A General Framework for the Related-key Linear Attack against Block Ciphers with Linear Key Schedules

Jung-Keun Lee, Bonwook Koo, and Woo-Hwan Kim

{jklee,bwkoo,whkim5}@nsr.re.kr

Abstract. We present a general framework for the related-key linear attack that can be applied to iterative block ciphers with linear key schedules. The attack utilizes a newly introduced *related-key linear approximation* that is obtained directly from a linear trail. The attack makes use of a known related-key data consisting of triplets of a plaintext, a ciphertext, and a key difference such that the ciphertext is the encrypted value of the plaintext under the key that is the xor of the key to be recovered and the specified key difference. If such a block cipher has a linear trail with linear correlation ϵ, it admits attacks with related-key data of size $O(\epsilon^{-2})$ just as in the case of classical Matsui's Algorithms. But since the attack makes use of a related-key data, the attacker can use a linear trail with the squared correlation less than 2^{-n}, n being the block size, in case the key size is larger than n. Moreover, the standard key hypotheses seem to be appropriate even when the trail is not dominant as validated by experiments.
The attack can be applied in two ways. First, using a linear trail with squared correlation smaller than 2^{-n}, one can get an effective attack covering more rounds than existing attacks against some ciphers, such as Simon48/96, Simon64/128 and Simon128/256. Secondly, using a trail with large squared correlation, one can use related-key data for key recovery even when the data is not suitable for existing linear attacks.

Keywords: related-key attack, linear cryptanalysis, linear key schedule, Simon

1 Introduction

In recent years many lightweight block ciphers have been proposed targeting resource-constrained platforms. They adopt simple key schedules to get competitive performance figures in terms of the resource requirements. In this regard not a few of them have linear key schedules. (e.g. Gift [3], Skinny [6], Midori [2], Simon [4], Zorro [21], Prince [14], Led [22], Piccolo [39], Katan [15].) However, there are little cryptanalytic techniques that are applicable to general block ciphers of such a kind. In this work, we will present a framework for the related-key linear attack that can be applied to generic iterative block ciphers with linear key

Table 1: Attack results on SIMON

cipher (# rounds)	# attacked rounds	computation	data	$\Pr_{\text{success}}$	Attack	Ref.
SIMON32/64	**23**	$\mathbf{2^{46.65}}$	$\mathbf{2^{46.3}}$	0.5	RKLC	Here
(32)	23	$2^{56.3}$	$2^{31.19}$	0.28	LC	[16]
	21	$2^{55.25}$	2^{31}	0.51	DC	[41]
SIMON48/96	**28**	$\mathbf{2^{71.07}}$	$\mathbf{2^{70.9}}$	0.5	RKLC	Here
(36)	25	$2^{88.28}$	$2^{47.92}$	0.445*	LC	[16]
	24	$2^{87.25}$	2^{47}	0.48	DC	[41]
SIMON64/128	**34**	$\mathbf{2^{95.5}}$	$\mathbf{2^{95.32}}$	0.5	RKLC	Here
(44)	31	2^{120}	$2^{63.53}$	0.316*	LC	[16]
	29	$2^{116.25}$	2^{63}	0.46	DC	[41]
SIMON128/256	**62**	$\mathbf{2^{190.76}}$	$\mathbf{2^{190.4}}$	0.5	RKLC	Here
(72)	**55**	$\mathbf{2^{175}}$	$\mathbf{2^{174.73}}$	0.5	RKLC	Here
	53	$2^{248.01}$	$2^{127.6}$	0.315*	LC	[16]
	50	$2^{247.25}$	2^{127}	0.48	DC	[41]

* estimates based on [8] under an assumption on the distribution of correlations [19]

schedules. Since the linear attack was publicized by M. Matsui [33], there have been many extensions such as attacks using linear hulls [34, 36], multiple linear attacks [7, 27], multidimensional linear attacks [17, 23, 25] and zero-correlation attacks [13]. Though there are lots of works regarding the related-key attacks against block ciphers using differential characteristics, there are not many dealing with the related-key linear attacks. The current work tries to address this issue.

1.1 Our Contributions

- We present a general framework for the related-key linear attack that is applicable to block ciphers with linear key schedules. It is based on classical Matsui's Algorithms and makes use of a related-key linear approximations that can be obtained from an ordinary linear trail in a straightforward way. We also provide a statistical model for the attack from which we derive estimates for the success probability and the attack complexities.
- We present experimental results that confirm the validity of our framework including the appropriateness of the statistical model we presume. We consider small-scale variants of SIMON and a variant of PRESENT with a linear key schedule for the experiments.
- We present related-key linear attacks on SIMON whose results are summarized in Table 1. The attacks cover more rounds than existing attacks and can be regarded to be better than the generic related-key attack [28] with known key differences and random plaintexts in terms of the attack complexities.

1.2 Related Works

Related-key linear attacks. The idea of using related keys in linear attacks appears in a small number of previous works. P. Vora et al. [40] described an attack against a round-reduced DES based on a coding theory framework claiming that using related keys one can marginally improve the single-key linear attacks. M. Hermelin et al. [24] claim a related-key linear attack against the full PRESENT-128 using very special types of chosen key differences based on some assumption regarding the capacity of multidimensional approximations. A. Bogdadov et al. [10] presented a key recovery attack using related-key linear distinguishers with chosen key differences. Their method works against block ciphers whose key schedules admit certain invariance property. The related-key attack presented in this work uses keys with known differences though it works against block ciphers with linear key schedules.

Linear attacks using a linear approximation with the small correlation. A linear approximation of a block cipher with correlation $< 2^{-n/2}$ is usually considered not of much use except when it is exploited in a multiple linear attack together with other approximations. C. Beierle et al. [6] argue that the SKINNY ciphers are secure against related-tweakey linear attacks by presenting bounds on the correlations of linear trails as the number of rounds increases, taking into account the fact that the attacker may utilize the tweakey as the additional data source. T. Kranz et al. [31] show that the linear tweak trails in such ciphers which can be used in the linear attack are exactly those coming from the ordinary trails. T. Ashur et al. [1] described a χ^2 distinguisher detecting correlations smaller than $2^{-n/2}$ in the multi-key setting. But they were not able to use the distinguisher for key recovery.

Generic related-key attacks. J. Kelsey et al. [28] mentioned a related-key attack that can be applied to any block cipher, referring to [42]. The attack uses keys with known differences and the product of the number of related keys and the computational complexity is 2^k in the attack. But the plaintexts are required to be the same regardless of the keys in the attack. So in the known plaintext setting the attacker needs to get about $M2^n$ pairs of key difference and plaintext to get $M \gg 1$ related keys with the same plaintext. Thus the product of the data size and the computational complexity is about 2^{k+n} in such a setting.

1.3 Organization of the Paper

In Sect. 2 we introduce the terminology and notations used in the paper. In Sect. 3 we describe the framework for the related-key linear cryptanalysis against block ciphers with linear key schedules. In Sect. 4 we present attack results on SIMON obtained from the framework presented in Sect. 3 together with some dedicated analysis. In Sect. 5 we provide experimental results that corroborate the claims of the paper. In Sect. 6 we discuss the validity and the usefulness of the framework in more detail. We conclude in Sect. 7.

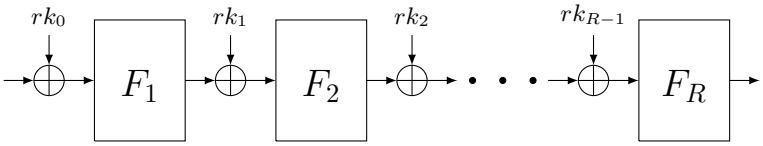

Fig. 1: a long-key cipher

2 Terminology and Notations

$\mathbb{F}_2$ and $\mathbb{Z}$ denote the field with two elements and the ring of integers, respectively. A word is a bit string of the length w=12, 16, 24, 32, or 64. For integers i, j with $i \leq j$, $[i..j]$ denotes the set of integers x such that $i \leq x \leq j$. The LSB (least significant bit) of a word is indexed as 0 and is located at the rightmost position. The $(i+1)$-th rightmost bit of a word x is denoted by $x[i]$ so that $x[0]$ denotes the LSB of x. For a w-bit word x, $x[i]$ with $i \notin [0..(w-1)]$ means $x[i \mod w]$. Also $x[i..j]$ denotes the bit string $x[j] \| \ldots \| x[i]$ for $0 \leq i \leq j < w$. $x \lll a$ and $x \ggg a$ denote the circular shift of a word x to the left and right by a bits, respectively. $\wedge$ represents the bitwise-and of two words. The inner product of a w-bit mask γ and a w-bit value x is defined to be $\oplus_{i=0}^{w-1} \gamma[i]x[i]$ and is denoted by $\langle \gamma, x \rangle$. For a Boolean function $G : \mathbb{F}_2^l \to \mathbb{F}_2$, the correlation of G is defined to be the imbalance $(|\{x : G(x) = 0\}| - |\{x : G(x) = 1\}|)/2^l$. For a vectorial Boolean function $F : \mathbb{F}_2^l \to \mathbb{F}_2^m$, an l-bit mask γ, and an m-bit mask λ, the (linear) correlation of F with respect to the mask pair (γ, λ) is defined to be the correlation of the Boolean function G given by $G(x) = \langle \gamma, x \rangle \oplus \langle \lambda, F(x) \rangle$ and is denoted by $\varepsilon_F(\gamma, \lambda)$. The Hamming weight of a word x, denoted by $\mathrm{wt}(x)$, is the number of the nonzero bits of x. For a bit string X with even length, X_{L} and X_{R} denote the left half and right half of X, respectively. The support of a w-bit word x is defined to be the set of indices $\{i \in [0..(w-1)] : x[i] \neq 0\}$ and is denoted by $\mathsf{supp}(x)$. $\|$ denotes the concatenation of bit strings. Bit strings are expressed in the hexadecimal representations. For example, `c201` represents the bit string $1, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1$. For real numbers μ and $\sigma > 0$, $\mathcal{N}(\mu, \sigma^2)$ denotes the normal distribution with the mean μ and the standard deviation σ. Φ denotes the cumulative distribution function of the standard normal distribution.

3 Description of the Framework

In our related-key linear attack, the attacker takes advantage of related-key data such that each entry in the data is a triplet $(P, C, \Delta K)$ of a plaintext P, a ciphertext C, and a key difference ΔK for which the ciphertext is the encrypted value $E_{K^* \oplus \Delta K}(P)$ of the plaintext under the key that is the xor of the unknown base key K^* to be recovered and the key difference ΔK. The attack proceeds as in the classical Matsui's Algorithms [33]. In the classical Algorithm 2 using a linear trail, for example, the attacker uses a linear approximation that involves

masked intermediate values and a parity bit expressed as an xor of masked base round keys. But in our related-key linear attack, the attacker makes use of a *related-key linear approximation* that involves masked key differences together with the masked intermediate values and the parity bit. The related-key linear attack is based on the following features of the block ciphers with linear key schedules:

- When a key K is the xor of an unknown base key K^* and a known key difference ΔK, the difference of round keys derived from K and K^* can be computed directly from ΔK though two keys are unknown.
- The intermediate state obtained from a plaintext by performing several encryption rounds with a key K can be also computed using P, ΔK, and the round keys derived from K^*. The same holds for the decryption rounds.

