as showcased in 3(1), then the probability that the boomerang returns when the output difference of this S-box is γ is

$$\bar{r} = \frac{\mathtt{DDT}(\alpha,\gamma)}{2^n} \cdot \sum_{\beta} \frac{\#\{y \in \mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma) : y \oplus \beta \in \mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma)\}}{\#\mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma)} \cdot \mathtt{Pr}(y_1 \oplus y_3 = \beta). \tag{5}$$

When we take all possible output differences γ of this S-box into account, we have

$$r = \sum_{\gamma} \bar{r} = \sum_{\beta} \frac{\mathtt{BCT}(\alpha,\beta)}{2^n} \cdot \mathtt{Pr}(y_1 \oplus y_3 = \beta). \tag{6}$$

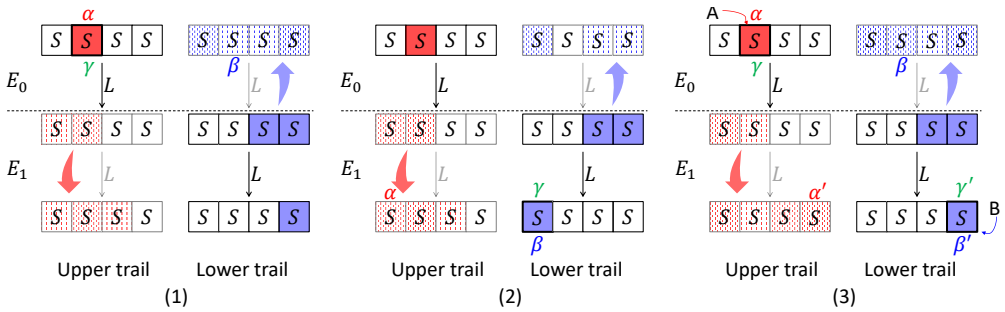

Figure 3: Toy examples of (1) active S-boxes in E_0, (2) active S-boxes in E_1 and (3) interrelated active S-boxes, where 'S' denotes an n-bit S-box, 'L' denotes the linear layer and the red (resp. blue) arrows stand for extensions of the upper (resp. lower) differential trails with probability 1.

Particularly, if the lower crossing difference β is constant, then

$$\bar{r} = \frac{\mathtt{DDT}(\alpha,\gamma)}{2^n} \cdot \frac{\#\{y \in \mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma) : y \oplus \beta \in \mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma)\}}{\#\mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma)},$$

and $r = \sum_\gamma \bar{r} = \frac{\mathtt{BCT}(\alpha,\beta)}{2^n}$ is exactly the same as Eq. 3. If β is always 0,

$$\bar{r} = \frac{\mathtt{DDT}(\alpha,\gamma)}{2^n} \text{ and } r = 1.$$

If the lower crossing difference β is uniformly distributed, *i.e.*, for any $a \in \mathbb{F}_2^n$, $\mathtt{Pr}(\beta = a) = \frac{1}{2^n}$, then

$$\bar{r} = \left(\frac{\mathtt{DDT}(\alpha,\gamma)}{2^n}\right)^2,$$

which becomes identical to the computation by p^2q^2 in the classical boomerang attack.

Active S-boxes in E_1. For an active S-box in the lower differential trail of E_1, similar results can be obtained. Suppose the output and input differences β, γ of this S-box are specified by the lower differential trail, as shown in Figure 2(3) and 3(2). In this case, α is the difference propagated from the upper differential trail and called the **upper crossing difference**. The value of α may not be fixed. From Eq. 4, it can be seen that only when $\mathcal{X}_{\mathtt{DDT}}(\gamma,\beta) \cap (\mathcal{X}_{\mathtt{DDT}}(\gamma,\beta) \oplus \alpha)$ is not empty will the boomerang return. That is, both x_1 and $x_2 = x_1 \oplus \alpha$ should belong to $\mathcal{X}_{\mathtt{DDT}}(\gamma,\beta)$ (see Figure 2(3)). If the distribution of the upper crossing difference α is independent of the lower differential trail, *i.e.*, the value of α is not affected by the lower differential trail as showcased in 3(2), then the probability that the boomerang returns when the input difference of this S-box is γ is

$$\bar{r} = \frac{\mathtt{DDT}(\gamma,\beta)}{2^n} \cdot \sum_\alpha \frac{\#\{x \in \mathcal{X}_{\mathtt{DDT}}(\gamma,\beta) : x \oplus \alpha \in \mathcal{X}_{\mathtt{DDT}}(\gamma,\beta)\}}{\#\mathcal{X}_{\mathtt{DDT}}(\gamma,\beta)} \cdot \mathtt{Pr}(x_1 \oplus x_2 = \alpha). \quad (7)$$

When we take all possible input difference γ of this S-box into account, we have

$$r = \sum_\gamma \bar{r} = \sum_\alpha \frac{\mathtt{BCT}(\alpha,\beta)}{2^n} \cdot \mathtt{Pr}(x_1 \oplus x_2 = \alpha). \quad (8)$$

Particularly, if the upper crossing difference α is constant, then

$$\bar{r} = \frac{\mathtt{DDT}(\gamma,\beta)}{2^n} \cdot \frac{\#\{x \in \mathcal{X}_{\mathtt{DDT}}(\gamma,\beta) : x \oplus \alpha \in \mathcal{X}_{\mathtt{DDT}}(\gamma,\beta)\}}{\#\mathcal{X}_{\mathtt{DDT}}(\gamma,\beta)},$$

and $r = \sum_\gamma \bar{r} = \frac{\mathtt{BCT}(\alpha,\beta)}{2^n}$ is exactly the same as Eq. 4. If α is always 0,

$$\bar{r} = \frac{\mathtt{DDT}(\gamma,\beta)}{2^n} \text{ and } r = 1.$$

If the upper crossing difference α is uniformly distributed, *i.e.*, for any $a \in \mathbb{F}_2^n$, $\mathtt{Pr}(\alpha = a) = \frac{1}{2^n}$, then

$$\bar{r} = \left(\frac{\mathtt{DDT}(\gamma,\beta)}{2^n}\right)^2,$$

which becomes identical to the computation by p^2q^2 in the classical boomerang attack.

