Refined Probability of Differential Characteristics Including Dependency Between Multiple Rounds

Anne Canteaut[1], Eran Lambooij[2], Samuel Neves[3], Shahram Rasoolzadeh[4], Yu Sasaki[5] and Marc Stevens[6]

[1] Inria, France, Anne.Canteaut@inria.fr
[2] Technische Universiteit Eindhoven, The Netherlands, e.lambooij@student.tue.nl
[3] CISUC, Dept. of Informatics Engineering, University of Coimbra, Portugal, sneves@dei.uc.pt
[4] Ruhr-Universität Bochum, Germany, Rasoolzadeh.shahram@gmail.com
[5] NTT Secure Platform Laboratories, Japan, sasaki.yu@lab.ntt.co.jp
[6] CWI Amsterdam, The Netherlands, marc.stevens@cwi.nl

Abstract. The current paper studies the probability of differential characteristics for an unkeyed (or with a fixed key) construction. Most notably, it focuses on the gap between two probabilities of differential characteristics: probability with independent S-box assumption, p_{ind}, and exact probability, p_{exact}. It turns out that p_{exact} is larger than p_{ind} in Feistel network with some S-box based inner function. The mechanism of this gap is then theoretically analyzed. The gap is derived from interaction of S-boxes in three rounds, and the gap depends on the size and choice of the S-box. In particular the gap can never be zero when the S-box is bigger than six bits. To demonstrate the power of this improvement, a related-key differential characteristic is proposed against a lightweight block cipher RoadRunneR. For the 128-bit key version, p_{ind} of 2^{-48} is improved to p_{exact} of 2^{-43}. For the 80-bit key version, p_{ind} of 2^{-68} is improved to p_{exact} of 2^{-62}. The analysis is further extended to SPN with an almost-MDS binary matrix in the core primitive of the authenticated encryption scheme Minalpher: p_{ind} of 2^{-128} is improved to p_{exact} of 2^{-96}, which allows to extend the attack by two rounds.

Keywords: differential cryptanalysis · independent S-box · fixed key · unkeyed construction · exact probability · RoadRunneR · Minalpher

1 Introduction

Differential cryptanalysis [BS90, BS93] is one of the most fundamental cryptanalytic approaches targeting symmetric-key primitives. While its basic concept in an idealized environment under several assumptions can easily be understood, predicting the actual behavior of concrete algorithms is quite complex and a lot of research has been done regarding this topic.

Most block ciphers are designed to iterate a small keyed permutation, called the round function, with many rounds being performed to build a conversion between the plaintext and ciphertext. The plaintext x_0 is updated by round function RF_i in the ith round by processing $x_{i+1} \leftarrow RF_i(x_i)$ for $i = 0, 1, 2, \cdots$. The most common approach for evaluating the effect of differential analysis consists in applying the Markov assumption to the cipher [LMM91] and evaluating the probability of differential propagation for each round. The probability of the differential characteristic over the entire cipher is then equal to the product of the probabilities of the differentials of all rounds.

IACR Transactions on Symmetric Cryptology Vol. 2017, No.2, pp.203—227,
DOI:10.13154/tosc.v2017.i2.203-227

Given a pair of differences (a_i, a_{i+1}), the corresponding probability $p_i \triangleq \Pr_{x\in\mathbf{P}}[RF_i(x)\oplus RF_i(x \oplus a_i) = a_{i+1}]$ is searched for each i, where $\mathbf{P}$ is the plaintext space, and $\Pi_i p_i$ is the probability of the characteristic $(a_0, a_1, \ldots, a_r)$ for the entire r-round cipher.

The hidden argument in the above explanation is the treatment of a key k or subkeys k_i. The Markov assumption can be established when the state x_i is first xored with a subkey k_i and all subkeys are chosen independently uniformly at random. Therefore, most analyses are based on bounds on the *expected* probability of a differential characteristic, i.e., the probability averaged over all keys. However, the implementation environment for symmetric-key primitives does not allow to store all independent subkeys, thus k_i is usually expanded from k, and the Markov assumption collapses.

Moreover, subkeys may not be xored in every round to all state bits, which can be seen in designs of lightweight cryptographic schemes such as Simon [BSS+13], SKINNY [BJK+16] and LED [GPPR11]. Also some primitives, like hash functions or Even-Mansour schemes [DKS12, EM91, EM97], are based on an iterated permutation which does not involve any key at all. In such a case, the evaluation using the Markov assumption may still give some insight about the security against differential analysis, but never leads to the exact probability of the differential propagation for multiple rounds.

To conclude, evaluating the probability of differential propagations for multiple rounds precisely without the Markov assumption is a big challenge.

1.1 Related Work on Precise Evaluation of Differential Probability

Our work then focuses on the evaluation of the probability of a differential characteristic for a primitive with a fixed key, or for a keyless primitive. It is worth noticing that both contexts are similar in the sense that the absence of a key can equivalently be seen as the insertion of an all-zero key. Conversely, a structure with a fixed key is equivalent to an unkeyed one with different building blocks. For instance, using an S-box S with a fixed round-key k is equivalent to using $S' : x \mapsto S_k(x)$ as an S-box without any key. Let E be a block cipher with a fixed key and let ΔP and ΔC be the plaintext and ciphertext differences, respectively. Suppose that the goal is to precisely evaluate the probability of $\Pr[E(x) \oplus E(x \oplus \Delta P) = \Delta C]$, where the probability is taken over all plaintexts x. Besides the issue of subkeys for multiple rounds, there are several aspects to precisely evaluate this probability.

The first issue we would like to mention is the contrast between differential characteristics and differential effect. The differential characteristics specify not only $(\Delta P, \Delta C)$ but also differences in intermediate states, often the initial difference in each round, and evaluate the probability of each section and multiplies all the probabilities. On the contrary, the differential effect sums up the probabilities of all possible differential characteristics, thus gives a more precise probability. A lot of research has been done to evaluate the exact maximum expected differential probability (and the maximum expected linear potential) in particular for AES, e.g. [HLL+00, KMT01, PSC+02, PSLL03, DR06, KS07, CR15], and for Feistel or MISTY networks, e.g. [NK92, Mat96]. Those researches are different from the current paper with respect to the point that all state bits are xored by subkeys which are assumed to be chosen independently and uniformly at random.

In contrary, our work focuses on determining the exact probability of a differential characteristic when the key is fixed. This fixed-key probability has been determined in a very few cases only. The most prominent example is the AES, for which the probabilities of 2-round characteristics have been determined, for all possible values of the key [DR07].

Alternative approaches can be used when such a theoretical analysis is out of reach. One approach is carrying out some experiment, which exhaustively chooses plaintexts $P \in \mathbf{P}$ and actually computes $E_K(x) \oplus E_K(x \oplus \Delta P)$. The experiment is then iterated for several keys (see e.g. [BG10]). The experiment can include any complex event, however, the lack of theoretical analysis limits its versatility to be applied to other ciphers. Of

course the approach can only be applied to ciphers with small block sizes, often 32-bit block sizes, such as SIMON and KATAN [DDGS15, CDK09]. Another approach introduced in [BBL13] consists in computing the maximal expected probability of a characteristic and deriving a bound on the probability of the existence of characteristics whose fixed-key probability exceeds a given value. This result can be used by designers to guarantee that characteristics with high probability are very unlikely. However, this bound exhibits a large gap between the fixed-key and the expected probabilities (see Table 1 in [BBL13]). It is then of little use to the cryptanalyst who needs to estimate the exact probability of some characteristic for a given key.

1.2 Our Contributions

In this paper, we evaluate the exact probabilities of the differential characteristics in some unkeyed constructions. In particular, we provide an in-depth study of the probabilities of the differential characteristics over three rounds of an unkeyed Feistel network. Most notably, when the inner function follows an SPN construction with an S-box having differential uniformity 4, the exact probability of a 3-round characteristic is either zero or a value which is greater than or equal to the usual estimate with independent S-box assumption, p_{ind}. A more thorough analysis is then provided when the inner function consists of a single n-bit S-box with differential uniformity 4. We show that, in this case, the exact probability of any 3-round characteristic with only active Sboxes is either zero, or exceeds p_{ind} by a factor of 2^{ℓ} where $\ell \geq \max(0, n-6)$.

The above analysis is then applied to the lightweight 64-bit block cipher RoadRunneR [BS15]. It adopts a Feistel construction and its inner function starts and ends with the S-box application without applying any subkey, therefore the above generic analysis can be applied. Although no security is claimed against related-key attacks, the designers mentioned related-key differential characteristics with 24 active S-boxes on the full (12) rounds of RoadRunneR-128, whose probability is expected to be $2^{-2\cdot 24} = 2^{-48}$. The designers also speculated that the number of active S-boxes could be reduced further with more careful analysis. In this paper, we first concretize the related-key characteristic with 24 active S-boxes and show that the exact probability is higher than the original expectation. The comparison of two probabilities is shown in Table 1. The attack is implemented up to 8 rounds and the improved factor is verified. We prove that the minimum number of active S-boxes is 24 by using a SAT solver, thus our characteristic is fairly tight.

Finding related-key differential characteristics is much harder in RoadRunneR-80 due to its key schedule. We propose an 8-round characteristic with $p_{\text{ind}} = 2^{-68}$ which are unlikely to be satisfied even with a full codebook, but the improvement with p_{exact} increases it to 2^{-62}.

We then extend the application of our observations to SPN-based structures with almost-MDS binary matrices. In particular, we analyze p_{exact} of the differential characteristic in an authenticated encryption scheme Minalpher [STA+14], which offers 128-bit security. The previous differential characteristic reaches 2^{-128} for 6 (out of 17.5) rounds. We show that for this characteristic a refined estimate of the exact probability is 2^{-96}. This significant increase enables us to extend the attack by two rounds. The comparison of the probabilities are given in Table 1.

1.3 Paper Outline

The paper is organized as follows. Section 2 provides theoretical analysis of p_{exact} for 3-round Feistel structure. Section 3 applies the observation to RoadRunneR with 128-bit key. Section 4 extends the application to SPN with almost-MDS matrices in Minalpher.

Table 1: Improved probability of characteristics for RoadRunneR-128 and Minalpher.

Rounds	1	2	3	4	5	6	7	8	9	10	11	12
RoadRunneR-128												
p_{ind}	-4	-8	-12	-16	-20	-24	-28	-32	-36	-40	-44	-48
p_{exact}	-4†	-8†	-12†	-15†	-19†	-22†	-26†	-29†	-33	-36	-40	-43
RoadRunneR-80												
p_{ind}	-8	-17	-26	-34	-42	-51	-60	-68				
p_{exact}	-8†	-17†	-25†	-32†	-39†	-47	-55	-62				
Minalpher												
p_{ind}	-16	-48	-64	-80	-112	-128						
p_{exact}	-16†	-40‡	-48	-64	-88	-96	-112	-128				

Numbers denote logarithm of the probabilities. Probabilities with † were experimentally verified. Probability with ‡ was experimentally verified only for the essential part, namely the probability of passing through S-boxes that are affected by our analysis was verified.

2 Probabilities of 3-Round Characteristics in some Keyless Feistel Networks

In this section, we evaluate the exact probability of a differential characteristic over three rounds of an unkeyed Feistel network whose inner function is seen as a single S-box application. We then want to determine the probability over all possible inputs (x_0, x_1) of the three-round characteristic depicted in Figure 1, where the difference at the output of the ith S-box is defined as $b_i = a_{i+1} \oplus a_{i-1}$. It is worth noticing that the differential probabilities for an unkeyed 3-round Feistel have been previously investigated in order to determine the smallest differential uniformity we can get for an S-box which follows this construction [LW14, CDL15]. However, these papers focus on the maximum possible probability for a 3-round differential characteristic, while we want to obtain a formula which captures any given characteristic.