3.1 A Related-key Linear Approximation

Let R, r, and s be integers with $0 \le s \le s+r \le R$ and let $E : \mathbb{F}_2^k \times \mathbb{F}_2^n \to \mathbb{F}_2^n$ be an R-round key-alternating iterative block cipher with k-bit keys and n-bit blocks. Let $\tilde{E}$ be the long-key cipher corresponding to E and ψ be the key scheduling function that is linear. That is, $\tilde{E}$ is a function $\mathbb{F}_2^{Rn} \times \mathbb{F}_2^n \to \mathbb{F}_2^n$ defined by

$$\tilde{E}(rk_0 \| rk_1 \| \cdots \| rk_{R-1}, x) = F_R(rk_{R-1} \oplus \cdots F_2(rk_1 \oplus F_1(rk_0 \oplus x)) \cdots)$$

as in Fig. 1, where each F_i is a fixed n-bit permutation, ψ is a function $\mathbb{F}_2^k \to \mathbb{F}_2^{Rn}$, and $E(K, x) = \tilde{E}(\psi(K), x)$ for $(K, x) \in \mathbb{F}_2^k \times \mathbb{F}_2^n$. Suppose that we have an r-round linear trail $[\gamma_s, \gamma_{s+1}, \ldots, \gamma_{s+r}]$ for $\tilde{E}$ such that the correlation $\varepsilon_{F_{i+1}}(\gamma_i, \gamma_{i+1})$ for the $(i+1)$-th round is ϵ_i for each $i \in [s..(s+r-1)]$. It is well-known that the average of the correlations over long keys is $\epsilon = \epsilon_s \cdots \epsilon_{s+r-1}$. That is,

$$\Pr_{\mathbf{rk},x}(\langle \gamma_s, x \rangle \oplus \langle \gamma_{s+r}, \tilde{E}_s^{s+r-1}(\mathbf{rk}, x) \rangle \oplus \bigoplus\nolimits_{i=0}^{r-1} \langle \gamma_{s+i}, rk_{s+i} \rangle = 0) = \frac{1+\epsilon}{2}, \quad (1)$$

where $\mathbf{rk} = rk_0 \| rk_1 \| \cdots \| rk_{R-1}$ and $\tilde{E}_i^j$ is the subcipher of $\tilde{E}$ spanning from the $(i+1)$-th round to the $(j+1)$-th round. (See e.g. [35].) Let K^* be the fixed unknown key to be recovered. Let $\psi(K^*) = \mathbf{rk}^* = rk_0^* \| rk_1^* \| \cdots \| rk_{R-1}^*$. For each key K, let $\Delta K = K \oplus K^*$. Since ψ is linear, $\delta\mathbf{rk} := \psi(K) \oplus \psi(K^*)$ is determined by ΔK. Let $\delta\mathbf{rk} = \delta rk_0 \| \delta rk_1 \| \cdots \| \delta rk_{R-1}$. By (1),

$$\Pr_{\mathbf{rk} \in \mathrm{Im}(\psi),x}(\langle \gamma_s, x \rangle \oplus \langle \gamma_{s+r}, \tilde{E}_s^{s+r-1}(\mathbf{rk}, x) \rangle \oplus \bigoplus\nolimits_{i=0}^{r-1} \langle \gamma_{s+i}, rk_{s+i} \rangle = 0) \approx \frac{1+\epsilon}{2},$$

which means that the correlation of the approximation

$$\langle \gamma_s, x \rangle \oplus \langle \gamma_{s+r}, E_s^{s+r-1}(K^* \oplus \Delta K, x) \rangle \oplus \bigoplus\nolimits_{i=0}^{r-1} \langle \gamma_{s+i}, rk_{s+i}^* \oplus \delta rk_{s+i} \rangle, \quad (2)$$

i.e. the imbalance of the approximation as $(x, \Delta K)$ takes all the values in $\mathbb{F}_2^{n+k}$, is the same regardless of K^* and is very close to ϵ. Since ψ is linear, we have a

linear function L_ψ and a constant C_ψ with $\psi(K) = L_\psi(K) \oplus C_\psi$ for each key K. So for each $i \in [0..(r-1)]$, we have a linear relation

$$\langle \bar{\gamma}_{s+i}, \Delta K \rangle \oplus \langle \gamma_{s+i}, \delta rk_{s+i} \rangle = 0 \tag{3}$$

where $\bar{\gamma}_{s+i}$ is a mask determined from γ_{s+i} and L_ψ. Now, using the approximation (2) and the relations (3), we get a linear approximation

$$\langle \gamma_s, x \rangle \oplus \langle \gamma_{s+r}, E_s^{s+r-1}(K^* \oplus \Delta K, x) \rangle \oplus \bigoplus_{i=0}^{r-1} (\langle \bar{\gamma}_{s+i}, \Delta K \rangle \oplus \langle \gamma_{s+i}, rk^*_{s+i} \rangle) = 0, \tag{4}$$

whose correlation, i.e. the imbalance of the approximation as $(x, \Delta K)$ takes all the values in $\mathbb{F}_2^{n+k}$, is very close to ϵ. We will call each of (2) and (4) a *related-key linear approximation*.

Assumption 1 *The correlation of the related-key linear approximation (2) is ϵ.*

We will also call $\bigoplus_{i=0}^{r-1} \langle \gamma_{s+i}, rk^*_{s+i} \rangle$ the *parity bit* of the approximation. In Matsui's Algorithm 1, the attacker tries to recover only the parity bit and the number of attacked rounds is the same as the number of the rounds that the linear approximation spans over. Matsui's Algorithm 2 tries to add outer rounds to the linear approximation and recover some outer round key bits and, if possible, the parity bits. We will describe the corresponding attacks in our related-key setting.

3.2 Description of Algorithm RKLC-1

In this variant of Matsui's Algorithm 1, we try to recover the parity bit without added outer rounds. So $s = 0$ and $r = R$. Suppose that we have a related-key linear approximation (4) with the correlation ϵ. Let a random related-key data $D = \{(P_i, C_i, \Delta K_i) : i = 1, \ldots, N\}$ for a key K^* be given. Compute

$$\begin{aligned} \tau_0(K^*, D) := & |\{i : \langle \gamma_0, P_i \rangle \oplus \langle \gamma_R, C_i \rangle \oplus \bigoplus_{j=0}^{r-1} \langle \bar{\gamma}_{s+j}, \Delta K_i \rangle = 0\}| \\ & - |\{i : \langle \gamma_0, P_i \rangle \oplus \langle \gamma_R, C_i \rangle \oplus \bigoplus_{j=0}^{r-1} \langle \bar{\gamma}_{s+j}, \Delta K_i \rangle = 1\}|. \end{aligned}$$

If $\epsilon\tau_0(K^*, D) > 0$, then determine the parity bit to be 0 and otherwise determine it to be 1.

3.3 Description of Algorithm RKLC-2

Now we will describe the related-key attack that tries to add outer rounds to an r-round related-key linear approximation (4) and recover some of the outer round key bits together with the parity bit. Assume that we have the approximation (4) with the correlation ϵ. Let a random related-key data $D = \{(P_i, C_i, \Delta K_i) : i = 1, \ldots, N\}$ be given. We let $z_I^* = \bigoplus_{i=0}^{r-1} \langle \gamma_{s+i}, rk^*_{s+i} \rangle$ be the parity bit and z_I be the candidate value for z_I^*. We perform an attack using (4) as the distinguisher: We identify the positions of the outer round key bits that are required to compute

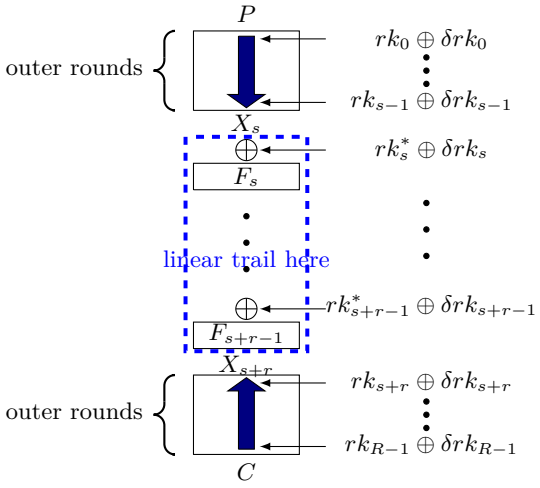

Fig. 2: The outline of RKLC-2

$\langle \gamma_s, X_s^K \rangle \oplus \langle \gamma_{s+r}, X_{s+r}^K \rangle$ with the triplets of plaintext, ciphertext and the key difference where X_s^K and X_{s+r}^K are the intermediate states for the start of the s-th round and the end of the $(s+r-1)$-th round, respectively, that is $X_s^K = E_0^{s-1}(K \oplus \Delta K, P)$ and $E_{s+r}^{R-1}(K \oplus \Delta K, X_{s+r}^K) = C$. (See Fig. 2. Here bits of δrk_i for outer rounds that are used in the computation are computable directly from ΔK.) By allocating a bit value in each of the positions and then concatenating, we get a candidate *outer key* z. Thus $\langle \gamma_s, X_s^K \rangle \oplus \langle \gamma_{s+r}, X_{s+r}^K \rangle \oplus \bigoplus_{i=0}^{r-1} \langle \bar{\gamma}_{s+i}, \Delta K \rangle$ can be expressed as $g(z, P, C, \Delta K)$ for some function g. We denote

$$|\{i : g(z, P_i, C_i, \Delta K_i) = 0\}| - |\{i : g(z, P_i, C_i, \Delta K_i) = 1\}|$$

by $\tau(K^*, D, z)$ and calll $\tau(K^*, D, z)/N$ the observed sample imbalance. Let z^* denote the correct outer key, i.e. the value of z obtained from K^*. Heuristically, if z is correct, then $\tau(K^*, D, z)/N$ is likely to be close to $(-1)^{z_I^*}\epsilon$ and otherwise the imbalance is likely to be close to 0. For the actual attack one usually performs the data compression first to reduce the computational complexity. The data compression in a linear attack is a process that collapses the data into a new data with multiplicity considering the outer round computations. It is part of the distillation [7] and is also called the linear compression by others [32]. But the data compression in our related-key setting needs to handle the key differences unlike that in the single-key linear attacks. So the "compression function" $H_c : \mathbb{F}_2^{2n+k} \to \mathbb{F}_2^d$ with $2^d \ll N$ we need to get for the data compression is one such that the computation of $g(z, P, C, \Delta K)$ can be carried out using z and $H_c(P, C, \Delta K)$ or such that there is a function h such that $g(z, P, C, \Delta K) = h(z, H_c(P, C, \Delta K))$ for any $(P, C, \Delta K)$. Once we have a compression function, we apply it to the data to get the compressed data $\{(v, n_v) \in \mathbb{F}_2^d \times \mathbb{Z} : n_v = |\{(P, C, \Delta K) \in D : H_c(P, C, \Delta K) = v\}|\}$. We determine whether (z, z_I) is correct

Alg. 1 Algorithm RKLC-2

1. Perform the data compression to get the compressed set of size 2^d and set $\tau(K^*, D, z) = 0$ for each z.
2. For each entry (v, n_v) in the compressed data,
   - For each z, compute $h(z, v)$ and increment or decrement $\tau(K^*, D, z)$ by n_v depending on whether $h(z, v)$ is 0 or 1.
3. For each (z, z_I) for which $(-1)^{z_I}\tau(K^*, D, z)\epsilon \geq tN\epsilon^2$, try to recover the whole key bits by trial encryption.

or not by the decision rule $(-1)^{z_I}\tau(K^*, D, z)\epsilon \geq tN\epsilon^2$. Here t is the threshold parameter that enables us to get a tradeoff between the computational complexity and the success probability. To summarize, the attack proceeds as in Alg. 1. If $h(z, v)$ can be expressed as $h'(z \oplus v)$, we can use the FWHT to reduce the computational complexity as described in [18]. Consider the list of (z, z_I)'s for which $(-1)^{z_I}\tau(K^*, D, z)\epsilon \geq tN\epsilon^2$ in the attack. The attack is successful if (z^*, z_I^*) is in the list, and the list may also contain many wrong entries that are called the false alarms.