Interrelated active S-boxes. It is possible that active S-box `A` in E_0 and active S-box `B` in E_1 affect each other, as showcased in Figure 3(3). Suppose the differential of S-box `A` is $\alpha \to \gamma$, according to the upper differential trail. Similarly, the differential of S-box `B` is $\gamma' \to \beta'$, according to the lower differential trail. The interrelation here refers to two things. One is that the upper crossing difference α' of S-box `B` is propagated from S-box `A`, and the other is that the lower crossing difference β of S-box `A` is propagated from S-box `B`. To calculate the probability, we further introduce

$$\begin{aligned}\mathcal{D}_{\mathtt{BCT}}(\alpha,\beta,\gamma) \triangleq \{x \in \mathbb{F}_2^n : S^{-1}(S(x) \oplus \beta) \oplus S^{-1}(S(x \oplus \alpha) \oplus \beta) = \alpha,\\ x \oplus S^{-1}(S(x) \oplus \beta) = \gamma\}.\end{aligned}$$

Then, we have

$$\begin{aligned}\bar{r} = \sum_{\alpha'} \frac{\mathtt{DDT}(\alpha,\gamma)}{2^n} \cdot \mathtt{Pr}(\gamma \to \alpha') \frac{\mathcal{D}_{\mathtt{BCT}}(\alpha',\beta',\gamma')}{2^n} \cdot \mathtt{Pr}(\gamma' \to \beta') \cdot\\ \frac{\#\{y \in \mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma) : y \oplus \beta \in \mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma)\}}{\#\mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma)}.\end{aligned}$$

Let (y_1, y_2, y_3, y_4) be the output quartet of S-box `A` and (x'_1, x'_2, x'_3, x'_4) be the input quartet of S-box `B`. The above formula means that both the condition for S-box `A` that y_1 and $y_3 = y_1 \oplus \beta$ belong to $\mathcal{Y}_{\mathtt{DDT}}(\alpha,\gamma)$ and the condition for S-box `B` that x'_1 and $x'_2 = x'_1 \oplus \alpha'$ belong to $\mathcal{X}_{\mathtt{DDT}}(\gamma',\beta')$ should be satisfied simultaneously.

When we consider all possible output differences γ of S-box `A` and all possible input differences γ' of S-box `B`, we have

$$r = \sum_\gamma \sum_{\gamma'} \bar{r}. \tag{9}$$

If more than two active S-boxes affect each other, a similar analysis can be performed to calculate the probability r. Such examples can be found in Section 4.3.

Boundaries of E_m. From the above analysis, it can be deduced that the upper boundary of E_m is delineated by the round where the lower crossing differences for its active S-boxes are distributed (almost) uniformly. Also, the lower boundary of E_m is marked by the round where the upper crossing differences for its active S-boxes are distributed (almost) uniformly. Due to this, the length of E_m heavily depends on the diffusion properties of the cipher, which will be exemplified by the application to SKINNY and AES in the following two sections.

3.3 Algorithm for Evaluating r

Given two differential trails over E_0, E_1 respectively, we are to find the middle part E_m that contains dependency and evaluate its probability r for generating a right quartet. Once this is done, the probability of the full boomerang distinguisher of $E = \tilde{E_1} \circ E_m \circ \tilde{E_0}$ can be closely modeled as $\tilde{p}^2\tilde{q}^2 r$ where $\tilde{p}$ (resp. $\tilde{q}$) is the probability of the differential trail over $\tilde{E_0}$ (resp. $\tilde{E_1}$).

Given two differential trails over E_0 and E_1 respectively, we take the following steps to find the boundaries of E_m and calculate the probability r.

1. Extend the upper differential trail forwards with probability 1; also, extend the lower differential trail backwards with probability 1. With the extensions, it can be told whether an active S-box of the differential trail is affected by the other differential trail or not.
2. Initialize E_m with the last round of E_0 and the first round of E_1.
3. Prepend E_m with one more round,
   (a) Check whether the lower crossing differences for this newly added round are distributed uniformly or not. If yes, peel off the first round of E_m and go to step 4.
   (b) Go to step 3.
4. Append one more round to E_m,
   (a) Check whether the upper crossing differences for the newly added round are distributed uniformly or not. If yes, peel off the last round of E_m and go to step 5.
   (b) Go to step 4.
5. Calculate r using formulas in Section 3.2.

In step 4 and 5 of the algorithm, the upper boundary and the lower boundary are determined respectively, and these two steps can be swapped. If the returned r is 0, it means the two differential trails are incompatible. The time complexity of the algorithm depends on the properties of the cipher and the two differential trails of the boomerang distinguisher. We will discuss it in more details in Section 6.

4 Application to SKINNY

In [LGS17] Liu *et al.* mounted related-tweakey rectangle attacks against `SKINNY`. The attacks evaluated the probability of the distinguishers by taking into account the amplified probability but did not consider the dependency of two differential trails. In [CHP+18] which introduced `BCT`, the authors accurately evaluated the probability of generating the right quartet for two middle rounds by applying the BCT.

In this section, we revisit this issue by applying the generalized framework of `BCT` to `SKINNY`. With the generalized framework of `BCT`, we are able to identify the actual boundaries of E_m and accurately calculate the probability r for E_m. Most notably, accurate evaluations of the probability of full distinguishers of `SKINNY` become possible. The results show that there exist dependency in 5 or 6 middle rounds, which makes real probability much higher than previously evaluated.

In the remainder of this section, we first give a brief description of `SKINNY`, followed by a review of boomerang distinguishers proposed in [LGS17], for which we show how the generalized framework of `BCT` helps to evaluate the probability r. At last, an analysis of the results is added.

4.1 Description of SKINNY

SKINNY [BJK+16] is a family of lightweight block ciphers which adopt the substitution-permutation network and elements of the TWEAKEY framework [JNP14]. Members of SKINNY are denoted by SKINNY-n-t, where $n \in \{64, 128\}$ is the block size and $t \in \{n, 2n, 3n\}$ is the tweakey size. The internal states of SKINNY are represented as 4×4 arrays of cells with each cell being a nibble in case of $n = 64$ bits and a byte in case of $n = 128$ bits. The tweakey state is seen as a group of z 4×4 arrays, where, $z = t/n$. The arrays are marked as $TK1$, $(TK1, TK2)$ and $(TK1, TK2, TK3)$ for $z = 1, 2, 3$ respectively.

Encryption. SKINNY iterates a round function for N_r rounds and each round consists of the following five steps.

1. *SubCells* (SC) - A 4-bit (resp. 8-bit) S-box whose maximal differential probability is 2^{-2} is applied to all cells when n is 64 (resp. n is 128). The boomerang uniformity of the 8-bit S-box is 256.
2. *AddConstants* (AC) - This step involves XORing constants to the internal state.
3. *AddRoundTweakey* (ART) - The first two rows of the internal state absorb the first two rows of TK, where
$$TK = \bigoplus_{i=1}^{z} TK_i.$$
The tweakey states TK_i are then updated by a *tweakey scheduling algorithm*.
4. *ShiftRows* (SR) - Each cell in row j is rotated to the right by j cells.
5. *MixColumns* (MC) - Each column of the internal state is multiplied by matrix M. The inverse MixColumns operation employs M^{-1} instead. Note, the branch number of MC is only 2.
$$M = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix} \qquad M^{-1} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

Tweakey Scheduling Algorithm. The tweakey schedule of SKINNY is a linear algorithm. The t-bit tweakey is first loaded into z 4×4 tweakey states. After each ART step, the tweakey states are updated as follows.

- *Permutation Phase*: A cell-wised permutation P is applied to each tweakey state, where P is defined as: $P = [9, 15, 8, 13, 10, 14, 12, 11, 0, 1, 2, 3, 4, 5, 6, 7]$
- *LFSR Update Phase*: Cells in the first two rows of all tweakey states but TK_1 are individually updated using LFSRs.