Using that $x_3 = S(x_2) \oplus x_1$, we get that the probability of the three-round characteristic defined by $(a_0, \ldots, a_4)$ is equal to the following probability:

$$\begin{aligned} p_{\text{exact}} = &\Pr_{x_1, x_2 \in \mathbb{F}_2^n}[S(S(x_2) \oplus x_1 \oplus a_1 \oplus b_2) \oplus S(S(x_2) \oplus x_1) = b_3 \\ &\text{and } S(x_2 \oplus a_2) \oplus S(x_2) = b_2 \text{ and } S(x_1 \oplus a_1) \oplus S(x_1) = b_1]\,. \end{aligned}$$

We will show that this probability may differ from the usual estimate obtained when assuming that the inputs of the three S-boxes are independent, i.e. from

$$\begin{aligned} p_{\text{ind}} = &\Pr_{x_3 \in \mathbb{F}_2^n}[S(x_3 \oplus a_1 \oplus b_2) \oplus S(x_3) = b_3] \times \Pr_{x_2 \in \mathbb{F}_2^n}[S(x_2 \oplus a_2) \oplus S(x_2) = b_2] \\ &\times \Pr_{x_1 \in \mathbb{F}_2^n}[S(x_1 \oplus a_1) \oplus S(x_1) = b_1]\,. \end{aligned}$$

The difference between the two probabilities mainly comes from some dependencies due to the fact that the input of the S-box in the third round is the sum of two elements, x_1 and $S(x_2)$, where x_1 and x_2 respectively conform to the S-box differentials (a_1, b_1) and (a_2, b_2). Also, we show that the size of the S-box and, for a given size, the choice of the S-box may affect the factor between the exact probability and the usual estimate.

More precisely, we first show that, in many cases, including when S has an SPN structure based on an S-box with differential uniformity at most 4, the factor λ between these two probabilities is either zero or a power of 2 whose exponent corresponds to the dimension of a well-defined linear space. Most notably, if S corresponds to a single S-box with differential uniformity at most 4, then

$$p_{\text{exact}} = \lambda p_{\text{ind}}.$$

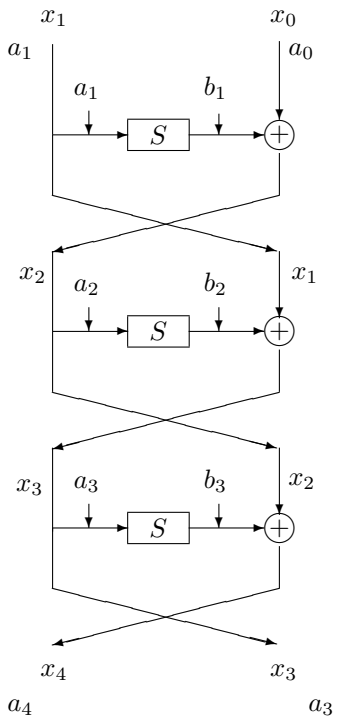

Figure 1: Differential characteristic of a three-round Feistel network where $b_i = a_{i+1} \oplus a_{i-1}$.

with $\lambda \in \{0, 2^\ell, \text{ with } \max(0, n-6) \leq \ell \leq n-2\}$, unless one of the three S-boxes in the differential path is inactive, which corresponds to $p_{\text{exact}} = p_{\text{ind}}$.

2.1 General result

The technique used in the proof is similar to the one used by Daemen and Rijmen for computing the fixed-key probabilities of the differentials over two rounds of the AES [DR07]. It mainly relies on the algebraic structure of the sets of inputs (resp. of outputs) of the S-box conforming to a given differential. These sets are defined as follows.

Definition 1. Let S be an n-bit to n-bit S-box. For any pair (a, b) of differences, we use the following notation:

$$\mathcal{X}_S(a,b) \triangleq \{x \in \mathbb{F}_2^n \colon S(x \oplus a) \oplus S(x) = b\},$$

and

$$\mathcal{Y}_S(a,b) \triangleq \{S(x) \in \mathbb{F}_2^n \colon S(x \oplus a) \oplus S(x) = b\}.$$

Remark 1. In the following, we will use some relationships between the sets $\mathcal{X}_S(a,b)$ and $\mathcal{Y}_S(a,b)$. Obviously,

$$\mathcal{Y}_S(a,b) = S\left(\mathcal{X}_S(a,b)\right) \ .$$

Moreover, if S is a permutation,

$$\mathcal{Y}_S(a,b) = \mathcal{X}_{S^{-1}}(b,a) \ .$$

Indeed, $y \in \mathcal{Y}_S(a,b)$ if and only if $x = S^{-1}(y)$ satisfies

$$S(x \oplus a) \oplus S(x) = b \ .$$

Then, we have

$$S(S^{-1}(y) \oplus a) = y \oplus b$$

which is equivalent to

$$S^{-1}(y) \oplus a = S^{-1}(y \oplus b)\,,$$

i.e., $y \in \mathcal{X}_{S^{-1}}(b, a)$.

Now, we focus on the following data transformation depicted in Figure 2:

$$z = S(x_2) \oplus x_1 \text{ such that } x_1 \in \mathcal{X}_S(a_1, b_1) \text{ and } x_2 \in \mathcal{X}_S(a_2, b_2).$$

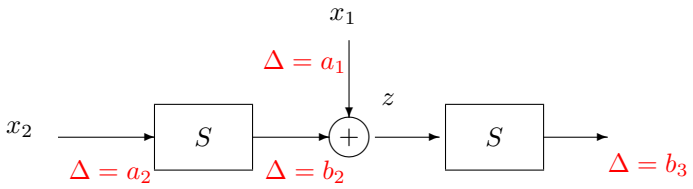

Figure 2: Target structure.

When the three sets $\mathcal{X}_S(a_1, b_1)$, $\mathcal{Y}_S(a_2, b_2)$ and $\mathcal{X}_S(a_1 \oplus b_2, b_3)$ are affine subspaces, we get the following result.

Theorem 1. *Let S be a permutation of $\mathbb{F}_2^n$, and let a_1, b_1, a_2, b_2, b_3 be five elements in $\mathbb{F}_2^n$. Assume that there exist $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{F}_2^n$ and three linear subspaces $V_1, V_2, V_3 \subseteq \mathbb{F}_2^n$ such that*

$$\mathcal{X}_S(a_1, b_1) = \alpha_1 + V_1,\ \mathcal{Y}_S(a_2, b_2) = \alpha_2 + V_2,\ \text{and } \mathcal{X}_S(a_1 \oplus b_2, b_3) = \alpha_3 + V_3\,.$$

Then, the multiset

$$\{(x_1, x_2) \in \mathcal{X}_S(a_1, b_1) \times \mathcal{X}_S(a_2, b_2) \colon S(S(x_2) \oplus x_1 \oplus a_1 \oplus b_2) \oplus S(S(x_2) \oplus x_1) = b_3\}$$

is either empty or has size 2^d with

$$d = \dim V_1 + \dim V_2 + \dim V_3 - \dim(V_1 + V_2 + V_3)$$

where $V_1 + V_2 + V_3$ denotes the linear space formed by all elements of the form $v_1 + v_2 + v_3$ with $v_i \in V_i$.

Proof. We first observe that we do not need to restrict ourselves to the situation where the input differences of all S-boxes are nonzero. Indeed, if the input difference of one S-box is zero (i.e. $a_1 = 0$ or $a_2 = 0$ or $a_1 = b_2$), either the corresponding output difference is nonzero, which implies that $p_{\text{exact}} = 0$ and the multiset we consider is empty, or the corresponding output difference is zero, and the associated set (i.e. $\mathcal{X}_S(a_1, b_1)$ or $\mathcal{Y}_S(a_2, b_2)$ or $\mathcal{X}_S(a_1 \oplus b_2, b_3)$) satisfies the hypothesis since it equals the whole space $\mathbb{F}_2^n$.

Let us now define the following set (without multiplicity)

$$\mathcal{Z} = \{(S(x_2) \oplus x_1) \colon (x_1, x_2) \in \mathcal{X}_S(a_1, b_1) \times \mathcal{X}_S(a_2, b_2)\}\,.$$

Then, $\mathcal{Z} = (\alpha_1 \oplus \alpha_2) + (V_1 + V_2)$, and each element in $\mathcal{Z}$ corresponds to 2^r pairs (x_1, x_2) in $\mathcal{X}_S(a_1, b_1) \times \mathcal{X}_S(a_2, b_2)$ with $r = \dim V_1 + \dim V_2 - \dim(V_1 + V_2)$. We want to determine the size of the set

$$\mathcal{S} = \{z \in \mathcal{Z} \colon S(z \oplus a_1 \oplus b_2) \oplus S(z) = b_3\}\,.$$

Clearly, this set corresponds to the intersection between $\mathcal{Z}$ and $\mathcal{X}_S(a_1 \oplus b_2, b_3)$, which are both affine subspaces of $\mathbb{F}_2^n$. Since the intersection between two affine subspaces is either empty or a coset of the intersection between the corresponding linear subspaces, we deduce that, if $\mathcal{S} \neq \emptyset$, then there exists some s such that

$$\mathcal{S} = s + ((V_1 + V_2) \cap V_3)\,.$$

Recall that, for any two linear subspaces U and V,

$$\dim(U+V) = \dim U + \dim V - \dim(U \cap V)\,. \tag{1}$$

It follows from (1) that, if $\mathcal{S} \neq \emptyset$, we have

$$\dim \mathcal{S} = \dim((V_1 + V_2) \cap V_3) = \dim(V_1 + V_2) + \dim V_3 - \dim(V_1 + V_2 + V_3)\,.$$

Since each element in $\mathcal{Z}$ and then in $\mathcal{S}$ corresponds to 2^r pairs (x_1, x_2) in $\mathcal{X}_S(a_1, b_1) \times \mathcal{X}_S(a_2, b_2)$, we deduce that the multiset

$$\{(x_1, x_2) \in \mathcal{X}_S(a_1, b_1) \times \mathcal{X}_S(a_2, b_2) \colon S(S(x_2) \oplus x_1 \oplus a_1 \oplus b_2) \oplus S(S(x_2) \oplus x_1) = b_3\}$$

is either empty or has size 2^d with

$$\begin{aligned} d &= r + \dim \mathcal{S} \\ &= \dim V_1 + \dim V_2 - \dim(V_1 + V_2) + \dim(V_1 + V_2) + \dim V_3 - \dim(V_1 + V_2 + V_3) \\ &= \dim V_1 + \dim V_2 + \dim V_3 - \dim(V_1 + V_2 + V_3)\,. \end{aligned}$$

□ □

Remark 2. For the sake of simplicity, the previous theorem considers a 3-round Feistel network with the same keyless S-box. However, since the result only relies on the structure of the three sets $\mathcal{X}_S(a_1, b_1)$, $\mathcal{Y}_S(a_2, b_2)$ and $\mathcal{X}_S(a_1 \oplus b_2, b_3)$, it clearly appears that Theorem 1 also holds for a Feistel network with three different S-boxes, S_1, S_2 and S_3, as soon as $\mathcal{X}_{S_1}(a_1, b_1)$, $\mathcal{Y}_{S_2}(a_2, b_2)$ and $\mathcal{X}_{S_3}(a_1 \oplus b_2, b_3)$ are affine subspaces.

As a direct consequence of Theorem 1, we get the following corollary.