3.4 Statistical Model, Success Probability and Attack Complexities

For our related-key attacks, we presume the "standard" key hypotheses that are similar to the ones accepted as valid in the ordinary linear attack using a dominant trail. But for that, we assume that the data D is random and that the round function of the cipher in consideration is not too simple. We will see in Sect. 5.2 that when the trail is not dominant and the number of plaintexts per each key difference in the data gets larger, such hypotheses get less pertinent. Under the standard hypotheses, we get the same estimates for the success probability and the attack complexities in terms of the data size and the correlation as in many previous works (e.g. [8, 38]). But we will clarify our hypotheses in the related-key setting and elaborate on the details. We let $c_{N,\epsilon} := \sqrt{N}|\epsilon|$ for each $N > 0$ and ϵ.

Algorithm RKLC-1. Let us consider the attack in Sect. 3.2 that uses a random data D of size N. If we fix K^* and let D vary, $\tau_0(K^*, D)/N$ can be regarded as a random variable. For the attack, we presume the following:

Hypothesis 1 *For each K^*, $\tau_0(K^*, D)/N$ follows $\mathcal{N}((-1)^b\epsilon, 1/N)$ where b is the parity bit.*

With this hypothesis, the success probability of the attack is $\Phi(\sqrt{N}|\epsilon|)$ by Lemma 1.

Lemma 1. *Let $\sigma > 0, b, \mu$ be real numbers and let Y be a random variable with $Y \sim \mathcal{N}(\mu, \sigma^2)$. Then $\Pr(Y \geq b) = \Phi((\mu - b)/\sigma)$ and $\Pr(Y \leq b) = \Phi((b - \mu)/\sigma)$.*

Algorithm RKLC-2. Now we consider the attack in Sect. 3.3. We fix K^* and let z^* be the correct outer key. For each outer key z, we can regard $\tau(K^*, D, z)/N$ as a random variable letting D vary. The right key hypothesis and the wrong key hypothesis we presume are the following:

Hypothesis 2 (Right Key Hypothesis) *For each K^*, $\tau(K^*, D, z^*)/N$ follows $\mathcal{N}((-1)^{b^*}\epsilon, (1-\epsilon^2)/N) \approx \mathcal{N}((-1)^b\epsilon, 1/N)$, where b^* is the parity bit.*

Hypothesis 3 (Wrong Key Hypothesis) *For each K^*, $\tau(K^*, D, z)/N$ follows $\mathcal{N}(0, 1/N)$ when $z \neq z^*$.*

We let $\mathcal{Z}$ be the set of the candidate outer keys and let $k_O = \log_2 |\mathcal{Z}|$. Let t be the threshold parameter. The success probability of the attack is $\Pr((-1)^{b^*}\tau(K^*, D, z^*)\epsilon \geq tN\epsilon^2)$, which equals $\Phi((1-t)c_{N,\epsilon})$ by Lemma 1 under Hypothesis 2. Using the same Lemma, we also have

- for (z, z_I) with $z \neq z^*$, the probability that $(-1)^{z_I}\tau(K^*, D, z)\epsilon \geq tN\epsilon^2$ is $\Phi(-tc_{N,\epsilon})$, and
- for (z, z_I) with $z = z^*$ and $z_I \neq z_I^*$, the probability that $(-1)^{z_I}\tau(K^*, D, z)\epsilon \geq tN\epsilon^2$ is $\Phi((-1-t)c_{N,\epsilon})$.

under Hypothesis 3. So the false alarm probability is

$$\begin{aligned}&\Pr(\mathbf{cond}, (z, z_I) \neq (z^*, z_I^*)) = \Pr(\mathbf{cond}, z \neq z^*) + \Pr(\mathbf{cond}, z = z^*, z_I \neq z_I^*)\\ &= \Pr(\mathbf{cond} \mid z \neq z^*)\Pr(z \neq z^*) + \Pr(\mathbf{cond} \mid z = z^*, z_I \neq z_I^*)\Pr(z = z^*, z_I \neq z_I^*)\\ &= (2^{k_O} - 1)\Phi(-tc_{N,\epsilon})/2^{k_O} + \Phi((-1-t)c_{N,\epsilon})/2^{k_O+1},\end{aligned}$$

where **cond** is short for the statement $(-1)^{z_I}\tau(K^*, D, z)\epsilon \geq tN\epsilon^2$.

Theorem 1. *With N, ϵ, t as described, the success probability of RKLC-2 is $\Phi((1-t)c_{N,\epsilon})$ and the false alarm probability $p_{\mathrm{fa}}(t)$ is $(2^{k_O} - 1)\Phi(-tc_{N,\epsilon})/2^{k_O} + \Phi((-1-t)c_{N,\epsilon})/2^{k_O+1}$.*

Note that $p_{\mathrm{fa}}(t) \approx \Phi(-tc_{N,\epsilon})$ when k_O is not too small. To compare the computational complexity of the attack with that of the exhaustive key search, we say that the complexity of 1 encryption (including the key schedule) is 1. Let c_p be the complexity of 1 computation of H_c and c_o be the complexity of 1 computation of h using an entry in the compressed data and a candidate outer key. Then the computational complexity of RKLC-2 is $c_pN + c_o2^{d+k_O} + 2^k p_{\mathrm{fa}}(t)$ by Theorem 1. The amount of memory required for the attack is $O(2^d)$. Let c_a be the complexity of addition or subtraction of two integers. In many cases, we can reduce the computational complexity by using FWHT:

Theorem 2. *The computational complexity of RKLC-2 using FWHT is*

$$c_pN + 3c_ak_O2^{k_O+1} + 2^k p_{\mathrm{fa}}(t),$$

with the success probability $\Phi((1-t)c_{N,\epsilon})$.

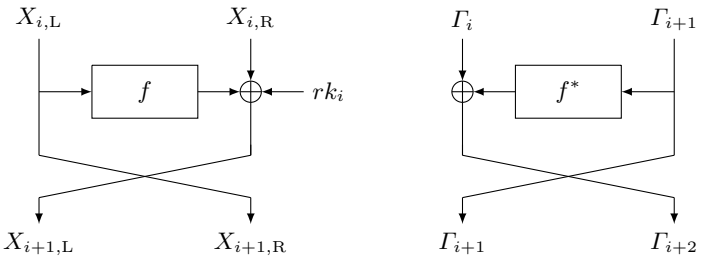

Fig. 3: A round of SIMON and a 1-round linear trail

But in this case, the memory complexity is $O(2^{k_O})$. Here we have assumed that the restored outer key z and the parity bit reveal simple independent relations between the bits of K^*, meaning that using the $(k_O + 1)$-bit information that (z^*, z_I^*) reveals about K^*, we can recover the whole k bits of K^* using other simple $(k - k_O - 1)$ relations between the bits of K^*. This is mostly the case when the key schedule is linear.

4 Related-Key Linear Attacks on Round-reduced Simon

The NSA published two families of lightweight block ciphers SIMON and SPECK [4]. They have remarkable performance figures on most software and hardware platforms and SIMON is the more hardware-oriented of the two. They have been the subject of intensive security analysis since their publication. The designers of SIMON expect that it is secure against related-key attacks [5].

4.1 The Simon Family of Block Ciphers

SIMONn/k is a block cipher of the classical Feistel structure with k-bit keys and n-bit blocks. Its round function f sends an $n/2$-bit input x onto $((x \lll 8) \wedge (x \lll 1)) \oplus (x \lll 2)$.(See Fig. 3.) It has a linear key schedule. We focus on the following ciphers whose key lengths are double the block lengths: SIMON32/64, SIMON48/96, SIMON64/128, and SIMON128/256. The details of this section can be applied equally well to the variant of SIMON to be used in Sect. 5.

4.2 Related-Key Linear Approximations of Simon

Since SIMONn/k has the classical Feistel structure, an r-round linear trail can be represented as a sequence of $(r+2)$ $n/2$-bit masks: $\Gamma_s.\Gamma_{s+1}.\cdots.\Gamma_{s+r+1}$ represents a linear trail such that at the $(i+1)$-th round, the input and output masks are $\Gamma_i \| \Gamma_{i+1}$ and $\Gamma_{i+1} \| \Gamma_{i+2}$, respectively, for each $i \in [s..(s+r-1)]$. (See Fig. 3.) Such a linear trail leads to the related-key linear approximation

$$\langle \Gamma_s, X_{s,\mathrm{L}} \rangle \oplus \langle \Gamma_{s+1}, X_{s,\mathrm{R}} \rangle \oplus \langle \Gamma_{s+r}, X_{s+r,\mathrm{L}} \rangle \oplus \langle \Gamma_{s+r+1}, X_{s+r,\mathrm{R}} \rangle \\ \oplus \langle \Lambda, \Delta K \rangle \oplus \bigoplus_{i=0}^{r-1} \langle \Gamma_{s+i+1}, rk^*_{s+i} \rangle = 0 \qquad (5)$$

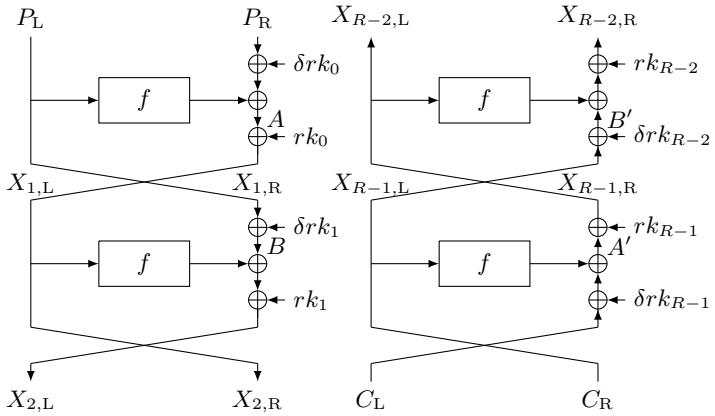

Fig. 4: 2-round computations of SIMON

between X_s, ΔK, X_{s+r}, where Λ is a mask with $\langle \Lambda, \Delta K \rangle = \bigoplus_{i=0}^{r-1} \langle \Gamma_{s+i+1}, \delta rk_{s+i} \rangle$. Such a mask Λ can be easily obtained.

4.3 Adding Outer Rounds

One of the pivotal processes of the related-key attacks is to get effective data compression with the related-key linear approximation. For SIMON, we can get effective data compression for prepending and appending many rounds when both the initial mask and the final mask of the linear trail have small Hamming weights. In this subsection, we will explain in detail how to add 2+2 rounds, i.e., how to prepend 2 rounds and append 2 rounds at the same time. How to add $s+s$ rounds for $s = 3, 4, 5$ will be explained in Sect. C of the Appendix, from which how to add $s + s'$ rounds for $s \neq s'$ and $2 \leq s, s' \leq 5$ will be obvious. For simplicity $a + b$ and ab denote the XOR and AND of $a, b \in \mathbb{F}_2$ in this section, respectively.