4.2 Previous Boomerang Attacks

In [LGS17], Liu *et al.* proposed boomerang distinguishers for SKINNY-n-$2n$ and SKINNY-n-$3n$ by connecting two short differential trails $\alpha \to \beta$ and $\gamma \to \delta$. For completeness, the differential trails are copied to Table 4, 7, and 8. The probabilities of these boomerang distinguishers are evaluated by $\hat{p}^2\hat{q}^2$, where $\hat{p} = \sqrt{\Sigma_i \mathrm{Pr}^2(\alpha \to \beta_i)}$ and $\hat{q} = \sqrt{\Sigma_j \mathrm{Pr}^2(\gamma_j \to \delta)}$. Since the probabilities are too small to be verified experimentally, the authors of [LGS17] verified two middle rounds, the last round of E_0 and the first round of E_1, to exclude incompatibility between the trails.

For boomerang distinguishers, two things are of concern. The first is the compatibility of two differential trails. If the differential trails are compatible, then the second concern is the exact probability. In [CHP+18], the probabilities of the two middle rounds of these distinguishers were evaluated with BCT. However, the problem of accurately evaluating the probability of a full distinguisher remains unsolved. Next, we will show that the generalized framework of BCT provides us with the first solution to this problem.

4.3 Probability Evaluation of Boomerang Distinguishers

Here we take the boomerang distinguisher of SKINNY-128-256 as an example. Both the upper and lower differential trails have 9 rounds, as shown in Table 4. Following the algorithm in Section 3.3, we extend both differential trails with probability 1 towards each other. Around the connecting point, there are 5 rounds containing dependency, as depicted in Figure 4.

Table 4: Differential trails of SKINNY-128-256 where each non-zero cell is given in hexadecimal. The master tweakey difference is denoted by ΔK. For each round of SKINNY, input/output differences of the S-box layer in each round, as well as the the round tweakey difference are presented.

	9-round upper trail $p = 2^{-34.42}$	9-round lower trail $q = 2^{-20}$
ΔK	0,0,cc,0, 0,0,0,0, 0,0,0,0, ff,0,0,0 0,0,f3,0, 0,0,0,0, 0,0,0,0, 9f,0,0,0	fc,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 ff,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0
R1	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0a 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,3f 0,0,3f,0, 0,0,0,0	80,0,0,0, 0,0,01,0, 0,01,0,0, 01,0,0,0 03,0,0,0, 0,0,20,0, 0,20,0,0, 20,0,0,0 03,0,0,0, 0,0,0,0
R2	0,0,0,0, 0,0,3f,0, 0,0,0,0, 0,0,3f,0 0,0,0,0, 0,0,41,0, 0,0,0,0, 0,0,e3,0 0,0,0,0, 0,0,c0,0	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,20 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,80 0,0,0,0, 0,0,0,0
R3	0,e3,0,0, 0,0,0,0, 0,0,0,81, 0,0,0,0 0,2a,0,0, 0,0,0,0, 0,0,0,2a, 0,0,0,0 0,0,0,0, 2a,0,0,0	0,0,80,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,02,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,02,0, 0,0,0,0
R4	0,0,0,0, 0,2a,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,80,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,80,0,0	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0
R5	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0
R6	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0
R7	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,55,0,0	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,04,0
R8	0,0,0,0, 0,0,0,0, 0,0,55,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,01,0, 0,0,0,0 0,0,0,0, 0,0,0,01	0,0,0,0, 0,0,0,0, 0,0,0,04, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,01, 0,0,0,0 0,0,0,0, 0,0,0,0
R9	01,0,0,0, 0,0,0,0, 0,0,0,0, 01,0,0,0 20,0,0,0, 0,0,0,0, 0,0,0,0, 20,0,0,0 0,0,0,ff, 0,0,0,0	0,01,0,0, 0,0,0,0, 0,01,0,0, 0,01,0,0 0,20,0,0, 0,0,0,0, 0,20,0,0, 0,20,0,0 0,0,0,0, 0,0c,0,0

Understanding Figure 4. The vertical line in the figure separates E_0 and E_1, and the final E_m is composed of two rounds from E_0 and three rounds from E_1. The upper and lower differential trails are marked in black, while their extensions are marked in gray. Colored squares stand for active cells where "$*$" means non-zero differences, "?" means any differences and hex numbers are exact differences specified by the differential trails. For convenience, we use the row index and the column index to describe the position of a cell. Different colors are used to show the differential propagation of certain cells. For example, the cell at (2,2) after SC and AC of R1 is outlined in dark green, meaning its lower crossing difference propagates through the dark green cells from the lower differential trail.

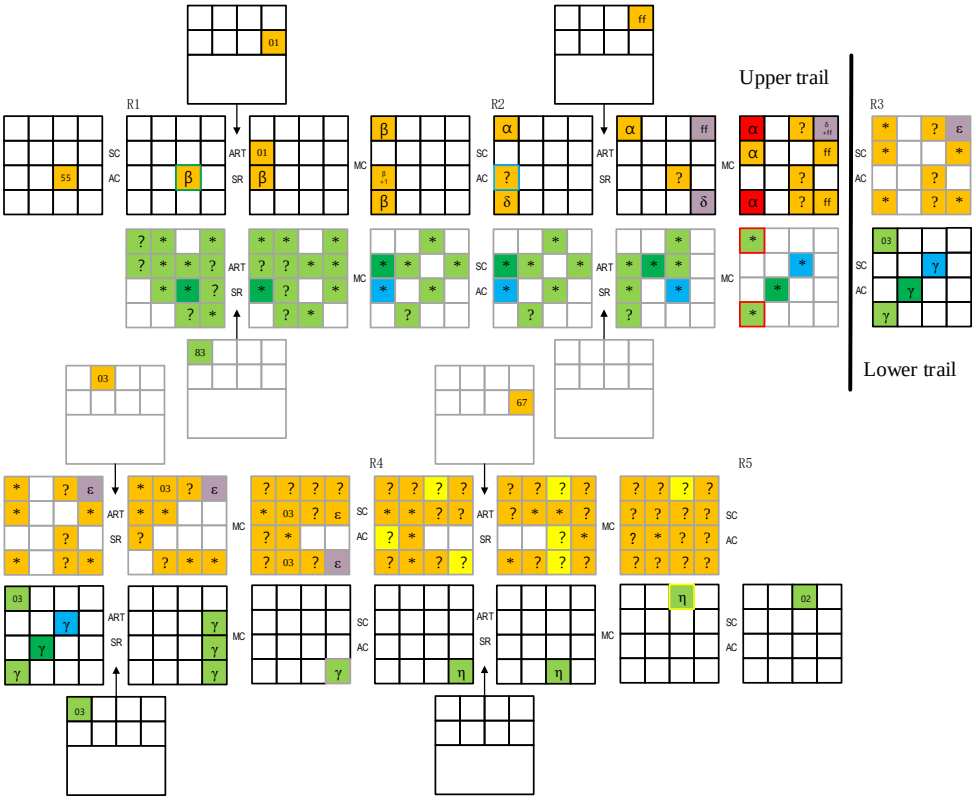

Figure 4: The five middle rounds of the 18-round boomerang distinguisher of SKINNY-128-256

Similarly, the cell at (3,3) before SC and AC of R4 is outlined in purple, meaning its upper crossing difference propagates through the purple cells from the upper differential trail.