Corollary 1. *Let S be a permutation of $\mathbb{F}_2^n$, and let a_1, b_1, a_2, b_2, b_3 be five elements in $\mathbb{F}_2^n$. Assume that there exist $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{F}_2^n$ and three linear subspaces $V_1, V_2, V_3 \subseteq \mathbb{F}_2^n$ such that*

$$\mathcal{X}_S(a_1, b_1) = \alpha_1 + V_1,\ \mathcal{Y}_S(a_2, b_2) = \alpha_2 + V_2,\ \textit{and}\ \mathcal{X}_S(a_1 \oplus b_2, b_3) = \alpha_3 + V_3\,.$$

Let

$$\begin{aligned} p_{\text{exact}} \;=\; & \Pr_{x_1, x_2 \in \mathbb{F}_2^n}[S(S(x_2) \oplus x_1 \oplus a_1 \oplus b_2) \oplus S(S(x_2) \oplus x_1) = b_3 \textit{ and} \\ & S(x_2 \oplus a_2) \oplus S(x_2) = b_2 \textit{ and } S(x_1 \oplus a_1) \oplus S(x_1) = b_1] \end{aligned}$$

and

$$\begin{aligned} p_{\text{ind}} \;=\; & \Pr_{x_1 \in \mathbb{F}_2^n}[S(x_1 \oplus a_1) \oplus S(x_1) = b_1] \times \Pr_{x_2 \in \mathbb{F}_2^n}[S(x_2 \oplus a_2) \oplus S(x_2) = b_2] \\ & \times \Pr_{z \in \mathbb{F}_2^n}[S(z \oplus a_1 \oplus b_2) \oplus S(z) = b_3] \end{aligned}$$

Then, either $p_{\text{exact}} = 0$ or

$$p_{\text{exact}} = 2^{\ell} p_{\text{ind}} \textit{ with } \ell = n - \dim(V_1 + V_2 + V_3)\,.$$

Most notably, $0 \leq \ell \leq n - 2$.

Proof. Let us focus on the case where $p_{\text{exact}} \neq 0$. We deduce from Theorem 1 that

$$p_{\text{exact}} = 2^{\dim V_1 + \dim V_2 + \dim V_3 - \dim(V_1+V_2+V_3) - 2n}.$$

Since $p_{\text{ind}} = 2^{\dim V_1 + \dim V_2 + \dim V_3 - 3n}$, we obtain that

$$\lambda = \frac{p_{\text{exact}}}{p_{\text{ind}}} = 2^{\ell}$$

with

$$\begin{aligned}\ell &= \dim V_1 + \dim V_2 + \dim V_3 - \dim(V_1 + V_2 + V_3) - 2n - (\dim V_1 + \dim V_2 + \dim V_3 - 3n) \\ &= n - \dim(V_1 + V_2 + V_3)\,.\end{aligned}$$

Since $V_1 + V_2 + V_3$ is a subspace of $\mathbb{F}_2^n$, its dimension does not exceed n. On the other hand, when $p_{\text{ind}} \neq 0$, V_1 (resp. V_2) contains at least two elements, 0 and a_1 (resp. 0 and b_2). It follows that, if $a_1 \neq b_2$, then $V_1 + V_2$ contains the linear space spanned by a_1 and b_2, i.e. $\langle a_1, b_2 \rangle$, which has dimension 2, implying that $\dim(V_1 + V_2 + V_3) \geq 2$. This lower bound also holds when $a_1 = b_2$ since this corresponds to $V_3 = \mathbb{F}_2^n$, leading to $\dim(V_1 + V_2 + V_3) = n$. Therefore, we have proved that

$$2 \leq \dim(V_1 + V_2 + V_3) \leq n$$

implying

$$0 \leq \ell \leq n - 2\,.$$

□ □

The hypothesis required for applying by this result, i.e., the fact that the three sets $\mathcal{X}_S(a_1, b_1)$, $\mathcal{Y}_S(a_2, b_2)$ and $\mathcal{X}_S(a_1 \oplus b_2, b_3)$ are affine subspaces, is satisfied in many practical cases. Indeed, when an S-box σ has differential uniformity at most 4, i.e., when 4 is the maximal value in the difference distribution table of σ, all sets $\mathcal{X}_\sigma(a, b)$ and $\mathcal{Y}_\sigma(a, b)$ are affine subspaces (see e.g., Lemma 2 in [DR07]). Therefore, the hypothesis is satisfied when S has an SPN structure based on a smaller differentially 4-uniform S-box σ: in this case, $\mathcal{X}_S(a, b)$ (resp. $\mathcal{Y}_S(a, b)$) corresponds to the Cartesian product of sets of the form $\mathcal{X}_\sigma(a, b)$ (resp. $\mathcal{Y}_\sigma(a, b)$).

An interesting observation deduced from the previous corollary is that, in all the previously mentioned situations, if the exact probability of a 3-round differential characteristic is non-zero, then it is greater than or equal to the usual estimate p_{ind}.

2.2 When S is differentially 4-uniform

There is a specific case where the factor λ between the two probabilities can be easily lower-bounded: when S itself is a function with differential uniformity at most 4.

Theorem 2. *Let S be a permutation of $\mathbb{F}_2^n$ with differential uniformity at most 4. Let a_1, b_1, a_2, b_2, b_3 be five nonzero elements in $\mathbb{F}_2^n$. Let*

$$\begin{aligned}p_{\text{exact}} \;=\; &\Pr_{x_1, x_2 \in \mathbb{F}_2^n}\big[S(S(x_2) \oplus x_1 \oplus a_1 \oplus b_2) \oplus S(S(x_2) \oplus x_1) = b_3 \text{ and} \\ &S(x_2 \oplus a_2) \oplus S(x_2) = b_2 \text{ and } S(x_1 \oplus a_1) \oplus S(x_1) = b_1\big]\end{aligned}$$

and

$$\begin{aligned}p_{\text{ind}} \;=\; &\Pr_{x_1 \in \mathbb{F}_2^n}[S(x_1 \oplus a_1) \oplus S(x_1) = b_1] \times \Pr_{x_2 \in \mathbb{F}_2^n}[S(x_2 \oplus a_2) \oplus S(x_2) = b_2] \\ &\times \Pr_{z \in \mathbb{F}_2^n}[S(z \oplus a_1 \oplus b_2) \oplus S(z) = b_3]\end{aligned}$$

Then,

- *if* $a_1 = 0$ *or* $a_2 = 0$ *or* $a_1 = b_2$*, then*

$$p_{\text{exact}} = p_{\text{ind}} \; ;$$

- *if the three S-boxes are active, i.e.* $a_1 \neq 0$ *and* $a_2 \neq 0$ *or* $a_1 \neq b_2$*, then either* $p_{\text{exact}} = 0$ *or*

$$p_{\text{exact}} = 2^{\ell} p_{\text{ind}} \text{ with } \max(0, n-6) \leq \ell \leq n-2 \; .$$

Moreover, if all three differentials (a_1, b_1)*,* (a_2, b_2)*, and* $(a_1 \oplus b_2, b_3)$ *have probability* 2^{1-n}*, then* $\lambda \in \{0, 2^{n-2}\}$*.*

Proof. We know from Corollary 1 that $p_{\text{exact}} = 0$ or $p_{\text{exact}} = 2^{\ell} p_{\text{ind}}$ with $\ell = n - \dim(V_1 + V_2 + V_3)$. Since $V_1 + V_2 + V_3$ is a subspace of $\mathbb{F}_2^n$, its dimension does not exceed n and is also smaller than the sum of the dimensions of the three subspaces. Since the S-box has differential uniformity at most 4, all V_i have dimension at most 2 unless the corresponding S-box is inactive, which is equivalent to $V_i = \mathbb{F}_2^n$.

- Let us first assume that the input difference of one of the S-boxes is zero. If the corresponding output difference is nonzero, the transition is not valid. In this case, we have $p_{\text{exact}} = p_{\text{ind}} = 0$. If the corresponding output is zero, i.e. if the S-box is inactive, the associated linear space V_i equals the whole space. It follows that $\ell = n - \dim(V_1 + V_2 + V_3) = 0$, leading to $p_{\text{exact}} = p_{\text{ind}}$.

- Let us now assume that all the three S-boxes are active. Then, $\dim(V_1 + V_2 + V_3)$ is smaller than 6. We derive that

$$\max(0, n-6) \leq \ell \leq n-2 \, .$$

Moreover, when all three subspaces V_1, V_2, and V_3 have dimension 1, then

$$V_1 + V_2 + V_3 = \langle a_1, b_2 \rangle \, .$$

It follows that, in this case,

$$\lambda \in \{0, 2^{n-2}\} \, .$$

In other words,

$$p_{\text{exact}} \in \{0, 2^{-2n+1}\} \, .$$

□ □

Most notably, when $n > 6$, if the differential path contains three active S-boxes, then its exact probability can never be equal to the product of the probabilities of the three constituent transitions.

Example 1. Theorem 2 can be verified for instance when S is the AES S-box, which operates on $\mathbb{F}_2^8$. Most differentials for the AES S-box have probability 2^{-7}. For such differential paths, we can check that $p_{\text{exact}} \in \{0, 2^{-15}\}$. For instance, for $(a_1, b_1) = (\texttt{0x01}, \texttt{0xca})$, $(a_2, b_2) = (\texttt{0xe5}, \texttt{0x18})$, and $b_3 = \texttt{0xb3}$, there are exactly two pairs $(x_1, x_2) \in \mathcal{X}_S(a_1, b_1) \times \mathcal{X}_S(a_2, b_2)$ such that $(S(x_2) \oplus x_1)$ satisfies the differential $(a_1 \oplus b_2, b_3)$. Then, the probability of the whole differential path is 2^{-15} while all three differentials have probability 2^{-7}, i.e., $\lambda = 2^{-15+21} = 2^6$. This factor varies when some of the involved differentials have probability 2^{-6}. For $(a_1, b_1) = (\texttt{0x01}, \texttt{0x1f})$, $(a_2, b_2) = (\texttt{0x33}, \texttt{0x0f})$ and $b_3 = \texttt{0xb8}$, the probability of the whole differential path is again 2^{-15}, while the second differential has probability 2^{-7} and the other two have probability 2^{-6}. We then have $\lambda = 2^{-15+19} = 2^4$.

If all S-boxes are active, the highest possible value for p_{exact} is $2^{n-2} \times \left(2^{-(n-2)}\right)^3 = 2^{-2n+4}$. It is worth noticing that this also corresponds to the highest possible value for p_{exact} when only two S-boxes are active, i.e. $p_{\text{exact}} = \left(2^{-(n-2)}\right)^2 = 2^{-2n+4}$. We now give a simple necessary condition on a_1 and b_2 for obtaining differential paths with three active S-boxes achieving this maximal probability.

Proposition 1. *Let S be a permutation of $\mathbb{F}_2^n$ with differential uniformity exactly 4. If there exist nonzero $a_1, b_1, a_1, b_2, b_3 \in \mathbb{F}_2^n$ such that $p_{exact} = 2^{-2n+4}$, then there exist x and y in $\mathbb{F}_2^n$ such that the second-order derivatives of S and S^{-1} satisfy*

$$D_{a_1} D_{b_2} S(x) = 0 \text{ and } D_{a_1} D_{b_2} S^{-1}(y) = 0\,, \tag{2}$$

where $D_u D_v S(x) = S(x) \oplus S(x \oplus u) \oplus S(x \oplus v) \oplus S(x \oplus u \oplus v)$.

It is worth noticing that, if S is an involution, then there always exists a pair (a_1, b_2) such that Condition (2) holds for some x and y in $\mathbb{F}_2^n$.