2-round computation. Let rk_0 and rk_1 be the round keys derived from the candidate key K for the first 2 rounds. For a plaintext $P = P_{\mathrm{L}} \| P_{\mathrm{R}}$ and a key difference ΔK, let δrk_0 and δrk_1 be the derived round key differences for the first 2 rounds. Note that δrk_0 and δrk_1 can be computed directly from ΔK. We want to express each bit of $X_2 = X_{2,\mathrm{L}} \| X_{2,\mathrm{R}} = E_0^1(K \oplus \Delta K, P)$ in terms of P, rk_0, rk_1, δrk_0, and δrk_1. Let $A = f(P_{\mathrm{L}}) \oplus P_{\mathrm{R}} \oplus \delta rk_0$, and $B = P_{\mathrm{L}} \oplus \delta rk_1$. (See Fig. 4.) Since $X_{2,\mathrm{R}} = X_{1,\mathrm{L}} = f(P_{\mathrm{L}}) \oplus P_{\mathrm{R}} \oplus \delta rk_0 \oplus rk_0 = rk_0 \oplus A$ and $X_{2,\mathrm{L}} = f(X_{1,\mathrm{L}}) \oplus P_{\mathrm{L}} \oplus \delta rk_1 \oplus rk_1 = f(rk_0 \oplus A) \oplus B \oplus rk_1$,

$$\begin{aligned} X_{2,\mathrm{L}}[i] &= (rk_0[i-1] + A[i-1])(rk_0[i-8] + A[i-8]) \\ &\quad + \underline{rk_0[i-2]} + A[i-2] + B[i] + \underline{rk_1[i]}, \\ X_{2,\mathrm{R}}[i] &= \underline{rk_0[i]} + A[i] \end{aligned}$$

Here $X_{2,\mathrm{L}}[i]$ can be computed in terms of $rk_0[i-1]+A[i-1], rk_0[i-8]+A[i-8]$, up to $A[i-2]+B[i]$ xored with a constant determined only by rk_0, rk_1. Note that the underlined terms do not mingle with the plaintext so that we will xor them with the parity bit to get an "adjusted parity bit" in RKLC-2. Otherwise the number of round key bits to restore and, hence, the attack complexity can be increased. By symmetry of the cipher structure, we get similar expressions for bits of $X_{R-2,\mathrm{R}}$ and $X_{R-2,\mathrm{L}}$ in terms of $A' = f(C_\mathrm{R}) \oplus C_\mathrm{L} \oplus \delta rk_{R-1}$, $B' = C_\mathrm{R} \oplus \delta rk_{R-2}$, rk_{R-2}, and rk_{R-1}. (See Fig. 4.)

The data compression. The above arguments tell us how to compress the data when adding 2+2 rounds. Suppose that we want to make use of the related-key linear approximation represented as (5) with $s = 2$ and $s + r + 2 = R$. Let $w = n/2$ be the word size. Let $\mathcal{I}_\mathrm{L} = \mathsf{supp}(\Gamma_s) = \{i \in [0..(w-1)] : \Gamma_s[i] \neq 0\}$, $\mathcal{I}_\mathrm{R} = \mathsf{supp}(\Gamma_{s+1})$, $\mathcal{I}'_\mathrm{L} = \mathsf{supp}(\Gamma_{R-2})$, and $\mathcal{I}'_\mathrm{R} = \mathsf{supp}(\Gamma_{R-1})$. The compression function extracts the following values from each data entry $(P, C, \Delta K)$:

- $A[i]$ for i such that $(i+1) \bmod w \in \mathcal{I}_\mathrm{L}$ or $(i+8) \bmod w \in \mathcal{I}_\mathrm{L}$
- $A'[i]$ for i such that $(i+1) \bmod w \in \mathcal{I}'_\mathrm{R}$ or $(i+8) \bmod w \in \mathcal{I}'_\mathrm{R}$
- $\bigoplus_{i\in\mathcal{I}_\mathrm{L}} (A[i-2] \oplus B[i]) \oplus \bigoplus_{i\in\mathcal{I}_\mathrm{R}} A[i] \oplus \bigoplus_{i\in\mathcal{I}'_\mathrm{R}} (A'[i-2] \oplus B'[i]) \oplus \bigoplus_{i\in\mathcal{I}'_\mathrm{L}} A'[i] \oplus \langle \Lambda, \Delta K \rangle$

The outer keys consists of the following outer round key bits that we need to guess:

- $rk_0[i]$ for i such that $(i+1) \bmod w \in \mathcal{I}_\mathrm{L}$ or $(i+8) \bmod w \in \mathcal{I}_\mathrm{L}$
- $rk_{R-1}[i]$ for i such that $(i+1) \bmod w \in \mathcal{I}'_\mathrm{R}$ or $(i+8) \bmod w \in \mathcal{I}'_\mathrm{R}$

The adjusted parity bit is $\bigoplus_{i=0}^{r-1} \langle \Gamma_{s+i+1}, rk^*_{s+i} \rangle \oplus \bigoplus_{i\in\mathcal{I}_\mathrm{L}} (rk^*_0[i-2] \oplus rk^*_1[i]) \oplus \bigoplus_{i\in\mathcal{I}_\mathrm{R}} rk^*_0[i] \oplus \bigoplus_{i\in\mathcal{I}'_\mathrm{R}} (rk^*_{R-1}[i-2] \oplus rk^*_{R-2}[i]) \oplus \bigoplus_{i\in\mathcal{I}'_\mathrm{L}} rk^*_{R-1}[i]$. Note that the number k_O of guessed round key bits for outer rounds is at most $2\mathrm{wt}(\Gamma_s) + 2\mathrm{wt}(\Gamma_{R-1})$ and d, $\log_2$ of the size of the compressed data, is $k_O + 1$. Note also that using above data compression, we can use FWHT in RKLC-2.

4.4 Attacks on Round-reduced Simon

Now we will present the attacks on round-reduced Simon summarized in Table 1. Note that we compare our attacks with the current best linear attacks [16] and differential attacks [41] only since the differential/linear attacks are the most limiting attacks on Simon as noted in [5]. Note also that there does not exist a related-key attack that is more efficient than such attacks yet (cf. [30]).
Each of our attacks is an instance of the RKLC-2 and we will specify the positions of the outer round key bits that will constitute the outer keys. We use the linear trails presented in Sect. B of the Appendix. Note that each of them has squared correlation somewhat larger than $2^{-(n+k)/2}$ and the product of the computational complexity and the data complexity of the presented attack using it is less than

2^{k+n}. So each presented attack is an effective one not covered by the generic attack in [28]. We set the threshold parameter t for each attack to 1 so that the success probabilities are all $\Phi(0) = 0.5$ by Theorem 1. Also we have $k_O \ll \log_2(N)$ and use FWHT in each attack so that we estimate the computational complexity as $c_p N + 2^k \Phi(-\sqrt{N}|\epsilon|)$ by Theorem 2. We use the estimate $c_p = R_{\text{add}}/R$, where R is the number of rounds of the round-reduced SIMON and R_{add} is the number of the added outer rounds. The attack method presented in this work can be applied to other SIMONn/k with $k > n$ in a straightforward manner.

Simon32/64. We use a 16-round linear trail with the correlation 2^{-21} whose initial and final mask are `40000001` and `00400110`, respectively. We can add 4+3 rounds to this linear trail with $k_O = 35$: The guessed outer round key bits are $rk_0[0, 2, 3, 4, 6, 7, 9, 10\ 11, 12, 13, 14]$, $rk_1[4, 5, 6, 11, 12, 13, 14]$, $rk_2[6, 13]$, $rk_{21}[0, 3, 7, 12]$, and rk_{22} [1,2,4,5,6,8,10,11,14,15]. Letting $N = 2^{46.3}$, $c_{N,\epsilon} = 2^{2.15}$ so that the complexity of RKLC-2 on the 23-round reduced SIMON32/64 is $2^{46.65}$ by Theorem 2.

Simon48/96. We use a 20-round linear trail with the correlation 2^{-33} whose initial and final mask are `400000000001` and `400000100001`, respectively. We add 5+3 rounds to this linear trail with $k_O = 53$. Letting $N = 2^{70.9}$, $c_{N,\epsilon} = 2^{2.45}$ and the complexity of RKLC-2 on the 28-round reduced SIMON48/96 is $2^{71.07}$.

Simon64/128. We use a 26-round linear trail with the correlation 2^{-45} whose initial mask is `0000000100004044` and final mask is `0000100000004400`. We add 4+4 rounds to this linear trail with $k_O < 80$. Letting $N = 2^{95.32}$, $c_{N,\epsilon} = 2^{2.66}$ and the complexity of RKLC-2 on the 34-round reduced SIMON64/128 is $2^{95.5}$.

Simon128/256. We use a 51-round linear trail with the correlation 2^{-92} whose initial and final mask are `00...004‖00...00` and `00...001‖400...004`, respectively. We get an attack on the 62-round reduced SIMON128/256 with data complexity $2^{190.4}$ and computational complexity $2^{190.76}$ by adding 6+5 rounds. Also, using a 45-round subtrail with the correlation 2^{-84} whose initial and final mask are `100..` `.001‖4400...004` and `400...004‖100...00`, respectively, we get an attack on the 55-round reduced SIMON128/256 with data complexity $2^{174.73}$ and computational complexity 2^{175} by adding 5+5 rounds.

5 Experiments

In this section, we carry out experiments using three block ciphers. Two of them are small-scale variants of SIMON. The other is PRESENT-L that is a variant of PRESENT-128 that has a linear key schedule.

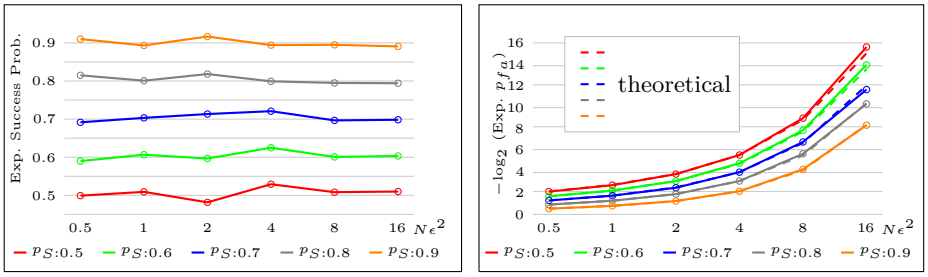

Fig. 5: Experimental results for 22-round key recovery on SIMON24

5.1 Experiments with Variants of Simon

Related-key attacks on the 22-round Simon24. We describe experimental results on SIMON24/48 that is a 22-round cipher with 48-bit keys and 24-bit blocks. The round function and the key schedule of the cipher are defined exactly in the same way as SIMON32/64. The 31-bit constant used in the key schedule is also the same. We try to add 2+2 rounds to the 18-round linear trail 0 01.000.001.410.001.000.001.410.001.000.001.410.001.000.001.410.001.000.0 01.400 with the correlation $\epsilon = 2^{-17}$. The guessed outer round key bits are $rk_0[4, 11], rk_{21}[2, 9]$, the number k_O of the guessed outer round key bits is 4, and the size of the compressed data is 2^5 by arguments in Sect. 4. The additional bit to be guessed is the adjusted parity bit $rk_0[10] \oplus rk_1[0] \oplus rk_{20}[10] \oplus rk_{21}[0] \oplus rk_{21}[8] \oplus \langle 000, rk_2 \rangle \oplus \langle 001, rk_3 \rangle \oplus \cdots \oplus \langle 001, rk_{19} \rangle$. In the experiment we repeat the key recovery tests using 1,000 different keys K^*. For each key, we generate data of size N for $N = 2^i \epsilon^{-2}$ with $i = -1, 0, 1, 2, 3$, and 4. The number ν of data entries per key difference was fixed to as large as 2^{16}. For each N, we compute the threshold parameters corresponding to $p_S = 0.5, 0.6, 0.7, 0.8$, and 0.9 using Theorem 1 and proceed as in Alg. 1. We count the number of the successful attempts and measure the average of the number of false alarms for the 1,000 tests. The result is shown in Fig. 5 from which we can see that the experimental probabilities are close to the theoretical ones.