Probability of E_m with R2 and R3. According to the upper differential trail, only two cells of the input of R2 are active and the input differences are both $\beta = 0x01$. For the lower differential trail, four cells in the first round (R3 in the figure) are active and their output differences are $0x03$, $0x20$, $0x20$ and $0x20$. From the differential trails and the extensions, it is known that the lower crossing differences of the two active S-boxes in R2 of the upper trail are always 0, while the upper crossing differences of the active S-boxes at (0,0) and (3,0) in R3 of the lower trail are α (as depicted in Figure 4) which is non-zero and propagated from the active S-box at (0,0) in R2 of the upper trail. Note that in the context of E_m with only R2 and R3, the propagation of $\beta \to \alpha$ is independent of the lower trail. By applying the generalized BCT (the exact formulas in use are indicated explicitly),

$$r = \sum_{\alpha} \frac{\mathtt{DDT}(0x01, \alpha)}{2^8} \cdot \frac{\mathtt{BCT}(\alpha, 0x03)}{2^8} \cdot \frac{\mathtt{BCT}(\alpha, 0x20)}{2^8} = 2^{-1.75} (\text{by Eq. } 8).$$

Probability of E_m with R1, R2 and R3. Now E_m starts from R_1. In this case, the analysis of the lower differential trail remains the same while the analysis for the upper differential trail changes, compared with the E_m with only R2 and R3. In the upper differential trail, only the S-box at (2,2) of R1 is active and has input difference $0x55$. However, its output difference β has multiple choices and might not be $0x01$. Therefore the S-box at (2,0) of R2 might be active. In R2, the lower crossing difference for S-box at (2,0) is the difference in blue that propagates from the lower differential trail and this difference

is independent of the upper differential trail. Another affected active S-box of the upper differential trail is the active S-box in R1. Specifically, its lower crossing difference is the difference in green propagated from S-box (2,1) in R3 of the lower differential trail through two rounds backwards. The probability of this E_m is then computed as

$$\begin{aligned}
p_0(\beta,\gamma) =& \sum_{\rho} \frac{\texttt{BCT}(\beta,\rho)\cdot\texttt{DDT}(\rho,\gamma)}{(2^8)^2} (\text{by Eq. 6}),\\
p_1(\beta,\gamma) =& \sum_{\alpha} \frac{\texttt{DDT}(\beta,\alpha)\cdot\texttt{BCT}(\alpha,0x03)\cdot\texttt{BCT}(\alpha,\gamma)}{(2^8)^3} (\text{by Eq. 8}),\\
p_2(\beta,\gamma) =& \frac{\texttt{DDT}(0x55,\beta)}{2^8}\cdot\sum_{d}\frac{\#\{y\in\mathcal{Y}_{\texttt{DDT}}(0x55,\beta): y\oplus d\in\mathcal{Y}_{\texttt{DDT}}(0x55,\beta)\}}{\#\mathcal{Y}_{\texttt{DDT}}(0x55,\beta)}\\
& \cdot \texttt{Pr}(d \xleftarrow{2 \text{ rounds}} \gamma) (\text{by Eq. 5}),\\
r =& \sum_{\beta} p_0(\beta\oplus 1, 0x20)\cdot p_1(\beta,0x20)\cdot p_2(\beta,0x20) = 2^{-6.06}.
\end{aligned}$$

Probability of E_m with R1, R2, R3 and R4. We do not prepend more rounds from E_0 with E_m since the three rounds ahead are fully passive and after propagating the lower differential trail by three more rounds backwards, the crossing differences can be seen as uniform. Then we try to append more rounds from E_1. First, we append R4 to E_m. The only active S-box in R4 is located at (3,3) and its upper crossing difference propagates through purple cells from the upper differential trail. We can see that the upper and lower differential trails strongly interrelate. However, compared with the previous situation where E_m is composed of R1, R2 and R3, the only extra effect is that the upper crossing difference in purple affects S-box (3,3) of R3. Still, we can calculate

$$\begin{aligned}
p_3(\beta,\gamma,\eta) =& \frac{\texttt{DDT}(\gamma,\eta)}{2^8}\cdot\sum_{d}\frac{\#\{x\in\mathcal{X}_{\texttt{DDT}}(\gamma,\eta): x\oplus d\in\mathcal{X}_{\texttt{DDT}}(\gamma,\eta)\}}{\#\mathcal{X}_{\texttt{DDT}}(\gamma,\eta)}\\
& \cdot \texttt{Pr}(\beta \xrightarrow[0xff]{2 \text{ rounds}} d) (\text{by Eq. 7}),\\
r =& \sum_{\gamma}\sum_{\beta} p_0(\beta\oplus 1,\gamma)\cdot p_1(\beta,\gamma)\cdot p_2(\beta,\gamma)\cdot p_3(\beta,\gamma,0x80) (\text{by Eq. 9})\\
=& 2^{-9.54}.
\end{aligned}$$

Probability of E_m with R1, R2, R3, R4 and R5. In fact, the upper crossing differences at R5 becomes so random that we can neglect the dependency in R5. However, due to the weak diffusion of the MC, the difference of the only active S-box in R4 of the lower trail does not diffuse to more cells. Actually, the output differences η of the active S-box in R4 on the two faces of the boomerang do not have to be identical. Considering this, we calculate

$$\begin{aligned}
p_4(\beta,\gamma,\eta_1,\eta_2) =& \sum_{d}\frac{\#\{x\in\mathcal{X}_{\texttt{DDT}}(\gamma,\eta_1): x\oplus d\in\mathcal{X}_{\texttt{DDT}}(\gamma,\eta_2)\}}{\#\mathcal{X}_{\texttt{DDT}}(\gamma,\eta_1)}\cdot\texttt{Pr}(\beta \xrightarrow[0xff]{2 \text{ rounds}} d),\\
p_5(\gamma,\eta_1,\eta_2) =& \frac{\texttt{DDT}(\gamma,\eta_1)\cdot\texttt{DDT}(\eta_1,0x02)\cdot\texttt{DDT}(\eta_2,0x02)}{(2^8)^3},\\
r =& \sum_{\eta_1}\sum_{\eta_2}\sum_{\gamma}\sum_{\beta} p_0(\beta\oplus 1,\gamma)\cdot p_1(\beta,\gamma)\cdot p_2(\beta,\gamma)\\
& \cdot p_4(\beta,\gamma,\eta_1,\eta_2)\cdot p_5(\gamma,\eta_1,\eta_2) = 2^{-11.45}.
\end{aligned}$$

Now we have identified the middle part E_m of the boomerang distinguisher of SKINNY-128-256. The E_m has 5 rounds and its probability of generating a right quartet is $r =$

$2^{-11.45}$. The probabilities of intermediate E_m with $2 \sim 4$ rounds and the final E_m with 5 rounds are confirmed by experiments. By adding three passive rounds to the front and four passive rounds to the rear, we obtain a 12-round boomerang distinguisher of the same probability, namely $2^{-11.45}$. For the full 18-round distinguisher, the probability of the first four rounds is $\tilde{p} = 2^{-25.19}$ by a simple calculation considering the clustering effect of differentials. The probability of the last two rounds is $\tilde{q} = 2^{-8}$ (no clustering effect). Therefore, the probability of the full distinguisher is $\tilde{p}^2\tilde{q}^2 r = 2^{-77.83}$, which is much higher than $2^{-103.84}$ calculated in [LGS17]. For other versions of SKINNY, a similar analysis can be done and we summarize the result as follows.