Proof. By hypothesis, all the three S-boxes are active. Then, $p_{\text{ind}} \leq 2^{-3n+6}$ and we know from Theorem 2 that $\lambda \leq 2^{n-2}$. It follows that $p_{\text{exact}} = 2^{-2n+4}$ if and only if $\lambda = 2^{n-2}$ (i.e., if $\dim(V_1 + V_2 + V_3) = 2$) and all the three involved differentials have probability $2^{-(n-2)}$. Since the differential (a_1, b_1) has probability $2^{-(n-2)}$, there exists $x, v_1 \in \mathbb{F}_2^n$ with $v_1 \neq \{0, a_1\}$ such that $\mathcal{X}_S(a_1, b_1) = x + \langle a_1, v_1 \rangle$. This implies that

$$S(x) \oplus S(x \oplus a_1) = b_1 = S(x \oplus v_1) \oplus S(x \oplus v_1 \oplus a_1)$$

leading to

$$D_{a_1} D_{v_1} S(x) = 0\,.$$

Similarly, a_2 is such that $\mathcal{Y}_S(a_2, b_2) = y + \langle b_2, v_2 \rangle$ for some $y, v_2 \in \mathbb{F}_2^n$ with $v_2 \notin \{0, b_2\}$. We now use the fact that, for any permutation S, $\mathcal{Y}_S(a, b) = \mathcal{X}_{S^{-1}}(b, a)$ (see Remark 1). From the same arguments as for v_1, we deduce that

$$D_{b_2} D_{v_2} S^{-1}(y) = 0\,.$$

But, since $\lambda = 2^{n-2}$, we know that

$$\dim(V_1 + V_2) = \dim\langle a_1, b_2, v_1, v_2 \rangle = 2\,.$$

It follows that $v_1 \in \{b_2, b_2 \oplus a_1\}$ and $v_2 \in \{a_1, a_1 \oplus b_2\}$. This implies that $D_{a_1} D_{b_2} S(x) = 0$ and $D_{a_1} D_{b_2} S^{-1}(y) = 0$.

It is well-known that there is no pair of nonzero distinct elements (a, b) such that $D_a D_b S$ takes the value 0 if and only if S is APN (i.e., its differential uniformity equals 2) [Nyb94]. In our case, S is not APN, implying that such a pair (a, b) exists. When S is an involution, is also satisfies $D_a D_b S^{-1}(y) = 0$ for some y. □ □

Example 2 (RoadRunneR S-box)**.** It is easy to check that, for the RoadRunneR [BS15] S-box, there is no pair of nonzero distinct elements (a_1, b_2) such that both $D_{a_1} D_{b_2} S$ and $D_{a_1} D_{b_2} S^{-1}$ vanish at some points. We then deduce that any differential path with three active S-boxes satisfies $p_{\text{exact}} \leq 2^{-5}$. By examining all second-order derivatives of this S-box which take the value 0, we have searched for all $(a_1, b_1, a_2, b_2, b_3)$ such that all three differentials have probability 2^{-2} and lead to a differential path with overall probability 2^{-5}. We have found 136 such configurations. One example is

$$a_1 = \texttt{0x1}, b_1 = \texttt{0x1}, a_2 = \texttt{0x8}, b_2 = \texttt{0x4}, b_3 = \texttt{0x8}\,.$$

Among these patterns, the only one which satisfies $a_2 = a_1 \oplus b_2$ and such that also the differentials (a_1, b_2) and (a_1, b_3) have probability 2^{-2} is the one we will use in the next section:

$$a_1 = \texttt{0xd}, a_2 = \texttt{0xc} \text{ and } b_1 = b_2 = b_3 = \texttt{0x1}\,,$$

and the configuration obtained by inverting the roles of a_1 and a_2.

Example 3 (Klein S-box)**.** The Klein [GNL11] S-box is an involution over $\mathbb{F}_2^4$. Then, there exist some pairs of nonzero distinct elements (a_1, b_2) such that both $D_{a_1}D_{b_2}S$ and $D_{a_1}D_{b_2}S^{-1}$ vanish at some points. For instance, $a_1 = \texttt{0x1}$ and $b_2 = \texttt{0x2}$ satisfy this property. For this S-box, the differential path defined by

$$a_1 = \texttt{0x1}, b_1 = \texttt{0x3}, a_2 = \texttt{0xd}, b_2 = \texttt{0x2}, \text{ and } b_3 = \texttt{0xe}$$

has overall probability 2^{-4}. In other words, any pair of elements (x_1, x_2) satisfying the first two differentials also leads to some $(S(x_2) \oplus x_1)$ which satisfies the third one.

All previous results hold in the keyless setting, but are still valid when the three S-boxes are distinct permutations with differential uniformity 4. This enables us to cover the fixed-key scenario since using S with a fixed round-key k is equivalent to using $S' : x \mapsto S_k(x)$. For instance, in the fixed-key scenario, Theorem 2 states that a differential path with three active S-boxes satisfies $p_{\text{exact}} = \lambda p_{\text{ind}}$ with $\lambda \in \{0, 2^\ell, \text{ with } \max(0, n-6) \leq \ell \leq n-2\}$. However, for a given differential path, the value of λ may vary with the key. For instance, the same differential path may have probability zero for some round-keys, and probability $p_{\text{exact}} > 0$ for the other ones.

3 Application to RoadRunneR

3.1 Description of RoadRunneR

RoadRunneR is a lightweight block cipher recently proposed by Baysal and Sahin [BS15]. It has a Feistel network structure with a 64-bit block size and it supports both 80 and 128-bit keys. In the 80-bit version, the number of rounds is 10, whereas in the 128-bit version the number of rounds is 12. Whitening keys (WK_0 and WK_1) are applied to the left half of the block in the first and last round. The general structure of RoadRunneR is depicted in Figure 3.

Round Function. RoadRunneR's round function, named F, takes as input a 32-bit block L_i, a 96-bit subkey K_i, and a 32-bit constant C_i. The constant C_i for round i is the 32-bit value $N_r - i$, where N_r is the total number of rounds of the cipher as defined above.

The round function in RoadRunneR consists of three subsequent applications of SLK, which is composed of a substitution layer followed by a linear layer and a key addition layer. After three SLK layers a single substitution layer (S) is performed. In between the second and third SLK layer the constant C_i is added (cf. Figure 3).

Key Schedule. The key expansion of the 128-bit RoadRunneR version chops the key up in four 32-bit words. The round keys are permutations of these words. Similarly, in the 80-bit version the key is split into five 16-bit words, and the key schedule is a permutation of six words. Table 2 lists the exact permutations for the round and whitening keys.

Substitution Layer. The substitution layer S consists of a parallel composition of the 4×4-bit S-box of Table 3[1] to every 4-bit nibble of the block.

Linear Layer. The linear layer L applies the function $L' : \mathbb{F}_2^8 \mapsto \mathbb{F}_2^8$ to each individual byte of the block

$$L'(x) = x \oplus (x \lll 1) \oplus (x \lll 2).$$

This construction is known to be invertible in general for distinct rotation offsets [Riv11], and the designers of RoadRunneR argue that this particular set of rotation offsets has good diffusion properties.

[1] This is the "optimal" S-box 13 in [UCI+11, Table 4].

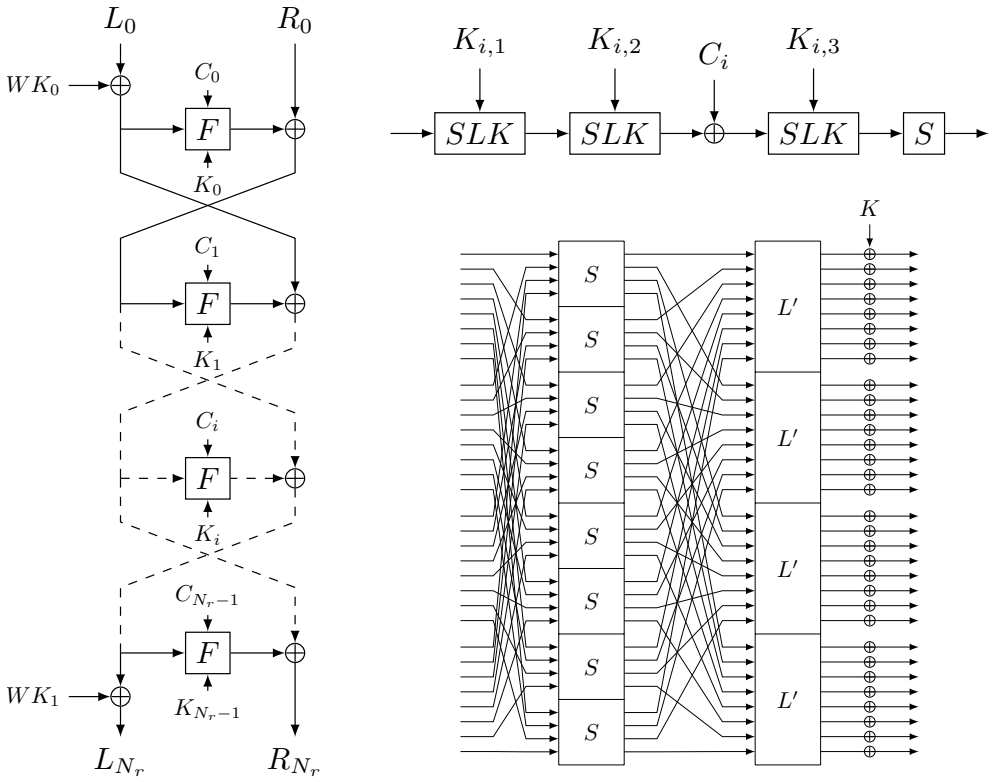

Figure 3: Overview of the RoadRunneR block cipher. Left: Feistel network with whitening keys xored in the first and last round. Top right: The round function F, taking in as input a 32-bit word, a 32-bit constant and a 96-bit round key. Bottom right: The core SLK function, which consists of an S-box layer followed by a linear diffusion layer and finally a key addition.

3.2 Security Analysis by the Designers

The designers claim no security in the related-key setting, due to the fact that the key schedule uses the master key without any change in between rounds. The designers in fact mention in the paper that each F can be passed with only two active S-boxes in a related key attack, with total of 24 active S-boxes, and that this total number may be further reduced in a more detailed analysis. We stress that no information about concrete characteristics, such as plaintext and subkey difference is provided.

In the single-key setting, the designers show that the minimum number of active S-boxes in an active F is 10 along with concrete propagation patterns. The authors experimentally checked that the probability of characteristics and differentials is correct. In their experiments they report that, the differential probability does not significantly increase from the theoretically calculated characteristic probability. Based on this experiment, the authors assume that each active S-box multiplies the probability with 2^{-2} and an active F has approximately a probability of 2^{-20}.

3.3 Applications of our Observations

By comparing Figure 1 and Figure 3, it is easy to see that the analysis in Section 2 can directly be applied to RoadRunneR when the number of rounds is more than two. We

Table 2: RoadRunneR's key schedule.

(a) 128-bit key.

Round Number	Key schedule
WK_0	A
WK_1	B
0 (mod 4)	$B\|\|C\|\|D$
1 (mod 4)	$A\|\|B\|\|C$
2 (mod 4)	$D\|\|A\|\|B$
3 (mod 4)	$C\|\|D\|\|A$

(b) 80-bit key.

Round Number	Key schedule
WK_0	$A\|\|B$
WK_1	$C\|\|D$
0 (mod 5)	$C\|\|D\|\|E\|\|A\|\|B\|\|C$
1 (mod 5)	$D\|\|E\|\|A\|\|B\|\|C\|\|D$
2 (mod 5)	$E\|\|A\|\|B\|\|C\|\|D\|\|E$
3 (mod 5)	$A\|\|B\|\|C\|\|D\|\|E\|\|A$
4 (mod 5)	$B\|\|C\|\|D\|\|E\|\|A\|\|B$

Table 3: The RoadRunneR S-box.

x	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
$S(x)$	0	8	6	D	5	F	7	C	4	E	2	3	9	1	B	A

emphasize that the observations can be applied both in the single-key and related-key settings. We also notice that the observation does not contradict the experiments by the designers that verified the probability of differentials within one round. What we are showing is that even before calculating the effect of collecting multiple differences, the actual probability of characteristics p_{exact} is higher than theoretically calculated one, p_{ind}, under the independent S-box assumption when the number of rounds is more than two.