Related-key attacks on the 16-round Simon32. We try to add 2+2 round to the 12-round trail 0005.0000.0005.c001.1005.0110.0040.0100.0000.0100. 0040.0110.0004.0111 with the correlation $\epsilon = -2^{-17}$. The number k_O of guessed outer round key bits is 10 and we proceed as in the preceding section. Then we get the results as in Fig. 6.

Single-key attacks on the 18-round Simon24. For comparison with the related-key linear attacks, we also perform a single-key linear attack using the linear hull containing the 14-round trail with correlation 2^{-13} represented as 001.000.001.410.001.000.001.410.001.000.001.410.001.000.001.400. We try

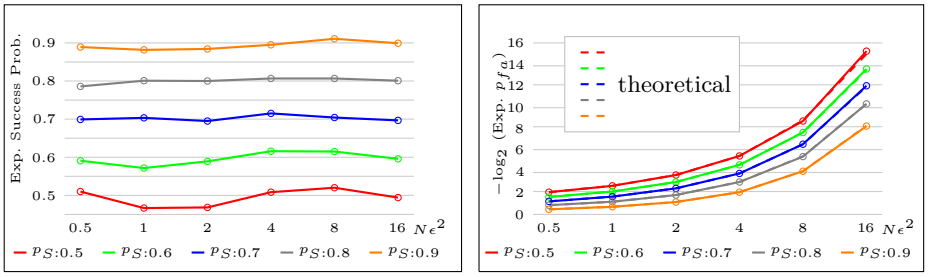

Fig. 6: Experimental results for 16-round key recovery on SIMON32

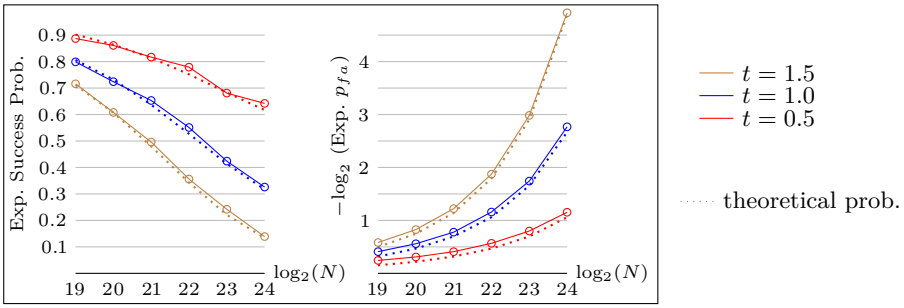

Fig. 7: Experimental results for 18-round single-key attack on SIMON24

to add 2+2 rounds to the linear hull. By analyzing the distribution of the squared correlation of the linear hull over 1,000 keys, we noticed that the linear probability of the linear hull is about 2^{-23}. We also observed that the distribution of the correlation of the linear hull is close to $\mathcal{N}(0, 2^{-23})$ as predicted by arguments in [19]. So we assume that the linear probability ϵ_H^2 of the linear hull is 2^{-23} and the distribution of the correlation of the linear hull is $\mathcal{N}(0, \epsilon_H^2)$. Then we apply Matsui's Algorithm 2 presuming some adjusted key hypotheses for attacks using data sampled without replacement [8]: For the attack we let the data size N be $2^{19}, 2^{20}, \ldots, 2^{23}$, or 2^{24}. We use the decision rule $|\tau(K^*, D, z)/N| \geq t|\epsilon_H|$ with threshold parameters $t = 0.5, 1.0, 1.5$. The theoretical success probability and the false alarm probability for each (t, N) are $2\Phi(-t\sqrt{N}|\epsilon_H|/\sqrt{1 - N/2^n + N\epsilon_H^2})$ and $2\Phi(-t\sqrt{N}|\epsilon_H|)$, respectively [8, 9]. We observe that the experimental probabilities are close to the theoretical ones as shown in Fig. 7, confirming the analyses in [8] with linear hulls. Considering the success probabilities for each fixed (a, N), where a is the advantage $-\log_2(p_{\mathrm{fa}}(t))$ and N is the data size, the single-key linear attack using the linear hull with the linear probability 2^{-23} is not so advantageous compared with the related-key linear attack (RKLC-2) using a linear trail with the linear correlation 2^{-13} as we see in Fig. 8.

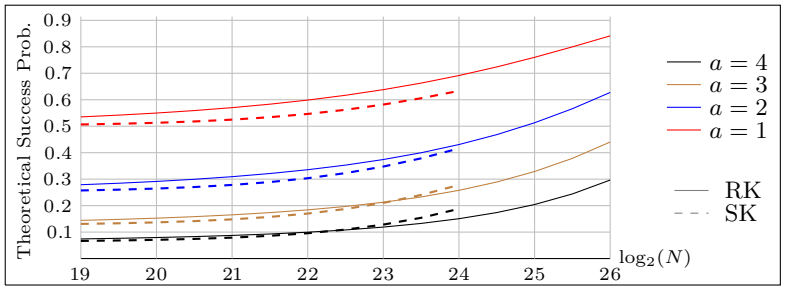

Fig. 8: Success probabilities for the single-key/related-key attack

5.2 Experiments with a Variant of Present-128

Let PRESENT-L be a block cipher that originates from the same long key cipher as PRESENT-128 and has a linear key schedule: The key schedule of PRESENT-L is the same as that of PRESENT-128 except that all the 4-bit S-boxes in the key schedule of PRESENT-128 are removed. So each key schedule round of PRESENT-L is just a rotation of the 128-bit state followed by xoring with a round constant. Let T-4R-B21 and T-4R-B42 be the 4-round trails with correlations 2^{-8} such that the input and output mask for each round is 0000000000200000 and 0000040000000000, respectively. Let T-6R-B21 and T-6R-B42 be the 6-round trails with correlations 2^{-12} defined similarly. Considering linear trails such that all the constituent masks have Hamming weight 1, we see that T-4R-B21 and T-4R-B42 have 2 and 1 other trails with the correlation $\pm 2^{-8}$ in their linear hulls, respectively (cf. [36]). We also see that T-6R-B21 and T-6R-B42 have at least 26 and 7 other trails with the correlation $\pm 2^{-12}$ in their linear hulls, respectively. In the experiments, we set the data size N to be ϵ^{-2}, $4\epsilon^{-2}$, or $16\epsilon^{-2}$. We also set the number ν of data entries per key difference to be 1, 8, or 64. We try to prepend 2 rounds before the 4-round trails or the 6-round trails. The results are as in Fig. 9–Fig. 10. When ν is large, the experimental probabilities may deviate considerably from the estimates given by Theorem 1. But when ν is close to 1, as in the case of random sampling, they are close to the theoretically predicted ones. Rather surprisingly, results with 6-round trails are closer to predicted ones than with the 4-round ones, though the former are far less dominant in their linear hulls than the latter. We suspect that this has been caused by the nonrandomness of the data we have used. The details of the data are provided in the Appendix.

6 Discussions

6.1 Statistical Models

The standard key hypotheses are not adequate for most single-key linear attacks [8, 9,11,12]. But we claim that the standard key hypotheses we presume are adequate for our attacks in the related-key scenario where the attacker has random related-key data though the results in Sect. 5.1 shows that sometimes such key hypotheses

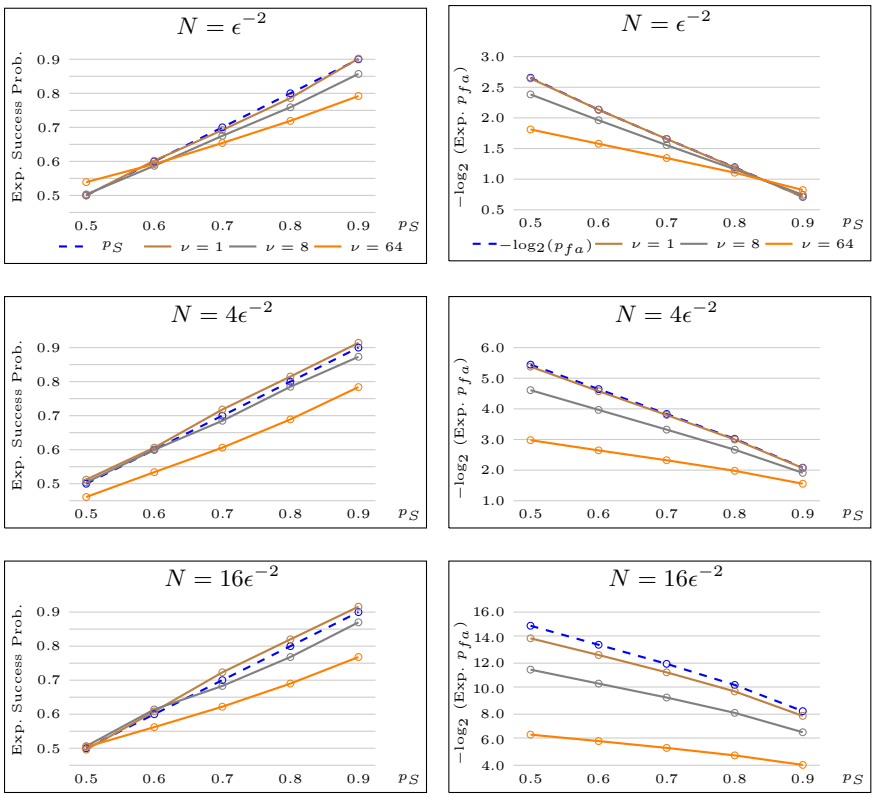

Fig. 9: Experimental results for 6-round key recovery using T-4R-B21

are suitable even when the trail is far from dominant and the number of data entries per key difference is quite large. First, the standard *right* key hypothesis is applicable in our related-key setting since the correlation of the related-key linear approximation is the same regardless of K^* as mentioned in Sect. 3 while that of the single-key linear approximation obtained from a linear trail varies greatly depending on the key if the trail is not dominant. We do not claim that the standard wrong key hypotheses are adequate in the related-key attack regardless of the round structure of the block cipher. But we claim that such hypotheses are appropriate in the related-key scenario with random data if the round function is not too simple. In the extreme nontypical case when we have only one related key, our attack would become one in the single-key setting and we might have to consider some adjusted wrong key hypotheses especially when the data size is large. The reason is that the distribution of the correlations for wrong keys in the single-key setting has deviation $O(2^{-n/2})$ [11] from J. Daemen and V. Rijmen's analysis [19] on the distribution of the correlations of linear approximations for n-bit permutations. But in our related-key setting where the