4.4 Results

The results of all the four versions of SKINNY-n-$2n$ and SKINNY-n-$3n$ are summarized in Table 1, where the fourth column presents the probabilities r of E_m computed by the algorithm in Section 3.3. We carry out experiments on E_m and the experimental probabilities are $2^{-12.95}, 2^{-11.37}, 2^{-10.51}$ and $2^{-9.89}$, which are close to the probabilities in the fourth column. The computation of r for the four versions is practical and takes 0.38, 189.26, 0.11 and 23.16 seconds respectively on a desktop. The source codes for calculating and verifying the probabilities r for E_m are available online[1].

The sixth column stands for the probabilities of the full boomerang distinguishers. It can be seen that the re-evaluated probabilities of the full distinguishers are much higher than the probabilities $\hat{p}^2\hat{q}^2$ evaluated before without considering the dependency of the two differential trails [LGS17]. Notably, the complexity of the full 17-round distinguisher of SKINNY-64-128 is $2^{29.78}$, which is practical. Indeed, 9 right quartets are found among 11×2^{29} quartets by an experiment while one could expect a right one in $2^{48.72}$ quartets according to [LGS17]. This big gap shows that the issue of dependency cannot be neglected in boomerang attacks.

Additionally, we have two interesting observations of the results. One observation is that the probabilities of the boomerang distinguishers of SKINNY in Table 1 are much higher than the probabilities of the differential trails of the same number of rounds. In Table 5, we copy the lowerbounds on the number of active S-boxes in SKINNY under the related-tweakey setting from [BJK+16]. For example, the minimal number of active S-boxes of 9-round SKINNY-n-$2n$ is 9, which means the probability of optimal differential trails could not be higher than 2^{-18}. Actually, the probability of the optimal differential trails of SKINNY-64-128 is 2^{-20}, as studied in [LGS17]. The differential trail in Table 7 is an example of the optimal differential trails. This probability can be increased to 2^{-18} by considering the clustering effect[2]. On the contrary, the boomerang distinguisher of SKINNY-64-128 with 6 up to 13 rounds has a much higher probability of $2^{-12.96}$.

Table 5: Lowerbounds on the number of active S-boxes in SKINNY under the related-tweakey setting [BJK+16]

#Rounds	9	10	11	12	13	14	15	16	17
SKINNY-n-$2n$	9	12	16	21	25	31	35	40	43
SKINNY-n-$3n$	3	6	10	13	16	19	24	27	31

The other observation of Table 1 is that the probability of the 17-round distinguisher of SKINNY-128-384 is slightly higher than the probability of the 17-round distinguisher of SKINNY-64-192. Even though an 8-bit S-box is used in the big versions of SKINNY, its optimal differential probability is 2^{-2} which is the same as the optimal differential

[1] https://drive.google.com/file/d/1cd9lNruJrHhUM2QIvI-XXKK7b8LNOgXv/view?usp=sharing

[2] There only exist 4 trails (each of probability 2^{-20}) when the input, output and key differences are fixed.

probability of the 4-bit S-box used in the small versions. Therefore, the probability of a boomerang distinguisher of the big versions is not necessarily lower than the probability of distinguishers of the small versions.

5 Application to AES

In [BK09], Biryukov *et al.* presented boomerang attacks on full AES-192 and AES-256 under the related-key setting, specifically, the related-subkey setting. Their attacks were based on high probability boomerang distinguishers constructed by applying the so-called boomerang switches. However, boomerang attacks or distinguishers of AES-128 were not covered in [BK09]. One reason might be that the boomerang switches do no work for AES-128 whose differences of a differential trail are much denser than those of AES-192 and AES-256.

In this section, we briefly review the specification of AES-128, and then search for differential trails of AES-128 under related-key setting. By choosing a pair of 3-round differential trails according to the generalized framework of BCT, we construct the 6-round boomerang distinguisher. The probability of the boomerang distinguisher is $2^{-109.42}$ under the related-subkey setting.

5.1 Description of AES

The Advanced Encryption Standard (AES) [DR02] is an iterated block cipher which encrypts 128-bit plaintexts with secret keys of sizes 128, 192, and 256 bits. In this paper, we focus on AES-128 which iterates 10 rounds using a 128-bit key. The internal state of AES can be represented as a 4×4 array of bytes. The round function consists of four basic steps as follows.

1. *SubBytes* (SB) - An 8-bit S-box is applied to each byte of the internal state. The details of the S-box could be found in [DR02].

2. *ShiftRows* (SR) - Each cell in row j is rotated to the left by j cells.

3. *MixColumns* (MC) - Each column of the internal state is multiplied by a Maximum Distance Separable (MDS) matrix over $\mathbb{F}_{2^8}$.

4. *AddRoundKey* (AK) - A round key is XORed with the internal state.

At the very beginning of the encryption, an additional whitening key addition is performed, and the last round does not contain MC.

The key schedule of AES-128 generates round keys which are used in each of the rounds. The 128-bit master key can be seen as 4 32-bit words $(W[0], W[1], W[2], W[3])$. Then $W[i]$ for $i \geqslant 4$ is computed as follows and each round key takes four consecutive words.

$$W[i] = \begin{cases} W[i-4] \oplus \mathtt{SB}(W[i-1] \lll 8) \oplus Rcon & i \equiv 0 \bmod 4, \\ W[i-8] \oplus W[i-1] & \text{otherwise} \end{cases}$$

Property of AES S-box. The best differential probability of AES S-box is 2^{-6}. Given any input difference α, there exists exactly one output difference β such that $\mathtt{DDT}(\alpha, \beta) = 4$ and $2^7 - 1$ output differences β such that $\mathtt{DDT}(\alpha, \beta) = 2$, and vice versa. Its boomerang uniformity is 6.

5.2 Search for Differential Trails

In the literature, several methods have been proposed to search differential trails for AES under the related-key setting, such as [BN10, GMS16, SGL+17, GLMS18]. The methods in [GMS16, GLMS18] employ Constraint Programming (CP) and perform well for all versions of AES. Therefore, we adopt the CP-based methods for searching differential trails of AES-128.