In the following sections, we demonstrate the power of our observations with applications to concrete attacks.

3.4 Attack on RoadRunneR-128

First, we concretize the characteristic having only two active S-boxes per round mentioned by the designers. Suppose that a 128-bit master key K is denoted by four 32-bit values and the difference of those values are denoted by $\Delta_0, \Delta_1, \Delta_2$ and Δ_3. By following the key schedule described in Table 2, the difference of the initial whitening key is $\Delta WK_0 = \Delta_0$. Then, subkey differences are $(\Delta_1, \Delta_2, \Delta_3)$ for the first round, $(\Delta_0, \Delta_1, \Delta_2)$ for the second round, $(\Delta_3, \Delta_0, \Delta_1)$ for the third round, and so on. Four rounds with those subkey differences are illustrated in Figure 4.

We then choose $\Delta_0, \Delta_1, \Delta_2$ and Δ_3. There are four S-layers in each round. Our strategy consists in canceling the difference from Δ_1 with Δ_2 after the S-layer, which makes the next S-layer inactive. Then canceling the difference from Δ_3 with Δ_0 after the S-layer, which makes the next S-layer inactive. By iterating this, non-active S-layers and active S-layers appear alternately, and we only have 2 active S-boxes per round.

As a result of our analysis, we construct a 4-round iterative characteristic by satisfying the following four conditions.

$$\Pr_{x\in\mathbb{F}_2^4}[S(x) \oplus S(x \oplus \delta_1) = \gamma_2] = 2^{-2}, \tag{3}$$

$$\Pr_{x\in\mathbb{F}_2^4}[S(x) \oplus S(x \oplus \delta_3) = \gamma_0] = 2^{-2}, \tag{4}$$

$$\Pr_{x\in\mathbb{F}_2^4}[S(x) \oplus S(x \oplus \delta_1) = \delta_1 \oplus \delta_3] = 2^{-2}, \tag{5}$$

$$\Pr_{x\in\mathbb{F}_2^4}[S(x) \oplus S(x \oplus \delta_3) = \delta_1 \oplus \delta_3] = 2^{-2}, \tag{6}$$

where δ_1 is a group of 4 bits in the 32-bit differences Δ_1 and the 4 bits gather into a single active S-box after the bit-permutation around the S-layer. δ_3 can similarly be defined.

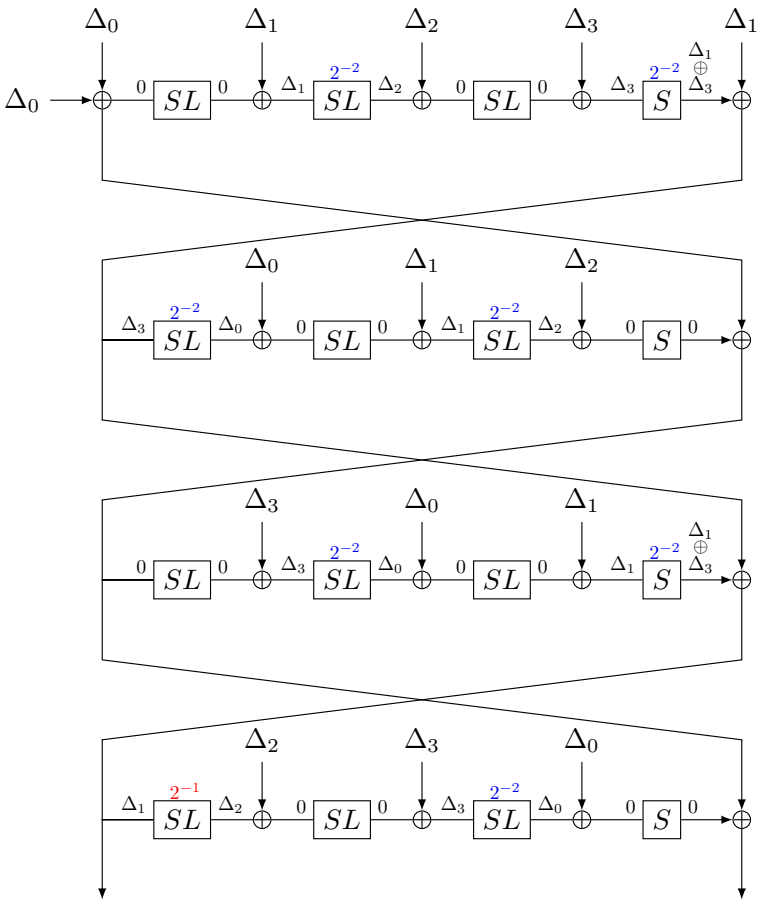

Figure 4: Four-round iterative differential characteristic against RoadRunneR-128.

The difference γ_0 (resp. γ_2) corresponds to the corresponding nibble of $L^{-1}(\Delta_0)$ (resp. of $L^{-1}(\Delta_2)$) where L denotes the whole linear layer. For example, when the active S-box position is fixed to the top in Figure 3, $\delta_1 = \texttt{0xf}$ corresponds to $\Delta_1 = \texttt{0x01010101}$.

We note that by setting $\Delta_0 = \Delta_2 = L(\Delta_1 \oplus \Delta_3)$, the first two conditions can always be satisfied when the last two conditions are satisfied. The characteristic is iterative after 4 rounds including subkey differences, and can be extended to 12 rounds easily.

By analyzing the differential distribution table (DDT) of the S-box, we chose $\delta_1 = \texttt{0xc}$ and $\delta_3 = \texttt{0xd}$ (or $\Delta_1 = \texttt{0x01010000}$ and $\Delta_3 = \texttt{0x01010001}$). Then, $\delta_1 \oplus \delta_3 = \texttt{0x1}$ ($\Delta_0 = \Delta_2 = L(\texttt{0x00000001})$). This configuration satisfies the above listed conditions.

Evaluation of p_{ind} and p_{exact}. From Eqs. (3) to (6), p_{ind} can be calculated from the transition probability for each S-box, 2^{-2}, and the number of active S-boxes, leading to $2^{-2 \cdot 24} = 2^{-48}$.

Recall that for any pair (a, b) of differences, we use the following notation: $\mathcal{X}_S(a, b) = \{x \in \mathbb{F}_2^4 : S(x) \oplus S(x \oplus a) = b\}$ and $\mathcal{Y}_S(a, b) = \{S(x) \in \mathbb{F}_2^4 : S(x) \oplus S(x \oplus a) = b\}$. By applying the analysis in Section 2, p_{exact} of the first S-layer in round 4 in Figure 4 is

$$\Pr_{x \in \mathcal{X}_S(\texttt{0xd},\texttt{0x1}), y \in \mathcal{Y}_S(\texttt{0xc},\texttt{0x1})} [x \oplus y \in \mathcal{X}_S(\texttt{0xc}, \texttt{0x1})]. \tag{7}$$

By analyzing the DDT, we obtain $\mathcal{X}_S(\texttt{0xd}, \texttt{0x1}) = \{\texttt{0x0}, \texttt{0x1}, \texttt{0xc}, \texttt{0xd}\}$, $\mathcal{Y}_S(\texttt{0xc}, \texttt{0x1}) = \{\texttt{0x4}, \texttt{0x5}, \texttt{0xe}, \texttt{0xf}\}$, and $\mathcal{X}_S(\texttt{0xc}, \texttt{0x1}) = \{\texttt{0x4}, \texttt{0x5}, \texttt{0x8}, \texttt{0x9}\}$, which leads to $p_{\text{exact}} = 2^{-1}$.

Similarly, p_{exact} of the first S-layer in rounds 6, 8, 10, and 12 are 2^{-1}, which leads to 2^{-43}.

It is important to notice that this probability is evaluated in the keyless scenario studied in the previous section because it is not affected by the values of the round-keys. Indeed, the round-key is inserted *after* applying the S-box and then does not affect $\mathcal{X}_S(\texttt{0xc}, \texttt{0x1})$ and $\mathcal{X}_S(\texttt{0xd}, \texttt{0x1})$. Moreover, the S-box involved in $\mathcal{Y}_S(\texttt{0xc}, \texttt{0x1})$ corresponds to the last S-box-layer in the third round and is independent from the key. It follows that, in this situation, p_{exact} takes the same value for any fixed-key.

Experiments. First of all, we experimentally proved that 24 active S-boxes in 12 rounds is minimal by using the SAT-solver based tool [MP13]. Differently from the expectation by the designers, the number of active S-boxes will not be further reduced.

We then implemented the attack up to 8 rounds. We refer back to Table 1 for the results, which clearly indicates the gap between p_{ind} and p_{exact} in rounds 4, 6 and 8.

3.5 Attack on RoadRunneR-80

In this part, we present an 8-round attack against RoadRunneR-80. Differently from RoadRunneR-128, the key is divided into 16-bit values (A, B, C, D, E) and each of them can be both the top half or the bottom half of 32-bit subkeys. Hence, constructing systematic subkeys is harder than in RoadRunneR-128.

By applying the bit-permutation around S, a group of 4 bits for a single S-box will move to symmetric positions in the 32-bit state. To exploit this fact, we set $\Delta A = \Delta B = \Delta C = \Delta D = \Delta E$ to make all 32-bit subkey differences identical and symmetric.

We set subkey difference to the xor of two differences ΔX and ΔZ. ΔX takes a role of input difference to the subsequent S-layer, and ΔZ cancels the difference from the previous S-layer. Namely, in every S-layer, cancellation and injection of differences are performed. The characteristic is illustrated in Figure 5, which is iterative after four rounds.

We then choose Δ_X and Δ_Z, where $\Delta_Z \triangleq L(\Delta_Y)$. We define δ_X, δ_Y similarly to the previous section, namely 4-bit difference in the 32-bit variable corresponding to an active S-box. Because subkey difference is symmetric, Δ_X and Δ_Y must be symmetric, which further limits δ_X, δ_Y to be symmetric (and non-zero). Therefore, $\delta_X, \delta_Y \in \{\texttt{5}, \texttt{a}, \texttt{f}\}$. According to the characteristic in Figure 5, we have the following two conditions;

$$\Pr_{x \in \mathbb{F}_2^4}[S(x) \oplus S(x \oplus \delta_X) = \delta_Y] > 0, \tag{8}$$

$$\Pr_{x \in \mathbb{F}_2^4}[S(x) \oplus S(x \oplus \delta_X \oplus \delta_Y) = \delta_Y] > 0, \tag{9}$$

$$\delta_X \neq \delta_Y. \tag{10}$$

From DDT, there is only one choice, $\delta_X = \texttt{5}$ ($\Delta_X = \texttt{0x00010001}$) and $\delta_Y = \texttt{a}$, which satisfies Conditions (8) and (9) with probability 2^{-2} and 2^{-3}, respectively.

Evaluation of p_{ind} and p_{exact}. We first evaluate p_{ind}. In every two rounds, there are seven active S-boxes with probability of 2^{-2} and there is one active S-box with probability of 2^{-3}. Thus p_{ind} is 2^{-17} in every 2 rounds and 2^{-68} for 8 rounds, which are unlikely to be satisfied with 2^{64} plaintexts of the full codebook.