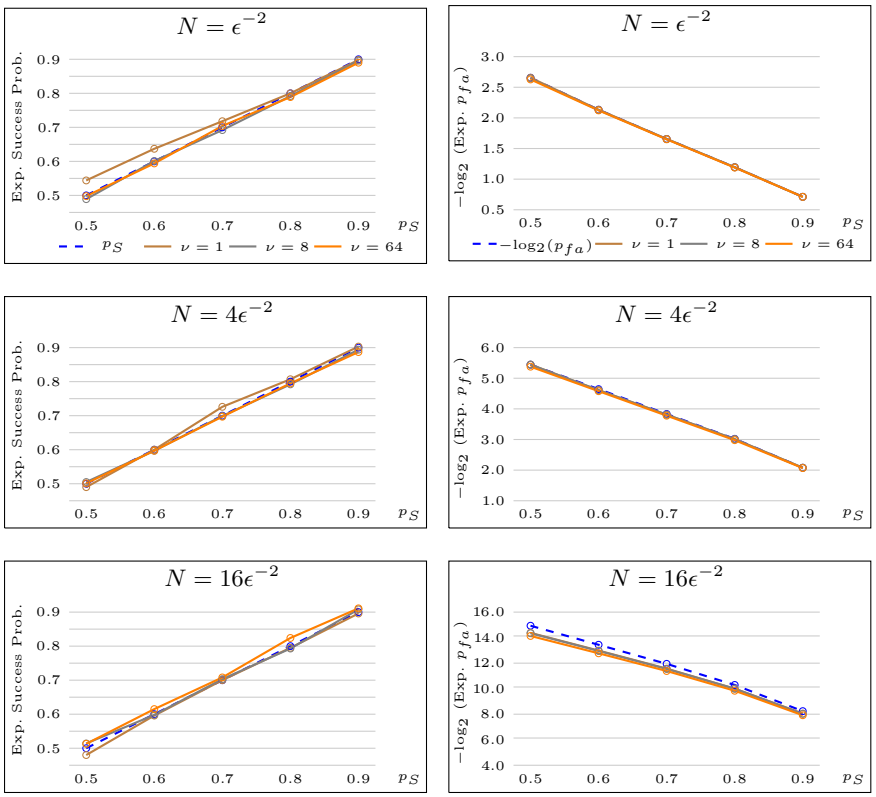

Fig. 10: Experimental results for 6-round key recovery using T-6R-B42

data is random, we just need to consider $(k+n)$-bit-to-n-bit functions where the additional k bits come from key differences so the deviation of the distribution might be $O(2^{-(k+n)/2})$ that is negligible compared to $1/\sqrt{N}$ with the data size N being much smaller than 2^{k+n}. The standard wrong key hypothesis in Matsui's Algorithm 2 in the single-key linear attack is not so valid especially when the linear probability of the linear hull is close to 2^{-n}. But the experimental results in Sect. 5.1 using a linear trail with the correlation whose absolute value is considerably less than $2^{-n/2}$ corroborate our claims on the wrong key hypotheses as well as the right key hypotheses in our related-key attacks. We also note that the choice between sampling data with replacement and without replacement yields little difference in the success probability and the attack complexities due to the size of the whole data space.

6.2 Linear Trails, Linear Hulls, and Multiple Linear Approximations

The formulations in Sect. 3 indicate that in our related-key linear attacks it seems natural to use linear trails. The experimental results in Sect. 5 show that

our estimates regarding the related-key linear attacks are accurate regardless of whether the trail is dominant or not when random data is used. In fact, none of the linear trail used in the experiments in Sect. 5 are dominant. For example, the linear correlation of the 18-round linear trail for SIMON24 is 2^{-17} whose square is quite less than the linear probability of the linear hull containing the trail that is larger than 2^{-30}. But the use of the key differences in our attacks seems to make it hard or inherently impossible to utilize the linear hulls. On the other hand, we would get related-key multiple linear attacks and related-key multidimensional attacks as straightforward derivations of the ones presented in [7, 23]. We expect that the key dependency issue in the original attacks will be mitigated in our related-key setting.

6.3 Comparison with Single-Key Linear Attacks

Our attack has significance even when it requires slightly more computation and data than the single-key linear attacks since it can utilize related-key data that the single-key attacks cannot exploit. But it is more meaningful when it is more effective than the generic related-key attack and covers more rounds than the single-key attacks. For a block cipher with a linear key schedule whose key length k is larger than the block length n, we need to find a linear trail with the correlation $\approx \pm 2^{-(k+n)/4}$ considering the generic related-key attack. The attack will be likely to cover longer rounds than the single-key linear attack only when we can find such a trail that is longer than any single-key linear approximation with the linear probability $\approx 2^{-n}$. The advantage of the related-key linear attack will be more visible for block ciphers with linear key schedules such that the linear hulls exploited in the single-key linear attacks contain dominant trails. For each of SIMON ciphers, we could not cover much longer rounds with the related-key attacks than with the single-key attacks mainly because it admits considerably longer single-key linear characteristics with the linear probability $\approx 2^{-n}$ originating from linear hulls than single-key linear trails with the correlation $\approx \pm 2^{-n/2}$. For example, the linear attack on the 25-round reduced SIMON48/k in [16] exploits a 16-round linear hull with linear probability $2^{-42.92}$, while there does not exist a 16-round linear trail whose squared correlation is larger than 2^{-50}.

6.4 Application to Tweakable Blockciphers

The statistical model in this work can be slightly modified to provide a relevant framework for the linear attack on the tweakable blockciphers constructed from the tweakey framework [6, 26]. Though the linear attack using a linear approximation with squared correlation much smaller than 2^{-n} against such a cipher is rather straightforward and was considered in [6, 31], it requires suitable statistical models. It seems that Assumption 1 can be adjusted to provide an adequate right key hypothesis for the attack. For example, it may be assumed that the average of the correlations over (K^*, T) as the tweak T varies are very close to the correlation of the linear trail regardless of the base key K^*. Note that the correlations of the related-key linear approximations are the same regardless of K^*, and we just

assume that they are the same as the correlation of the linear trail. A wrong key hypothesis that is similar to the one given in this work might be adopted for the attack.

7 Conclusions

We have introduced a general framework for the related-key linear cryptanalysis on block ciphers with linear key schedules. The attack is likely to cover more rounds than the existing linear attacks if the key length of the cipher is much larger than its block length. Using the framework, we are able to get effective related-key linear attacks on Simon that cover longer rounds than the previous attacks. Experiments with small-scale variants of Simon and a variant of Present corroborate the validity of the attack together with the suitability of the statistical model concerning the right keys and wrong keys. Some lightweight ciphers do not consider security against related-key attacks in its design criteria and the attack may be hard to apply in practice due to the requirement of large related-key data. But the attack is certainly one that should be taken into account in the design of a block cipher with a small block length and a linear key schedule that might be used in circumstances where frequent key update is inevitable. An additional feature of our attack is that the complexity of the attack does not depend much on the detail of the key schedule once the key schedule is linear, in contrast with other related-key attacks. As further works, one may try to apply the framework presented in this work to various block ciphers with linear key schedules other than Simon. Another important line of works would be to investigate the multiple linear attacks and multidimensional linear attacks in the related key setting to reduce the attack complexities.

Acknowledgements. We are grateful to the anonymous reviewers for their help in improving the quality of the paper. This work was supported by Institute for Information & Communications Technology Planning & Evaluation(IITP) grant funded by the Korean government(MSIT) (No.2017-0-00267).

References

1. Ashur, T., Bodden, D., Dunkelman, O.: Linear cryptanalysis using low-bias linear approximations. Cryptology ePrint Archive, Report 2017/204 (2017), http://eprint.iacr.org/2017/204
2. Banik, S., Bogdanov, A., Isobe, T., Shibutani, K., Hiwatari, H., Akishita, T., Regazzoni, F.: Midori: A block cipher for low energy. In: Iwata, T., Cheon, J.H. (eds.) ASIACRYPT 2015, Part II. LNCS, vol. 9453, pp. 411–436. Springer (2015)
3. Banik, S., Pandey, S.K., Peyrin, T., Sasaki, Y., Sim, S.M., Todo, Y.: GIFT: A small present - towards reaching the limit of lightweight encryption. In: Fischer, W., Homma, N. (eds.) CHES 2017. LNCS, vol. 10529, pp. 321–345. Springer (2017)
4. Beaulieu, R., Shors, D., Smith, J., Treatman-Clark, S., Weeks, B., Wingers, L.: The simon and speck families of lightweight block ciphers. Cryptology ePrint Archive, Report 2013/404 (2013), http://eprint.iacr.org/2013/404

5. Beaulieu, R., Shors, D., Smith, J., Treatman-Clark, S., Weeks, B., Wingers, L.: Notes on the design and analysis of simon and speck. Cryptology ePrint Archive, Report 2017/560 (2017)
6. Beierle, C., Jean, J., Kölbl, S., Leander, G., Moradi, A., Peyrin, T., Sasaki, Y., Sasdrich, P., Sim, S.M.: The SKINNY family of block ciphers and its low-latency variant MANTIS. In: Robshaw, M., Katz, J. (eds.) CRYPTO 2016, Part II. LNCS, vol. 9815, pp. 123–153. Springer (2016)
7. Biryukov, A., Cannière, C.D., Quisquater, M.: On multiple linear approximations. In: Franklin, M.K. (ed.) CRYPTO 2004. LNCS, vol. 3152, pp. 1–22. Springer (2004)
8. Blondeau, C., Nyberg, K.: Improved parameter estimates for correlation and capacity deviates in linear cryptanalysis. IACR Trans. Symmetric Cryptol. 2016(2), 162–191 (2016)
9. Blondeau, C., Nyberg, K.: Joint data and key distribution of simple, multiple, and multidimensional linear cryptanalysis test statistic and its impact to data complexity. Des. Codes Cryptography 82(1-2), 319–349 (2017)
10. Bogdanov, A., Boura, C., Rijmen, V., Wang, M., Wen, L., Zhao, J.: Key difference invariant bias in block ciphers. In: Sako, K., Sarkar, P. (eds.) ASIACRYPT 2013, Part I. LNCS, vol. 8269, pp. 357–376. Springer (2013)
11. Bogdanov, A., Tischhauser, E.: On the wrong key randomisation and key equivalence hypotheses in matsui's algorithm 2. In: Moriai, S. (ed.) FSE 2013, Revised Selected Papers. LNCS, vol. 8424, pp. 19–38. Springer (2013)
12. Bogdanov, A., Vejre, P.S.: Linear cryptanalysis of DES with asymmetries. In: Takagi, T., Peyrin, T. (eds.) ASIACRYPT 2017, Part I. LNCS, vol. 10624, pp. 187–216. Springer (2017)
13. Bogdanov, A., Wang, M.: Zero correlation linear cryptanalysis with reduced data complexity. In: Canteaut, A. (ed.) FSE 2012, Revised Selected Papers. LNCS, vol. 7549, pp. 29–48. Springer (2012)
14. Borghoff, J., Canteaut, A., Güneysu, T., Kavun, E.B., Knezevic, M., Knudsen, L.R., Leander, G., Nikov, V., Paar, C., Rechberger, C., Rombouts, P., Thomsen, S.S., Yalçin, T.: PRINCE - A low-latency block cipher for pervasive computing applications - extended abstract. In: Wang, X., Sako, K. (eds.) ASIACRYPT 2012. LNCS, vol. 7658, pp. 208–225. Springer (2012)
15. Cannière, C.D., Dunkelman, O., Knezevic, M.: KATAN and KTANTAN - A family of small and efficient hardware-oriented block ciphers. In: Clavier, C., Gaj, K. (eds.) CHES 2009. LNCS, vol. 5747, pp. 272–288. Springer (2009)
16. Chen, H., Wang, X.: Improved linear hull attack on round-reduced simon with dynamic key-guessing techniques. In: Peyrin, T. (ed.) FSE 2016, Revised Selected Papers. LNCS, vol. 9783, pp. 428–449. Springer (2016)
17. Cho, J.Y., Hermelin, M., Nyberg, K.: A new technique for multidimensional linear cryptanalysis with applications on reduced round serpent. In: Lee, P.J., Cheon, J.H. (eds.) ICISC 2008, Revised Selected Papers. Lecture Notes in Computer Science, vol. 5461, pp. 383–398. Springer (2008)
18. Collard, B., Standaert, F., Quisquater, J.: Improving the time complexity of matsui's linear cryptanalysis. In: Nam, K., Rhee, G. (eds.) ICISC 2007. LNCS, vol. 4817, pp. 77–88. Springer (2007)
19. Daemen, J., Rijmen, V.: Probability distributions of correlation and differentials in block ciphers. Cryptology ePrint Archive, Report 2005/212 (2005), http://eprint.iacr.org/2005/212
20. Gennaro, R., Robshaw, M. (eds.): CRYPTO 2015, Part I, LNCS, vol. 9215. Springer (2015)