As pointed out in [GLMS18], the minimal number of active S-boxes in three rounds of AES-128 under the related-key setting is 5. When we increase the number of rounds to four, the minimal number of active S-boxes becomes 12. Based on this, we aim to construct 6-round boomerang distinguishers of AES-128 from 3-round differential trails.

5.3 Boomerang Distinguisher

There exist only two 3-round differential trails with 5 active S-boxes and both have probability 2^{-31}. However, these two differential trails turn out to be incompatible, *i.e.*, the dependent part E_m of two differential trails could not generate a right quartet. How hard can we find a pair of compatible differential trails? According to the properties of the AES S-box, we give a rough estimation as follows.

For an active S-box in E_0 (resp. E_1) whose lower (resp. upper) crossing difference is non-zero and fixed, the differences are compatible (the BCT entry is greater than 0) with probability close to 2^{-1}. For a pair of interrelated active S-boxes, the differences are compatible (the probability in Eq. 9 is greater than 0) with probability close to 2^{-2}. Therefore, the sparser the differences are, the more easily we can get a pair of compatible differential trails.

We then search for more 3-round differential trails by allowing 6 active S-boxes and obtain 18 differential trails with probability $2^{-36}, 2^{-37}$ or 2^{-38} respectively. From these 3-round differential trails, we search for a compatible combination by the generalized framework of BCT. The best one we find is composed of a 3-round upper differential trail of probability 2^{-31} and a 3-round lower differential trail of probability 2^{-37}, as shown in Table 6.

Table 6: 6-round related-subkey boomerang distinguisher of AES-128 with probability $2^{-109.42}$

Round	Before AK	Subkey diff.	Before SB	After SB	After SR	p_r
R1	8c 1f 8c 00 01 99 01 00 8d 00 8d c2 37 00 8d 00	8c 00 8c 00 01 00 01 00 8d 00 8d 00 8d 00 8d 00	00 1f 00 00 00 99 00 00 00 00 00 c2 ba 00 00 00	00 a3 00 00 00 8d 00 00 00 00 00 46 97 00 00 00	00 a3 00 00 8d 00 00 00 00 46 00 00 00 97 00 00	$(2^{-6})^8$
R2	8c 8c 00 00 01 fe 00 00 8d 8d 00 00 8d 8d 00 00	8c 8c 00 00 01 01 00 00 8d 8d 00 00 8d 8d 00 00	00 00 00 00 00 ed 00 00 00 00 00 00 00 00 00 00	00 00 00 00 00 8d 00 00 00 00 00 00 00 00 00 00	00 00 00 00 8d 00 00 00 00 00 00 00 00 00 00 00	$(2^{-7})^2$
R3	8c 00 00 00 01 00 00 00 8d 00 00 00 8d 00 00 00	8c 00 00 00 01 00 00 00 8d 00 00 00 8d 00 00 00	00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00	1
R4	0a 87 0a 00 0c bc f6 00 06 00 06 fb 23 00 06 00	0a 00 0a 00 0c 00 0c 00 06 00 06 00 06 00 06 00	00 87 00 00 00 bc fa 00 00 00 00 fb 19 00 00 00	00 74 00 00 00 06 4e 00 00 00 00 6c 5c 00 00 00	00 74 00 00 06 4e 00 00 00 6c 00 00 00 5c 00 00	$2^{-33.42}$
R5	0a 0a 00 00 0c 00 00 00 06 06 00 00 06 06 00 00	0a 0a 00 00 0c 0c 00 00 06 06 00 00 06 06 00 00	00 00 00 00 00 0c 00 00 00 00 00 00 00 00 00 00	00 00 00 00 00 06 00 00 00 00 00 00 00 00 00 00	00 00 00 00 06 00 00 00 00 00 00 00 00 00 00 00	$(2^{-7})^2$
R6	0a 00 00 00 0c 00 00 00 06 00 00 00 06 00 00 00	0a 00 00 00 0c 00 00 00 06 00 00 00 06 00 00 00	00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00	00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00	1

The dependent part E_m of the 6-round boomerang distinguisher of AES-128 contains

only two rounds, as depicted in Figure 5. The probability of generating a right quartet is

$$r = \frac{\texttt{BCT}(0x8c, 0x74) \cdot \texttt{BCT}(0x01, 0x06) \cdot \texttt{BCT}(0x01, 0x4e) \cdot \texttt{BCT}(0x8d, 0x6c) \cdot \texttt{BCT}(0x8d, 0x5c)}{(2^8)^5}$$
$$= \frac{2 \cdot 6 \cdot 2 \cdot 2 \cdot 2}{(2^8)^5} = 2^{-33.42}.$$

Therefore, the probability for the 6-round boomerang distinguisher is

$$\tilde{p}^2 \tilde{q}^2 r = 2^{-31 \times 2} 2^{-7 \times 2} 2^{-33.42} = 2^{-109.42}.$$

The probability of the 2-round E_m is verified by experiments.

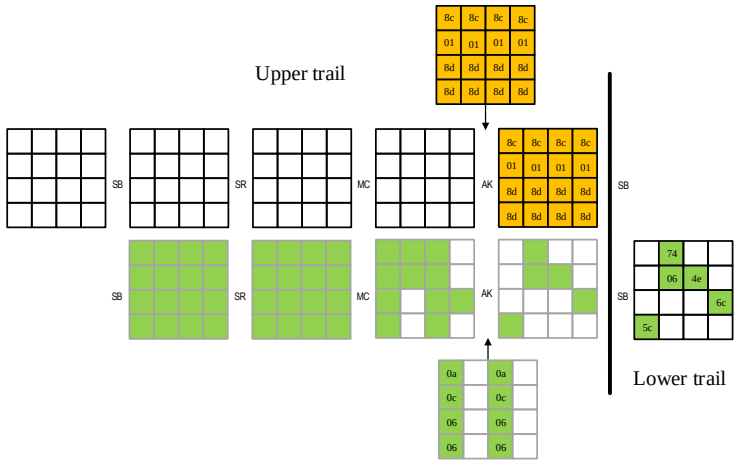

Figure 5: The two middle rounds of the 6-round boomerang distinguisher of AES-128

6 Discussions

As showcases, we apply our generalized framework of BCT to SKINNY and AES. Both SKINNY and AES are SPN block ciphers and share similar round functions. However, AES and SKINNY are typical examples of block ciphers with very strong and weak round functions respectively. Specifically, the AES S-box is differentially 4-uniform (6-uniform for BCT) and the AES MC has a branch number of 5. On the contrary, the SKINNY's 8-bit S-box is differential 64-uniform (256-uniform for BCT) and the branch number of its MC is only 2. Together with the analysis of the two block ciphers, we summarize two general properties of the dependent part E_m of the boomerang distinguisher.

Property 1 *The length of E_m is mainly determined by the diffusion effect of the linear layer*, even though it is also influenced by the density of differences of the trails. Note that, AES takes 2 rounds to diffuse an active byte to the full state while SKINNY takes 6 rounds to have the same effect. Compared with the 2-round E_m of AES, the E_m of SKINNY is quite long and can be 6 rounds, which can be seen from the analysis in Section 4 and 5.