The mechanism of occurring the advantage of p_{exact} is the same as in the attack against RoadRunneR-128, but we now have an active S-box at the beginning of the inner function in every round. Therefore, from the third round, p_{exact} is higher than p_{ind} by a factor of 2, which improves the probability of 8-rounds to $2^{-8-9-8-7-7-8-8-7} = 2^{-62}$.

In more details, p_{exact} of the first S-layer in rounds with $p_{\text{ind}} = 2^{-8}$ and $p_{\text{ind}} = 2^{-9}$ are

$$\Pr_{x \in \mathcal{X}_S(\texttt{0xf},\texttt{0xa}), y \in \mathcal{Y}_S(\texttt{0x5},\texttt{0xa})}[x \oplus y \in \mathcal{X}_S(\texttt{0x5}, \texttt{0xa})], \tag{11}$$

$$\Pr_{x \in \mathcal{X}_S(\texttt{0x5},\texttt{0xa}), y \in \mathcal{Y}_S(\texttt{0x5},\texttt{0xa})}[x \oplus y \in \mathcal{X}_S(\texttt{0xf}, \texttt{0xa})]. \tag{12}$$

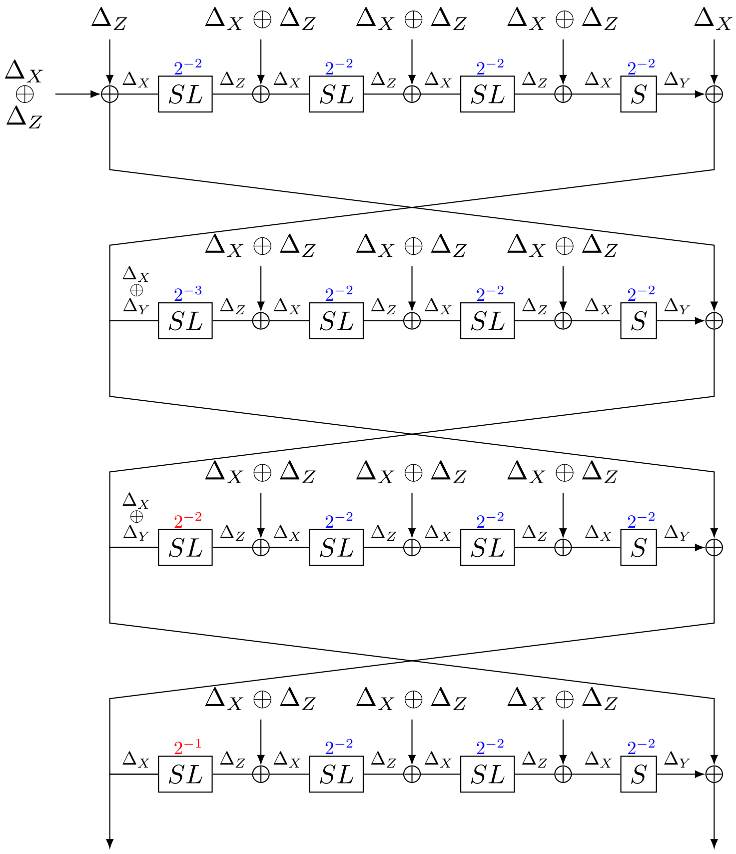

Figure 5: 4-round iterative characteristic for RoadRunneR-80. $\Delta_X = \texttt{0x5}$, $\Delta_Y = \texttt{0xA}$, $\Delta_Z = L(\Delta_Y)$. The transition probabilities in red are those which differ from the estimate with the independent S-box assumption.

Given that $\mathcal{X}_S(\texttt{0x5}, \texttt{0xa}) = \{\texttt{0x2}, \texttt{0x3}, \texttt{0x6}, \texttt{0x7}\}$, $\mathcal{Y}_S(\texttt{0x5}, \texttt{0xa}) = \{\texttt{0x6}, \texttt{0x7}, \texttt{0xc}, \texttt{0xd}\}$ and $\mathcal{X}_S(\texttt{0xf}, \texttt{0xa}) = \{\texttt{0x0}, \texttt{0xf}\}$, p_{exact} in eq. (11) is 2^{-1} instead of 2^{-2} and p_{exact} in eq. (12) is 2^{-2} instead of 2^{-3}.

Experiments. To ensure our estimates match reality, we performed some computational verification of the above differential characteristic:

- 1 round of RoadRunneR-80 yielded 65870 ($\approx 2^{16}$) matches over 2^{24} trials;
- 2 rounds of RoadRunneR-80 yielded 1011 ($\approx 2^{10}$) matches over 2^{27} trials;
- 3 rounds of RoadRunneR-80 yielded 124 ($\approx 2^{7}$) matches over 2^{32} trials;
- 4 rounds of RoadRunneR-80 yielded 28 ($\approx 2^{5}$) matches over 2^{37} trials;
- 5 rounds of RoadRunneR-80 yielded 16 ($= 2^{4}$) matches over 2^{43} trials.

These results are summarized in Table 1.

4 Extension to Almost-MDS Matrix in Minalpher-P

In this section, we show that improving the probability by evaluating p_{exact} can be extended to SPN with almost-MDS binary matrices. An example of such matrices is

$$\begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}, \tag{13}$$

which is actually adopted by Minalpher [STA+14]. The rotated version of the above matrix is more popular, which can be seen in several designs e.g. PRINCE [BCG+12], FIDES [BBK+13], and Midori [BBI+15]. Section 4.1 provides an overview of our observation. Section 4.2 introduces the specification of Minalpher-P. Section 4.3 introduces the previous best differential characteristic evaluated by p_{ind}. Section 4.4 improves the probability by evaluating p_{exact} and extends the attack by two rounds.

4.1 Overview

Let us consider a 1-column state consisting of four cells of size n bits, thus the state size is $4n$ bits. Suppose that the state is updated by an SPN, in which the S-layer applies an n-bit S-box to all of four cells and the P-layer applies the matrix in Eq. (13). With this structure, the number of active cells can be two per rounds owing to the following property: *When two cells have an identical difference, the matrix multiplication does not change the number of active cells and the differential value.*

Let us consider the 2-round characteristic shown in Figure 6, which assumes that $\Pr_{x\in\mathbb{F}_2^n}[S(x)\oplus S(x\oplus\Delta a)=\Delta b]=2^{-n+2}$ and $\Pr_{x\in\mathbb{F}_2^n}[S(x)\oplus S(x\oplus\Delta b)=\Delta c]=2^{-n+2}$. p_{ind} is $(2^{-n+2})^4$ because of the four active S-boxes, meanwhile we show that p_{exact} is

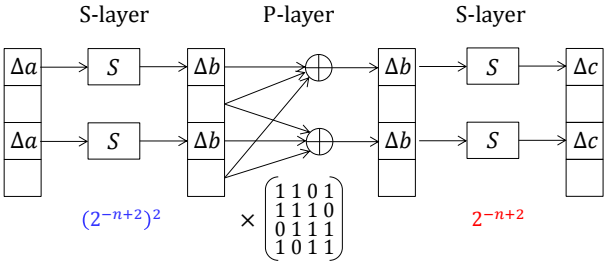

Figure 6: Overview: 2-round characteristic in SPN with single column.

$(2^{-n+2})^3$ in which the S-layer can be satisfied only with 2^{-n+2} from the second round. The state of SPN ciphers usually have more columns, thus the improvement by a factor of 2^{-n+2} can be amplified, which makes the improved factor significantly large.

4.2 Specification of Minalpher-P

The core part of Minalpher is the Even-Mansour construction in which a 256-bit plaintext is masked by a 256-bit secret value, and then a nibble-wise 256-bit permutation called Minalpher-P is computed. Finally, the output of Minalpher-P is masked by the 256-bit secret value. A 256-bit state is described as two 4×8 nibble-matrices denoted by A and B.

Let A_{i-1} and B_{i-1} be the inputs of the round function for round i. The states are updated to A_i and B_i with a round function, which consists of SubNibbles (SN), ShuffleRows (SR), SwapMatrices (SM), XorMatrix (XM) and MixColumns (MC), where

SN, SR and MC are functions from $\left\{\mathbb{F}_2^4\right\}^{4\times 8}$ to $\left\{\mathbb{F}_2^4\right\}^{4\times 8}$. In the end, the state is xored with the round constant. We use notations A^{op} and B^{op} to denote $\left\{\mathbb{F}_2^4\right\}^{4\times 8}$ data after operation op. See Figure 7 for its illustration.

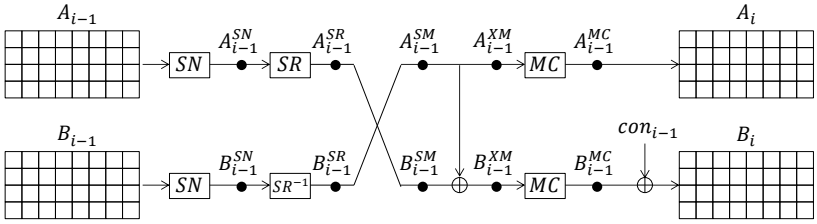

Figure 7: Illustration of the round function of Minalpher-P.

SubNibbles (SN). SN substitutes each nibble by using 4-bit involution S-box S.

x	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
$S(x)$	B	3	4	1	2	8	C	F	5	D	E	0	6	9	A	7

ShuffleRows (SR). SR shuffles nibble positions within each row. SR consists of two shuffle functions SR_1 and SR_2 defined as follows. Elements in 4×8 matrix A are moved according to the table below, and for B, SR^{-1} is applied instead of SR.

i	0	1	2	3	4	5	6	7
$SR_1(i)$	6	7	1	0	2	3	4	5
$SR_2(i)$	4	5	0	1	7	6	2	3
$SR_1^{-1}(i)$	3	2	4	5	6	7	0	1
$SR_2^{-1}(i)$	2	3	6	7	0	1	5	4

SwapMatrices (SM). SM swaps the matrix A and the matrix B.

XorMatrix (XM). The matrix B is xored with the matrix A.

MixColumns (MC). MC is a column-wise linear operation. As introduced before, MC is expressed as a multiplication by the matrix in Eq. (13).

Round Constant. The round constant con_{i-1} is xored to the matrix B. In this paper, the fact that the matrix A is not updated by round constant is important.

4.3 Differential Characteristics of Minalpher-P

The designers of Minalpher found a 6-round iterative truncated differential with 64 active S-boxes, which is shown in Figure 8. Note that this is not the one with minimal number of active S-boxes for 6 rounds. However, if it is iterated beyond 6 rounds, the number of active S-boxes matches the lower bound obtained by automated search.

Then, we convert the truncated differential to a specific characteristic by fixing the differential values. By calculating DDT of the 4-bit S-box, we observe that the input difference `0x4` will be mapped to the output difference `0x4` with probability 2^{-2}. So, we replace all filled cells in Figure 8 with the particular difference `0x4`.