21. Gérard, B., Grosso, V., Naya-Plasencia, M., Standaert, F.: Block ciphers that are easier to mask: How far can we go? In: Bertoni, G., Coron, J. (eds.) CHES 2013. LNCS, vol. 8086, pp. 383–399. Springer (2013)
22. Guo, J., Peyrin, T., Poschmann, A., Robshaw, M.J.B.: The LED block cipher. In: Preneel and Takagi [37], pp. 326–341
23. Hermelin, M., Cho, J.Y., Nyberg, K.: Multidimensional linear cryptanalysis. J. Cryptology 32(1), 1–34 (2019)
24. Hermelin, M., Nyberg, K.: Linear cryptanalysis using multiple linear approximations. IACR Cryptology ePrint Archive 2011, 093 (2011), http://eprint.iacr.org/2011/093
25. Huang, J., Vaudenay, S., Lai, X., Nyberg, K.: Capacity and data complexity in multidimensional linear attack. In: Gennaro and Robshaw [20], pp. 141–160
26. Jean, J., Nikolic, I., Peyrin, T.: Tweaks and keys for block ciphers: The TWEAKEY framework. In: Sarkar, P., Iwata, T. (eds.) 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. Lecture Notes in Computer Science, vol. 8874, pp. 274–288. Springer (2014)
27. Kaliski, B.S., Robshaw, M.J.B.: Linear cryptanalysis using multiple approximations. In: Desmedt, Y. (ed.) CRYPTO '94. LNCS, vol. 839, pp. 26–39. Springer (1994)
28. Kelsey, J., Schneier, B., Wagner, D.A.: Key-schedule cryptanalysis of idea, g-des,. gost, safer, and triple-des. In: Koblitz, N. (ed.) CRYPTO '96. LNCS, vol. 1109, pp. 237–251. Springer (1996)
29. Kölbl, S., Leander, G., Tiessen, T.: Observations on the SIMON block cipher family. In: Gennaro and Robshaw [20], pp. 161–185
30. Kondo, K., Sasaki, Y., Todo, Y., Iwata, T.: Analyzing key schedule of simon: Iterative key differences and application to related-key impossible differentials. In: Obana, S., Chida, K. (eds.) IWSEC 2017. LNCS, vol. 10418, pp. 141–158. Springer (2017)
31. Kranz, T., Leander, G., Wiemer, F.: Linear cryptanalysis: Key schedules and tweakable block ciphers. IACR Trans. Symmetric Cryptol. 2017(1), 474–505 (2017)
32. Liu, Z., Li, Y., Wang, M.: The security of simon-like ciphers against linear cryptanalysis. Cryptology ePrint Archive, Report 2017/576 (2017), http://eprint.iacr.org/2017/576
33. Matsui, M.: Linear cryptanalysis method for DES cipher. In: Helleseth, T. (ed.) EUROCRYPT '93. LNCS, vol. 765, pp. 386–397. Springer (1993)
34. Nyberg, K.: Linear approximation of block ciphers. In: Santis, A.D. (ed.) EUROCRYPT '94. LNCS, vol. 950, pp. 439–444. Springer (1994)
35. Nyberg, K.: Linear cryptanalysis. SAC Summer School (2015), http://sacworkshop.org/SAC2015/S3-linear-all.pdf
36. Ohkuma, K.: Weak keys of reduced-round PRESENT for linear cryptanalysis. In: Jr., M.J.J., Rijmen, V., Safavi-Naini, R. (eds.) SAC 2009, Revised Selected Papers. Lecture Notes in Computer Science, vol. 5867, pp. 249–265. Springer (2009)
37. Preneel, B., Takagi, T. (eds.): CHES 2011, LNCS, vol. 6917. Springer (2011)
38. Selçuk, A.A.: On probability of success in linear and differential cryptanalysis. J. Cryptology 21(1), 131–147 (2008)
39. Shibutani, K., Isobe, T., Hiwatari, H., Mitsuda, A., Akishita, T., Shirai, T.: Piccolo: An ultra-lightweight blockcipher. In: Preneel and Takagi [37], pp. 342–357
40. Vora, P.L., Mir, D.J.: Related-key linear cryptanalysis. ISIT '06: Proceedings of the 2006 IEEE International Symposium of Information Theory 2006, 1609–1613 (July 2006)

41. Wang, N., Wang, X., Jia, K., Zhao, J.: Differential attacks on reduced simon versions with dynamic key-guessing techniques. Cryptology ePrint Archive, Report 2014/448 (2014), http://eprint.iacr.org/2014/448
42. Winternitz, R.S., Hellman, M.E.: Chosen-key attacks on a block cipher. Cryptologia 11(1), 16–20 (1987)

A Searching Method for the Linear Trails of SIMON

Linear trails with the squared correlation less than 2^{-n} for SIMONn/k are not provided in the previous works. To find such trails, we apply an efficient search algorithm that we have designed by slightly modifying the searching method described in [32] that is based on M. Matsui's branch-and-bound algorithm. As noted before, an r-round linear trail of SIMONn/k can be represented as a sequence of $(r+2)$ $n/2$-bit masks: $\Gamma_s.\Gamma_{s+1}.\cdots.\Gamma_{s+r+1}$ represents a linear trail such that at the $(i+1)$-th round, the input and output masks are $\Gamma_i \| \Gamma_{i+1}$ and $\Gamma_{i+1} \| \Gamma_{i+2}$, respectively, for each $i \in [s..(s+r-1)]$. Such a linear trail leads to the related-key linear approximation (5). The linear correlation of the round function of SIMON with respect to various input-output mask pairs can be easily computed as explained in [29] and [32]. Throughout this section, f denotes the round function of SIMONn/k that sends an $n/2$-bit input x onto $((x \lll 8) \wedge (x \lll 1)) \oplus (x \lll 2)$. Note that if $\Gamma_s.\cdots.\Gamma_{s+r+1}$ is a linear trail with the correlation ϵ, so are the reversed trail $\Gamma_{s+r+1}.\cdots.\Gamma_s$ and the rotated trail $(\Gamma_s \lll l).\cdots.(\Gamma_{s+r+1} \lll l)$ for each l.
For each $l \geq 0$, let $\mathcal{L}_l$ be the list of mask pairs (α, β) such that $|\varepsilon_f(\alpha, \beta)| = 2^{-l}$. We denote $-\log_2(|\varepsilon_f(\beta \ggg 2, \beta)|)$ by $\mathrm{lac}(\beta)$. Then the mask pairs in $\mathcal{L}_l$ are exactly those (α, β)'s for which $\mathrm{lac}(\beta) = l$ and $\varepsilon_f(\alpha, \beta) \neq 0$. A mask β is called a rotational representative if $(\beta \lll j) \geq \beta$ for each j. For each $l \geq 0$, let $\mathcal{L}_l^{\mathrm{red}}$ be the set of mask pairs (α, β) in $\mathcal{L}_l$ such that β is a rotational representative. For an r-round trail $T = \Gamma_0.\Gamma_1.\cdots.\Gamma_{r+1}$ and $l \leq r+1$, T_l denotes Γ_l and $\mathrm{lac}(T)$ denotes $\sum_{l=1}^{r} \mathrm{lac}(\Gamma_l)$. Also for $l \leq r$, $T|_l$ denotes the l-round subtrail $\Gamma_0.\Gamma_1.\cdots.\Gamma_{l+1}$.
When we search for an r-round trail T, we impose the following restrictions on the masks in the trail:

- $\mathrm{lac}(T_i) \leq 4$ for $i = 0, \ldots, r$.
- If $T_1 = 0$, T_0 is a rotational representative. Otherwise, T_1 is a rotational representative.

It turns out that these restrictions let us quickly find out the trails suitable for our purposes. For each r, let B_r be the minimum of $\mathrm{lac}(T)$ when T runs among all the r-round trails such that each mask in the trail except the last one has $\mathrm{lac} \leq 4$. It is easy to see that $B_1 = 0$, $B_2 = 1$, and $B_3 = 2$. In the search algorithm, for each $r \geq 4$, we get B_r and an r-round linear trail T with $\mathrm{lac}(T) = B_r$ assuming that we have already computed B_1, …, B_{r-1}. The search algorithm is presented as Alg. 2. Before running it, we prepare two lists of $n/2$-bit masks in advance for acceleration: One with β's such that $\mathrm{lac}(\beta) \leq 4$ and the other with β's such that $\mathrm{lac}(\beta) \leq 4$ *and* β is a rotational representative. B_r's we have obtained are the

same as those presented in [32] for SIMONn/k with n=32, 48, or 64 whenever $B_r \leq n$. We remark that by modifying Alg. 2, we can also find many linear trails with various constraints on the intermediate masks and the correlation.