Property 2 *The probability of E_m is strongly affected by the* DDT *and* BCT *of the S-box.* For example, when we replace SKINNY-128-256's S-box in Figure 4 with the AES S-box, the probabilities of E_m with two and three middle rounds decrease from $2^{-1.75}, 2^{-6.06}$ to $2^{-15.87}, 2^{-31.67}$ respectively.

As can be seen, these properties are identical to the common criteria for designing symmetric-key primitives.

The time complexity of calculating the probability r of E_m mainly depends on the length of E_m and the S-box used in the cipher. Specifically, for a short E_m with small or weak S-boxes, calculating r is efficient, while for a long E_m with large and strong S-boxes, calculating r might be a time-consuming task, *i.e.*, the time complexity might be greater than 2^{35}.

7 Concluding Remarks

In this paper, we revisited the boomerang connectivity table (BCT) and provided a generalized framework of BCT which systematically handles the dependency of two differential trails in boomerang distinguishers. Particularly, our framework not only identifies the actual boundaries of the dependent part E_m of the boomerang distinguisher, but also calculates the probability r of E_m for generating a right quartet. With our generalized framework of BCT, the sandwich $E = \tilde{E}_1 \circ E_m \circ \tilde{E}_0$ now closely models the boomerang distinguisher with probability $\tilde{p}^2\tilde{q}^2 r$ where $\tilde{p}$ (resp. $\tilde{q}$) is the probability of the differential of $\tilde{E}_0$ (resp. $\tilde{E}_1$).

The power of the generalized framework of BCT was demonstrated by the application to SKINNY and AES. In the application to SKINNY, the probabilities of four boomerang distinguishers of SKINNY were accurately computed for the first time, which show that the actual probabilities are much higher than those previously computed by the formula $\hat{p}^2\hat{q}^2$. In the application to AES, a 6-round related-subkey boomerang distinguisher was constructed with the generalized framework of BCT.

We also discussed the general relation between the dependency of two differential trails in a boomerang distinguisher and the properties of the components of the cipher, and showed that the dependency is strongly influenced by both the diffusion property of the linear layer and differential properties of the non-linear layer.

Acknowledgments

The authors would like to thank the anonymous referees for their helpful comments. The authors are supported by NTU research grant M4080456 and the National Natural Science Foundation of China (Grants No. 61802399, 61802400, 61732021 and 61772519), the Youth Innovation Promotion Association CAS, and Chinese Major Program of National Cryptography Development Foundation (Grant No. MMJJ20180102).

References

[ALLW14] Farzaneh Abed, Eik List, Stefan Lucks, and Jakob Wenzel. Differential Cryptanalysis of Round-Reduced Simon and Speck. In Carlos Cid and Christian Rechberger, editors, *Fast Software Encryption - 21st International Workshop, FSE 2014, London, UK, March 3-5, 2014. Revised Selected Papers*, volume 8540 of *Lecture Notes in Computer Science*, pages 525–545. Springer, 2014.

[BC18] Christina Boura and Anne Canteaut. On the Boomerang Uniformity of Cryptographic Sboxes. *IACR Transactions on Symmetric Cryptology*, 2018(3):290–310, Sep. 2018.

[BDK01] Eli Biham, Orr Dunkelman, and Nathan Keller. The Rectangle Attack - Rectangling the Serpent. In Birgit Pfitzmann, editor, *Advances in Cryptology - EUROCRYPT 2001, International Conference on the Theory and Application*

of Cryptographic Techniques, Innsbruck, Austria, May 6-10, 2001, Proceeding, volume 2045 of *Lecture Notes in Computer Science*, pages 340–357. Springer, 2001.

[BDK05] Eli Biham, Orr Dunkelman, and Nathan Keller. Related-Key Boomerang and Rectangle Attacks. In Ronald Cramer, editor, *Advances in Cryptology - EUROCRYPT 2005, 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Aarhus, Denmark, May 22-26, 2005, Proceedings*, volume 3494 of *Lecture Notes in Computer Science*, pages 507–525. Springer, 2005.

[BJK+16] Christof Beierle, Jérémy Jean, Stefan Kölbl, Gregor Leander, Amir Moradi, Thomas Peyrin, Yu Sasaki, Pascal Sasdrich, and Siang Meng Sim. The SKINNY Family of Block Ciphers and Its Low-Latency Variant MANTIS. In Matthew Robshaw and Jonathan Katz, editors, *Advances in Cryptology - CRYPTO 2016 - 36th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 14-18, 2016, Proceedings, Part II*, volume 9815 of *Lecture Notes in Computer Science*, pages 123–153. Springer, 2016.

[BK09] Alex Biryukov and Dmitry Khovratovich. Related-Key Cryptanalysis of the Full AES-192 and AES-256. In Mitsuru Matsui, editor, *Advances in Cryptology - ASIACRYPT 2009, 15th International Conference on the Theory and Application of Cryptology and Information Security, Tokyo, Japan, December 6-10, 2009. Proceedings*, volume 5912 of *Lecture Notes in Computer Science*, pages 1–18. Springer, 2009.

[BN10] Alex Biryukov and Ivica Nikolic. Automatic Search for Related-Key Differential Characteristics in Byte-Oriented Block Ciphers: Application to AES, Camellia, Khazad and Others. In Henri Gilbert, editor, *Advances in Cryptology - EUROCRYPT 2010, 29th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Monaco / French Riviera, May 30 - June 3, 2010. Proceedings*, volume 6110 of *Lecture Notes in Computer Science*, pages 322–344. Springer, 2010.

[BS93] Eli Biham and Adi Shamir. *Differential Cryptanalysis of the Data Encryption Standard*. Springer, 1993.

[CHP+18] Carlos Cid, Tao Huang, Thomas Peyrin, Yu Sasaki, and Ling Song. Boomerang Connectivity Table: A New Cryptanalysis Tool. In Jesper Buus Nielsen and Vincent Rijmen, editors, *Advances in Cryptology - EUROCRYPT 2018 - 37th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tel Aviv, Israel, April 29 - May 3, 2018 Proceedings, Part II*, volume 10821 of *Lecture Notes in Computer Science*, pages 683–714. Springer, 2018.

[CLN+17] Anne Canteaut, Eran Lambooij, Samuel Neves, Shahram Rasoolzadeh, Yu Sasaki, and Marc Stevens. Refined Probability of Differential Characteristics Including Dependency Between Multiple Rounds. *IACR Trans. Symmetric Cryptol.*, 2017(2):203–227, 2017.

[DKS10] Orr Dunkelman, Nathan Keller, and Adi Shamir. A Practical-Time Related-Key Attack on the KASUMI Cryptosystem Used in GSM and 3G Telephony. In Tal Rabin, editor, *Advances in Cryptology - CRYPTO 2010, 30th Annual Cryptology Conference, Santa Barbara, CA, USA, August 15-19, 2010. Proceedings*, volume 6223 of *Lecture Notes in Computer Science*, pages 393–410. Springer, 2010.