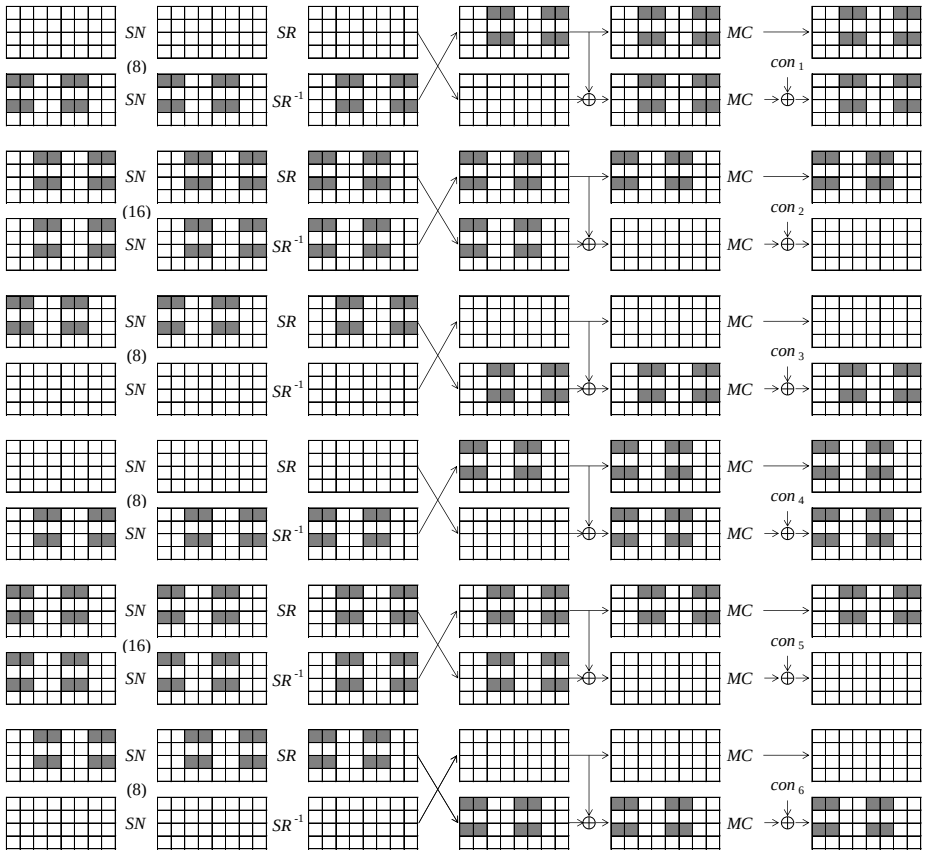

Figure 8: 6-round iterative truncated differential of Minalpher-P. Filled and empty cells denote active and inactive nibbles, respectively. Note that we rotated the original 6-round iterative characteristic by one round to optimize it in our analysis.

Let us evaluate the probability of the 6-round characteristic. Here we assume that the secret mask of the Even-Mansour construction prevents the attacker from choosing the plaintext or ciphertext to deterministically satisfy differential propagations through S-box in the first and the last rounds. The linear part is satisfied with probability 1, thus the probability only comes from the S-box, which is 2^{-2} per S-box. Because $8+16+8+8+16+8=64$ S-boxes are included in the characteristic, the probability is $(2^{-2})^{64}=2^{-128}$ when all transitions through all S-boxes are assumed to be independent. Considering that the security of Minalpher is claimed up to 128 bits, extending the characteristic by a few more rounds is impossible.

4.4 Analysis of Exact Probability

Preliminaries. Recall that for any pair (a,b) of differences, we use the following notation: $\mathcal{X}_S(a,b)=\{x\in\mathbb{F}_2^4 : S(x)\oplus S(x\oplus a)=b\}$ and $\mathcal{Y}_S(a,b)=\{S(x)\in\mathbb{F}_2^4 : S(x)\oplus S(x\oplus a)=b\}$. When S is involution as in Minalpher-P, $\mathcal{X}_S(a,a)$ is equal to $\mathcal{Y}_S(a,a)$ for any a. In particular, when $a=\mathtt{4}$ in the S-box of Minalpher-P, $\mathcal{X}_S(\mathtt{4},\mathtt{4})=\mathcal{Y}_S(\mathtt{4},\mathtt{4})=\{\mathtt{9},\mathtt{a},\mathtt{d},\mathtt{e}\}$. This is represented by an affine space $\langle\mathtt{3},\mathtt{4}\rangle+\mathtt{9}$, where $\langle x,y\rangle$ is a linear subspace.

Analysis of $p_{\mathbf{exact}}$. Here, we show that the probability of the 6-round characteristic is actually 2^{-96} instead of 2^{-128}, thus the number of attacked rounds can be extended. We

begin with the analysis of the simple case; SN and MC are iterated twice in a column, which is shown in Figure 9.

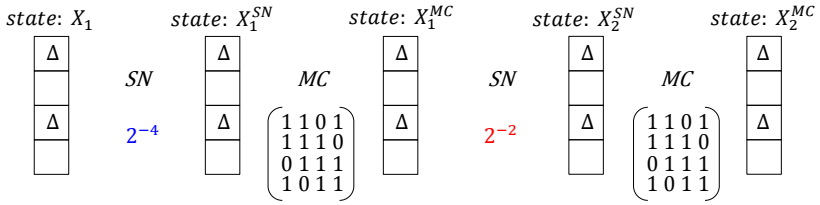

Figure 9: Analysis of simple case ($\Delta = \texttt{0x4}$). Probability is 2^{-8} if two SN operations are evaluated independently, while the exact probability is 2^{-6}.

As shown in Figure 9, the five states are denoted by $X_1, X_1^{SN}, X_1^{MC}, X_2^{SN}, X_2^{MC}$. Suppose that the 4-nibble value of X_1 is chosen uniformly at random. Then the probability of satisfying the first SN layer is $(2^{-2})^2 = 2^{-4}$. When this occurs, the value of $X_1^{SN}[0]$ and $X_1^{SN}[2]$ are limited to four choices in $\mathcal{Y}_S(4,4) = \{\texttt{9}, \texttt{a}, \texttt{d}, \texttt{e}\}$. From the specification of MC, the value of active nibbles in X_1^{MC} are calculated as

$$X_1^{MC}[0] = X_1^{SN}[0] \oplus X_1^{SN}[1] \oplus X_1^{SN}[3],$$
$$X_1^{MC}[2] = X_1^{SN}[1] \oplus X_1^{SN}[2] \oplus X_1^{SN}[3].$$

In order to satisfy the differential propagation in the second SN operation, both of $X_1^{MC}[0]$ and $X_1^{MC}[2]$ must be in the affine space of $\mathcal{X}_S(4,4) = \{\texttt{9}, \texttt{a}, \texttt{d}, \texttt{e}\}$. Considering that $X_1^{SN}[0]$ and $X_1^{SN}[2]$ are in the affine space, the condition that both of $X_1^{MC}[0]$ and $X_1^{MC}[2]$ are in the same affine space is $X_1^{SN}[1] \oplus X_1^{SN}[3]$ is in its linear subspace $\langle 3, 4\rangle = \{\texttt{0}, \texttt{3}, \texttt{4}, \texttt{7}\}$. This occurs with probability 2^{-2}, thus the probability of satisfying the second SN layer is 2^{-2}, instead of 2^{-4}.

Application to 6-Round Characteristic. All the differences in Figure 8 are fixed to `0x4`.

Round 1. Suppose that the lower half of the input state, B_0, is chosen uniformly at random. Then, the probability of satisfying the SN layer in round 1 is $(2^{-2})^8 = 2^{-16}$.

Round 2. The SR operation does not mix the value, thus irrelevant to this analysis. The state B_0^{SN} is next updated by MC and then passed to SN in round 2. Namely, the simple column-wise analysis discussed above appears in four columns. Thus the probability that the differences in A_1 are propagated to A_1^{SN} is $(2^{-2})^4 = 2^{-8}$ instead of 2^{-16}. Note that B_0^{XM} is xored with random state value B_0^{SM} and round constant, thus the probability between B_1 and B_1^{SN} is 2^{-16}. In total, the probability of round 2 is 2^{-24}.

Round 3. The same event as round 2 occurs. Namely B_1^{SN} is updated with MC and then SN in round 3. As discussed before, this probability is 2^{-8} instead of 2^{-16}.

Rounds 4–6. The probabilities for rounds 4, 5, and 6 are calculated round by round. The analysis becomes almost the same as round 1, 2, and 3, respectively because of the similarity of the active S-boxes positions. To avoid redundancy, we omit the round-by-round explanation. In the end, the probability for those rounds is $2^{-16-24-8} = 2^{-48}$.

From the above discussion we conclude that the probability of the 6-round differential characteristic in Figure 8 is 2^{-96}, which is significantly larger than p_{ind} of 2^{-128}.

Experimental Verification. The probability of the first three rounds already reach 2^{64}, which is infeasible in our environment. The gap between p_{ind} and p_{exact} first appears in state A_1^{SB} of the SN operation in the second round, which is independent of the propagation in state B_1^{SB}. We thus implement the state update from B_0^{SB} to A_1^{SB} with the limitation that values of active bytes are sampled randomly from $\mathcal{Y}_S(4,4)$.

We generated $65,536 (= 2^{16})$ random values at B_0^{SB}, and $250 (\approx 2^8)$ values satisfy the difference in A_1^{SB}, which confirms that the probability of the characteristic from B_0^{SB} to A_1^{SN} is actually $(2^{-2})^4 = 2^{-8}$ instead of $(2^{-4})^4 = 2^{-16}$.

Extension to 8 Rounds. We append 1 round to both of the beginning and the end of the 6-round iterative characteristic in Figure 8. Remember that the probability of the first round in the 6-round characteristic is 2^{-16}. Due to the iterative structure, with the same reason, the probability of the last extended round is 2^{-16}. The extended round at the beginning has eight active S-boxes. Because the advantage of p_{exact} cannot be exploited at the beginning, the probability is $(2^{-2})^8 = 2^{-16}$.

To conclude, the probability of the 8-round characteristic is $2^{-96-16-16} = 2^{-128}$. Considering that the previous 6-round characteristic has the same probability, we improved the previous attack by 2 rounds.

Note that a path with probability 2^{-128} cannot be a straightforward distinguisher with 2^{128} queries. Here our main focus is improving the previous analysis, and using the path with probability 2^{-128} is the same setting as the designers of Minalpher. Moreover, by combining with similar paths, the probability may be amplified to be greater than 2^{-128}.

5 Concluding Remarks

This paper studied the interaction between the differential transitions occurring in the multiple rounds of a fixed-key or unkeyed primitive. We showed that assuming independent input values for each S-box does not correspond to the actual situation, and p_{exact} can be much larger than p_{ind}. Our general analysis on the Feistel network showed that the gap between p_{exact} and p_{ind} depends on the S-box size and the S-box choice. In addition, having non-zero gap is inevitable when the S-box has differential uniformity 4 and a size larger than six bits (unless one Sbox is inactive).

This observation actually impacts the security of practical algorithms. We applied it to the lightweight block cipher RoadRunneRand the authenticated encryption scheme Minalpher. The results showed that with p_{exact} the number of attacked rounds could be improved compared to the evaluation with p_{ind}.

Symmetric-key primitives with unkeyed functions or public permutations are getting more popular due to its lightweight property and can be seen in many contemporary structures such as the sponge and the Even-Mansour constructions. This paper alerts us that the resistance against differential cryptanalysis needs to be analyzed carefully.

Acknowledgments

This work has been initiated during the Lorentz Center workshop on “High-Security Lightweight Cryptography”, held in Leiden, the Netherlands, in October 2016 and we would like to thank the organizers for inviting us. We also thank the anonymous reviewers for their careful reading and their valuable comments.

References

[BBI+15] Subhadeep Banik, Andrey Bogdanov, Takanori Isobe, Kyoji Shibutani, Harunaga Hiwatari, Toru Akishita, and Francesco Regazzoni. Midori: A block cipher for low energy. In Tetsu Iwata and Jung Hee Cheon, editors, *Advances in Cryptology - ASIACRYPT 2015 - 21st International Conference on the Theory and Application of Cryptology and Information Security, Auckland, New Zealand, November 29 - December 3, 2015, Proceedings, Part II*, volume 9453 of *Lecture Notes in Computer Science*, pages 411–436. Springer, 2015.