[DKS14] Orr Dunkelman, Nathan Keller, and Adi Shamir. A Practical-Time Related-Key Attack on the KASUMI Cryptosystem Used in GSM and 3G Telephony. *J. Cryptology*, 27(4):824–849, 2014.

[DR02] Joan Daemen and Vincent Rijmen. *The Design of Rijndael: AES - The Advanced Encryption Standard*. Information Security and Cryptography. Springer, 2002.

[GLMS18] David Gérault, Pascal Lafourcade, Marine Minier, and Christine Solnon. Revisiting AES Related-Key Differential Attacks with Constraint Programming. *Inf. Process. Lett.*, 139:24–29, 2018.

[GMS16] David Gerault, Marine Minier, and Christine Solnon. Constraint Programming Models for Chosen Key Differential Cryptanalysis. In Michel Rueher, editor, *Principles and Practice of Constraint Programming - 22nd International Conference, CP 2016, Toulouse, France, September 5-9, 2016, Proceedings*, volume 9892 of *Lecture Notes in Computer Science*, pages 584–601. Springer, 2016.

[JNP14] Jérémy Jean, Ivica Nikolic, and Thomas Peyrin. Tweaks and Keys for Block Ciphers: The TWEAKEY Framework. In Palash Sarkar and Tetsu Iwata, editors, *Advances in Cryptology - ASIACRYPT 2014 - 20th International Conference on the Theory and Application of Cryptology and Information Security, Kaoshiung, Taiwan, R.O.C., December 7-11, 2014, Proceedings, Part II*, volume 8874 of *Lecture Notes in Computer Science*, pages 274–288. Springer, 2014.

[KKS00] John Kelsey, Tadayoshi Kohno, and Bruce Schneier. Amplified boomerang attacks against reduced-round MARS and serpent. In Bruce Schneier, editor, *Fast Software Encryption, 7th International Workshop, FSE 2000, New York, NY, USA, April 10-12, 2000, Proceedings*, volume 1978 of *Lecture Notes in Computer Science*, pages 75–93. Springer, 2000.

[LGS17] Guozhen Liu, Mohona Ghosh, and Ling Song. Security Analysis of SKINNY under Related-Tweakey Settings. *IACR Trans. Symmetric Cryptol.*, 2017(3):37–72, 2017.

[Mur11] Sean Murphy. The Return of the Cryptographic Boomerang. *IEEE Trans. Information Theory*, 57(4):2517–2521, 2011.

[SGL+17] Siwei Sun, David Gerault, Pascal Lafourcade, Qianqian Yang, Yosuke Todo, Kexin Qiao, and Lei Hu. Analysis of AES, SKINNY, and Others with Constraint Programming. *IACR Trans. Symmetric Cryptol.*, 2017(1):281–306, 2017.

[Wag99] David A. Wagner. The Boomerang Attack. In Lars R. Knudsen, editor, *Fast Software Encryption, 6th International Workshop, FSE '99, Rome, Italy, March 24-26, 1999, Proceedings*, volume 1636 of *Lecture Notes in Computer Science*, pages 156–170. Springer, 1999.

A Differential Trails

Table 7: Differential trails of SKINNY-64-128 [LGS17]

	8-round upper trail $p = 2^{-12}$	9-round lower trail $q = 2^{-20}$
ΔK	0,0,0,0, 0,0,0,0, 6,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 9,0,0,0, 0,0,0,0	0,0,c,0, 0,0,0,0, 0,0,0,0, e,0,0,0 0,0,f,0, 0,0,0,0, 0,0,0,0, b,0,0,0
R1	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,1 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,8 0,0,0,0, 0,0,0,0	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,2 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,3 0,0,3,0, 0,0,0,0
R2	0,0,8,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,5,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,5,0, 0,0,0,0	0,0,0,0, 0,0,3,0, 0,0,0,0, 0,0,3,0 0,0,0,0, 0,0,d,0, 0,0,0,0, 0,0,c,0 0,0,0,0, 0,0,9,0
R3	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0	0,c,0,0, 0,0,0,0, 0,0,0,4, 0,0,0,0 0,2,0,0, 0,0,0,0, 0,0,0,2, 0,0,0,0 0,0,0,0, 2,0,0,0
R4	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0	0,0,0,0, 0,2,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,1,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,1,0,0
R5	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0
R6	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,b,0	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0
R7	0,0,0,0, 0,0,0,0, 0,0,0,b, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,1, 0,0,0,0 0,0,0,0, 0,0,0,0	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,4,0,0
R8	0,1,0,0, 0,0,0,0, 0,1,0,0, 0,1,0,0 0,8,0,0, 0,0,0,0, 0,8,0,0, 0,8,0,0 0,0,0,0, 0,c,0,0	0,0,0,0, 0,0,0,0, 0,0,4,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,2,0, 0,0,0,0 0,0,0,0, 0,0,0,2
R9		2,0,0,0, 0,0,0,0, 0,0,0,0, 2,0,0,0 6,0,0,0, 0,0,0,0, 0,0,0,0, 5,0,0,0 0,0,0,d, 0,0,0,0

Table 8: Differential trails of SKINNY-n-$3n$ [LGS17]

	11-round trail for SKINNY-64 $p = 2^{-20}$	11-round trail for SKINNY-128 $q = 2^{-21}$
ΔK	0,a,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,2,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,d,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0	0,aa,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,e6,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,cf,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0
R1	0,2,0,0, 1,0,0,0, 0,0,0,1, 0,0,1,0 0,5,0,0, b,0,0,0, 0,0,0,b, 0,0,b,0 0,5,0,0, 0,0,0,0	0,20,0,0, 10,0,0,0, 0,0,0,10, 0,0,10,0 0,83,0,0, 40,0,0,0, 0,0,0,40, 0,0,40,0 0,83,0,0, 0,0,0,0
R2	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,b,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,1,0,0 0,0,0,0, 0,0,0,0	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,40,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,04,0,0 0,0,0,0, 0,0,0,0
R3	1,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 8,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 8,0,0,0, 0,0,0,0	04,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 01,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 01,0,0,0, 0,0,0,0
R4	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0
R5	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0
R6	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0
R7	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0
R8	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0
R9	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,8,0	0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,0, 0,0,0,0 0,0,0,0, 0,0,01,0
R10	0,0,0,0, 0,0,0,0, 0,0,0,8, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,4, 0,0,0,0 0,0,0,0, 0,0,0,0	0,0,0,0, 0,0,0,0, 0,0,0,01, 0,0,0,0 0,0,0,0, 0,0,0,0, 0,0,0,20, 0,0,0,0 0,0,0,0, 0,0,0,0
R11	0,4,0,0, 0,0,0,0, 0,4,0,0, 0,4,0,0 0,2,0,0, 0,0,0,0, 0,2,0,0, 0,2,0,0 0,0,0,0, 0,5,0,0	0,20,0,0, 0,0,0,0, 0,20,0,0, 0,20,0,0 0,80,0,0, 0,0,0,0, 0,80,0,0, 0,80,0,0 0,0,0,0, 0,83,0,0