[BBK+13] Begül Bilgin, Andrey Bogdanov, Miroslav Knezevic, Florian Mendel, and Qingju Wang. FIDES: Lightweight authenticated cipher with side-channel resistance for constrained hardware. In Guido Bertoni and Jean-Sébastien Coron, editors, *Cryptographic Hardware and Embedded Systems - CHES 2013 - 15th International Workshop, Santa Barbara, CA, USA, August 20-23, 2013. Proceedings*, volume 8086 of *Lecture Notes in Computer Science*, pages 142–158. Springer, 2013.

[BBL13] Céline Blondeau, Andrey Bogdanov, and Gregor Leander. Bounds in shallows and in miseries. In Ran Canetti and Juan A. Garay, editors, *Advances in Cryptology - CRYPTO 2013*, volume 8042 of *Lecture Notes in Computer Science*, pages 204–221. Springer, 2013.

[BCG+12] Julia Borghoff, Anne Canteaut, Tim Güneysu, Elif Bilge Kavun, Miroslav Knezevic, Lars R. Knudsen, Gregor Leander, Ventzislav Nikov, Christof Paar, Christian Rechberger, Peter Rombouts, Søren S. Thomsen, and Tolga Yalçin. PRINCE - A low-latency block cipher for pervasive computing applications - extended abstract. In Xiaoyun Wang and Kazue Sako, editors, *Advances in Cryptology - ASIACRYPT 2012 - 18th International Conference on the Theory and Application of Cryptology and Information Security, Beijing, China, December 2-6, 2012. Proceedings*, volume 7658 of *Lecture Notes in Computer Science*, pages 208–225. Springer, 2012.

[BG10] Céline Blondeau and Benoît Gérard. Links between theoretical and effective differential probabilities: Experiments on PRESENT. Cryptology ePrint Archive, Report 2010/261, 2010. http://eprint.iacr.org/2010/261.

[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.

[BS90] Eli Biham and Adi Shamir. Differential cryptanalysis of DES-like cryptosystems. In Alfred Menezes and Scott A. Vanstone, editors, *Advances in Cryptology - CRYPTO '90, 10th Annual International Cryptology Conference, Santa Barbara, California, USA, August 11-15, 1990, Proceedings*, volume 537 of *Lecture Notes in Computer Science*, pages 2–21. Springer, 1990.

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

[BS15] Adnan Baysal and Sühap Sahin. RoadRunneR: A small and fast bitslice block cipher for low cost 8-bit processors. In Tim Güneysu, Gregor Leander, and

Amir Moradi, editors, *Lightweight Cryptography for Security and Privacy - 4th International Workshop, LightSec 2015, Bochum, Germany, September 10-11, 2015, Revised Selected Papers*, volume 9542 of *Lecture Notes in Computer Science*, pages 58–76. Springer, 2015.

[BSS+13] Ray Beaulieu, Douglas Shors, Jason Smith, Stefan Treatman-Clark, Bryan Weeks, and Louis Wingers. The SIMON and SPECK families of lightweight block ciphers. Cryptology ePrint Archive, Report 2013/404, 2013. http://eprint.iacr.org/2013/404.

[CDK09] Christophe De Cannière, Orr Dunkelman, and Miroslav Knezevic. KATAN and KTANTAN - A family of small and efficient hardware-oriented block ciphers. In Christophe Clavier and Kris Gaj, editors, *Cryptographic Hardware and Embedded Systems - CHES 2009, 11th International Workshop, Lausanne, Switzerland, September 6-9, 2009, Proceedings*, volume 5747 of *Lecture Notes in Computer Science*, pages 272–288. Springer, 2009.

[CDL15] Anne Canteaut, Sébastien Duval, and Gaëtan Leurent. Construction of lightweight S-Boxes using Feistel and MISTY structures. In Orr Dunkelman and Liam Keliher, editors, *Selected Areas in Cryptography - SAC 2015 - 22nd International Conference, Sackville, NB, Canada, August 12-14, 2015, Revised Selected Papers*, volume 9566 of *Lecture Notes in Computer Science*, pages 373–393. Springer, 2015.

[CR15] Anne Canteaut and Joëlle Roué. On the behaviors of affine equivalent Sboxes regarding differential and linear attacks. In Elisabeth Oswald and Marc Fischlin, editors, *Advances in Cryptology - EUROCRYPT 2015, Proceedings, Part I*, volume 9056 of *Lecture Notes in Computer Science*, pages 45–74. Springer, 2015.

[DDGS15] Itai Dinur, Orr Dunkelman, Masha Gutman, and Adi Shamir. Improved top-down techniques in differential cryptanalysis. In Kristin E. Lauter and Francisco Rodríguez-Henríquez, editors, *Progress in Cryptology - LATINCRYPT 2015 - 4th International Conference on Cryptology and Information Security in Latin America, Guadalajara, Mexico, August 23-26, 2015, Proceedings*, volume 9230 of *Lecture Notes in Computer Science*, pages 139–156. Springer, 2015.

[DKS12] Orr Dunkelman, Nathan Keller, and Adi Shamir. Minimalism in cryptography: The Even-Mansour scheme revisited. In David Pointcheval and Thomas Johansson, editors, *Advances in Cryptology - EUROCRYPT 2012 - 31st Annual International Conference on the Theory and Applications of Cryptographic Techniques, Cambridge, UK, April 15-19, 2012. Proceedings*, volume 7237 of *Lecture Notes in Computer Science*, pages 336–354. Springer, 2012.

[DR06] Joan Daemen and Vincent Rijmen. Understanding two-round differentials in AES. In Roberto De Prisco and Moti Yung, editors, *Security and Cryptography for Networks, SCN 2006*, volume 4116 of *Lecture Notes in Computer Science*, pages 78–94. Springer, 2006.

[DR07] Joan Daemen and Vincent Rijmen. Plateau characteristics. *IET Information Security*, 1(1):11–17, 2007.

[EM91] Shimon Even and Yishay Mansour. A construction of a cipher from a single pseudorandom permutation. In Hideki Imai, Ronald L. Rivest, and Tsutomu Matsumoto, editors, *Advances in Cryptology - ASIACRYPT '91, International Conference on the Theory and Applications of Cryptology, Fujiyoshida, Japan,*

November 11-14, 1991, Proceedings, volume 739 of *Lecture Notes in Computer Science*, pages 210–224. Springer, 1991.

[EM97] Shimon Even and Yishay Mansour. A construction of a cipher from a single pseudorandom permutation. *J. Cryptology*, 10(3):151–162, 1997.

[GNL11] Zheng Gong, Svetla Nikova, and Yee Wei Law. KLEIN: A new family of lightweight block ciphers. In Ari Juels and Christof Paar, editors, *RFID. Security and Privacy - 7th International Workshop, RFIDSec 2011, Amherst, USA, June 26-28, 2011, Revised Selected Papers*, volume 7055 of *Lecture Notes in Computer Science*, pages 1–18. Springer, 2011.

[GPPR11] Jian Guo, Thomas Peyrin, Axel Poschmann, and Matthew J. B. Robshaw. The LED block cipher. In Bart Preneel and Tsuyoshi Takagi, editors, *Cryptographic Hardware and Embedded Systems - CHES 2011*, volume 6917 of *Lecture Notes in Computer Science*, pages 326–341. Springer, 2011.

[HLL+00] Seokhie Hong, Sangjin Lee, Jongin Lim, Jaechul Sung, Dong Hyeon Cheon, and Inho Cho. Provable security against differential and linear cryptanalysis for the SPN structure. 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 273–283. Springer, 2000.

[KMT01] Liam Keliher, Henk Meijer, and Stafford E. Tavares. New method for upper bounding the maximum average linear hull probability for SPNs. 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 420–436. Springer, 2001.

[KS07] Liam Keliher and Jiayuan Sui. Exact maximum expected differential and linear probability for two-round Advanced Encryption Standard. *IET Information Security*, 1(2):53–57, 2007.

[LMM91] Xuejia Lai, James L. Massey, and Sean Murphy. Markov ciphers and differential cryptanalysis. In Donald W. Davies, editor, *Advances in Cryptology - EUROCRYPT '91, Workshop on the Theory and Application of of Cryptographic Techniques, Brighton, UK, April 8-11, 1991, Proceedings*, volume 547 of *Lecture Notes in Computer Science*, pages 17–38. Springer, 1991.

[LW14] Yongqiang Li and Mingsheng Wang. Constructing S-boxes for lightweight cryptography with Feistel structure. In Lejla Batina and Matthew Robshaw, editors, *Cryptographic Hardware and Embedded Systems - CHES 2014 - 16th International Workshop, Busan, South Korea, September 23-26, 2014. Proceedings*, volume 8731 of *Lecture Notes in Computer Science*, pages 127–146. Springer, 2014.

[Mat96] Mitsuru Matsui. New structure of block ciphers with provable security against differential and linear cryptanalysis. In Dieter Gollmann, editor, *Fast Software Encryption, Third International Workshop, Cambridge, UK, February 21-23, 1996, Proceedings*, volume 1039 of *Lecture Notes in Computer Science*, pages 205–218. Springer, 1996.

[MP13] Nicky Mouha and Bart Preneel. Towards finding optimal differential characteristics for ARX: Application to Salsa20. Cryptology ePrint Archive, Report 2013/328, 2013. http://eprint.iacr.org/2013/328.

[NK92] Kaisa Nyberg and Lars R. Knudsen. Provable security against differential cryptanalysis. In Ernest F. Brickell, editor, *Advances in Cryptology - CRYPTO '92, 12th Annual International Cryptology Conference, Santa Barbara, California, USA, August 16-20, 1992, Proceedings*, volume 740 of *Lecture Notes in Computer Science*, pages 566–574. Springer, 1992.

[Nyb94] Kaisa Nyberg. S-boxes and round functions with controllable linearity and differential uniformity. In Bart Preneel, editor, *Fast Software Encryption: Second International Workshop. Leuven, Belgium, 14-16 December 1994, Proceedings*, volume 1008 of *Lecture Notes in Computer Science*, pages 111–130. Springer, 1994.

[PSC+02] Sangwoo Park, Soo Hak Sung, Seongtaek Chee, E-Joong Yoon, and Jongin Lim. On the security of Rijndael-like structures against differential and linear cryptanalysis. In Yuliang Zheng, editor, *Advances in Cryptology - ASIACRYPT 2002, 8th International Conference on the Theory and Application of Cryptology and Information Security, Queenstown, New Zealand, December 1-5, 2002, Proceedings*, volume 2501 of *Lecture Notes in Computer Science*, pages 176–191. Springer, 2002.

[PSLL03] Sangwoo Park, Soo Hak Sung, Sangjin Lee, and Jongin Lim. Improving the upper bound on the maximum differential and the maximum linear hull probability for SPN structures and AES. In Thomas Johansson, editor, *Fast Software Encryption, 10th International Workshop, FSE 2003, Lund, Sweden, February 24-26, 2003, Revised Papers*, volume 2887 of *Lecture Notes in Computer Science*, pages 247–260. Springer, 2003.

[Riv11] Ronald L. Rivest. The invertibility of the XOR of rotations of a binary word. *Int. J. Comput. Math.*, 88(2):281–284, 2011.

[STA+14] Yu Sasaki, Yosuke Todo, Kazumaro Aoki, Yusuke Naito, Takeshi Sugawara, Yumiko Murakami, Mitsuru Matsui, and Shoichi Hirose. Minalpher v1. CAESAR Round 1 submission, 2014.

[UCI+11] Markus Ullrich, Christophe De Cannière, Sebastiaan Indesteege, Özgül Küçük, Nicky Mouha, and Bart Preneel. Finding optimal bitsliced implementations of 4×4-bit S-boxes. In *SKEW 2011 Symmetric Key Encryption Workshop, Copenhagen, Denmark, 16–17 February*, 2011.