Alg. 2 The search algorithm

```
Set B̄ = B_{r−1} − 1, and found = 0.
repeat
    B̄++
    ProcessR1( )
until found == 1
Output T, B̄

function ProcessR1( )
    ProcessR2A( )
    for l ← 1 to min(4, B̄ − B_{r−1}) do
        for (α, β) ∈ L_l^red do
            ProcessR2(α, β)
        end for
    end for
    return
end function

function ProcessR2A( )
    for l ← 1 to min(4, B̄ − B_{r−2}) do
        for (α, β) ∈ L_l^red do
            Set T_0 = β, T_1 = 0, T_2 = β, T_3 = α.
            ProcessR(3)
        end for
    end for
    return
end function

function ProcessR2(α,β)
    Set c = lac(β).
    for l ← 0 to min(4, B̄ − c − B_{r−2}) do
        for (α_1, β_1) ∈ L_l for which lac(α ⊕ β_1) ≤ 4 do
            Set T_0 = α ⊕ β_1, T_1 = β, T_2 = β_1, T_3 = β ⊕ α_1.
            ProcessR(3)
        end for
    end for
    return
end function
```

Alg. 2 The search algorithm (continued)

```
function ProcessR(m)
    Set β_m = T_m, c = lac(β_m).
    if m < r then
        if (lac(T|_{m-1}) + c + B_{r-m} > B̄) or (c > 4) then
            return
        else
            for α_m for which ε_f(α_m, β_m) ≠ 0 do
                Set T_{m+1} = α_m ⊕ T_{m-1}.
                ProcessR(m + 1)
            end for
        end if
    else
        if lac(T|_{m-1}) + c == B̄ then
            Choose an α_m for which ε_f(α_m, β_m) ≠ 0.
            Set T_{m+1} = α_m ⊕ T_{m-1}.
            Set found = 1, and exit all the functions.
        else
            return
        end if
    end if
end function
```

B Linear Trails of Simon

The linear trails we have used in Sect. 4.4 are as follows:

- a 16-round trail for Simon32/64 with the correlation $\pm 2^{-21}$:
  4000.0001.0000.0001.4000.1001.0400.1101.4040.0111.0004.0110.0040.0100.0000.0100.0040.0110.
- a 20-round trail for Simon48/96 with the correlation $\pm 2^{-33}$:
  400000.000001.000000.000001.400000.100001.040000.110001.404000.011001.018400.013001.40c000.110001.040000.100001.400000.000001.000000.000001.400000.100001
- a 26-round trail for Simon64/128 with the correlation $\pm 2^{-45}$:
  00000001.00004044.00001010.00004440.00000100.00004400.00001000.00004000.00000000.00004000.00001000.00004400.00000100.00004440.00001010.00004044.00000061.0000404c.00001030.00004440.00000100.00004400.00001000.00004000.00000000.00004000.00001000.00004400
- a 45(=32+13)-round trail for Simon128/256 with the correlation $\pm 2^{-84}$:
  0100000000000001.0440000000000004.0610000000000000.04c0000000000004.$\cdots$.0100000000000001.0440000000000004.0610000000000000.04c0000000000004.$\cdots$.0000000000000001.4000000000000004.1000000000000000
- a 51(=48+3)-round trail for Simon128/256 with the correlation $\pm 2^{-92}$:
  0000000000000004.0000000000000000.0000000000000004.0000000

000000001.$\cdots$.0000000000000004.0000000000000000.0000000000000004.0000000000000001.$\cdots$.0000000000000004.0000000000000001.4000000000000004

Linear trails for SIMON32, SIMON48, SIMON64 were found by our search algorithm, but the trail for SIMON128 was obtained as a subtrail of an iterative trail in [32].

C Adding More Rounds to Related-Key Linear Approximations of Simon

In this section, we will explain how to add $r_{\mathrm{pre}} + r_{\mathrm{post}}$ rounds for $(r_{\mathrm{pre}}, r_{\mathrm{post}}) =$ (3,3), (4,4), or (5,5). For simplicity $a + b$ and ab (or $a \bullet b$) denote the XOR and AND of $a, b \in \mathbb{F}_2$, respectively. Let $w = n/2$ be the word size as before.

C.1 Adding 3+3 Rounds

3-round computation. Let rk_0, rk_1, rk_2 be the round keys derived from the candidate key K for the first 3 rounds. For a plaintext $P = P_{\mathrm{L}} \| P_{\mathrm{R}}$ and a key difference ΔK, let $\delta rk_0, \delta rk_1, \delta rk_2$ be the derived round key differences for the first 3 rounds. Let $A = f(P_{\mathrm{L}}) \oplus P_{\mathrm{R}} \oplus \delta rk_0$, and $B = P_{\mathrm{L}} \oplus \delta rk_1$ as before. Then using the relations $X_{3,\mathrm{R}} = X_{2,\mathrm{L}}$ and $X_{3,\mathrm{L}} = f(X_{2,\mathrm{L}}) \oplus X_{2,\mathrm{R}} \oplus \delta rk_2 \oplus rk_2$, we can compute each bit of $X_3 = X_{3,\mathrm{L}} \| X_{3,\mathrm{R}} = E_0^2(K \oplus \Delta K, P)$ in terms of bits of $A, B, rk_0, rk_1, rk_2, \delta rk_0, \delta rk_1, \delta rk_2$ as follows:

$$\begin{aligned} X_{3,\mathrm{L}}[i] = {} & \big((rk_0 \oplus A)[i-9](rk_0 \oplus A)[i-2] + (rk_0 \oplus A)[i-3] + (rk_1 \oplus B)[i-1]\big) \bullet \\ & \big((rk_0 \oplus A)[i-16](rk_0 \oplus A)[i-9]) + (rk_0 \oplus A)[i-10] + (rk_1 \oplus B)[i-8]\big) \\ & + (rk_0 \oplus A)[i-10](rk_0 \oplus A)[i-3] + \underline{rk_0[i-4]} + A[i-4] + \underline{rk_0[i]} + A[i] \\ & + \underline{rk_1[i-2]} + B[i-2] + \underline{rk_2[i]} + \delta rk_2[i], \\ X_{3,\mathrm{R}}[i] = {} & X_{2,\mathrm{L}}[i]. \end{aligned}$$

Thus

- $X_{3,\mathrm{L}}[i]$ can be computed in terms of $(rk_0 \oplus A)[i-2, i-3, i-9, i-10, i-16]$, $(rk_1 \oplus B)[i-1, i-8]$, xored with $A[i-4] + B[i-2] + A[i] + \delta rk_2[i]$ and the underlined terms determined only by rk_0, rk_1, rk_2.
- $X_{3,\mathrm{R}}[i]$ can be computed in terms of $(rk_0 \oplus A)[i-1], (rk_0 \oplus A)[i-8]$, xored with $A[i-2] + B[i]$ and terms determined only by rk_0, rk_1, rk_2.

By symmetry of the cipher structure, we get similar expressions for bits of $X_{R-3,\mathrm{R}}$ and $X_{R-3,\mathrm{L}}$ in terms of $A' = f(C_{\mathrm{R}}) \oplus C_{\mathrm{L}} \oplus \delta rk_{R-1}$, $B' = C_{\mathrm{R}} \oplus \delta rk_{R-2}$, δrk_{R-3}, rk_{R-1}, rk_{R-2}, and rk_{R-3}.

The data compression. Suppose that we want to make use of the related-key linear approximation represented as (5) with $s = 3$ and $s + r + 3 = R$. Let $\mathcal{I}_{\mathrm{L}} = \mathsf{supp}(\Gamma_s)$, $\mathcal{I}_{\mathrm{R}} = \mathsf{supp}(\Gamma_{s+1})$, $\mathcal{I}'_{\mathrm{L}} = \mathsf{supp}(\Gamma_{R-3})$, and $\mathcal{I}'_{\mathrm{R}} = \mathsf{supp}(\Gamma_{R-2})$. The compression function extracts the following values from each data entry $(P, C, \Delta K)$:

- $A[i]$ for i such that one of $i+2, i+3, i+9, i+10, i+16 \bmod w$ is in $\mathcal{I}_{\mathrm{L}}$
- $B[i]$ for i such that one of $i+1, i+8 \bmod w$ is in $\mathcal{I}_{\mathrm{R}}$
- $A'[i]$ for i such that one of $i+2, i+3, i+9, i+10, i+16 \bmod w$ is in $\mathcal{I}'_{\mathrm{R}}$
- $B'[i]$ for i such that one of $i+1, i+8 \bmod w$ is in $\mathcal{I}'_{\mathrm{L}}$
- $\bigoplus_{i\in\mathcal{I}_{\mathrm{L}}}(A[i-4]+B[i-2]+A[i]+\delta rk_2[i]) \oplus \bigoplus_{i\in\mathcal{I}_{\mathrm{R}}}(A[i-2]+B[i]) \oplus \bigoplus_{i\in\mathcal{I}'_{\mathrm{R}}}(A'[i-4]+B'[i-2]+A'[i]+\delta rk_{R-3}[i]) \oplus \bigoplus_{i\in\mathcal{I}'_{\mathrm{L}}}(A'[i-2]+B'[i]) \oplus \langle \Lambda, \Delta K\rangle$

The outer round key bits we need to guess are as follows:

- $rk_0[i]$ for i such that one of $i+2, i+3, i+9, i+10, i+16 \bmod w$ is in $\mathcal{I}_{\mathrm{L}}$
- $rk_1[i]$ for i such that one of $i+1, i+8 \bmod w$ is in $\mathcal{I}_{\mathrm{L}}$
- $rk_0[i]$ for i such that one of $i+1, i+8 \bmod w$ is in $\mathcal{I}_{\mathrm{R}}$
- $rk_{R-1}[i]$ for i such that one of $i+2, i+3, i+9, i+10, i+16 \bmod w$ is in $\mathcal{I}'_{\mathrm{R}}$
- $rk_{R-2}[i]$ for i such that one of $i+1, i+8 \bmod w$ is in $\mathcal{I}'_{\mathrm{R}}$
- $rk_{R-1}[i]$ for i such that one of $i+1, i+8 \bmod w$ is in $\mathcal{I}'_{\mathrm{L}}$

Note that the number k_O of guessed round key bits for outer rounds is at most $7\mathrm{wt}(\Gamma_s) + 2\mathrm{wt}(\Gamma_{s+1}) + 2\mathrm{wt}(\Gamma_{R-3}) + 7\mathrm{wt}(\Gamma_{R-2})$ and d, $\log_2$ of the size of the compressed data, is $k_O + 1$.

C.2 Adding 4+4 Rounds

Let $A = f(P_{\mathrm{L}}) \oplus P_{\mathrm{R}} \oplus \delta rk_0$, and $B = P_{\mathrm{L}} \oplus \delta rk_1$ as before. Using the relations $X_{4,\mathrm{R}} = X_{3,\mathrm{L}}$ and $X_{4,\mathrm{L}} = f(X_{3,\mathrm{L}}) \oplus X_{3,\mathrm{R}} \oplus \delta rk_3 \oplus rk_3$, we see that $X_{4,\mathrm{L}}[i]$ (up to a constant determined only by rk_0, rk_1, rk_2, rk_3) is a function of the following terms

- $(rk_0 \oplus A)[i-1, i-3, i-4, i-5, i-8, i-10, i-11, i-12, i-17, i-18, i-24]$
- $(rk_1 \oplus B)[i-2, i-3, i-9, i-10, i-16]$
- $(rk_2 \oplus \delta rk_2)[i-1, i-8]$

xored with $A[i-6] + B[i] + B[i-4] + \delta rk_2[i-2] + \delta rk_3[i]$. Note also that $X_{4,\mathrm{R}}[i] = X_{3,\mathrm{L}}[i]$ and the backward computations can be carried out similarly. So when we use a related-key linear approximation represented as (5) with $s = 4$ and $s + r + 4 = R$, we have a compression with $k_O \leq 18\mathrm{wt}(\Gamma_s) + 7\mathrm{wt}(\Gamma_{s+1}) + 7\mathrm{wt}(\Gamma_{R-4}) + 18\mathrm{wt}(\Gamma_{R-3})$ and $d = k_O + 1$.

C.3 Adding 5+5 Rounds

Let $A = f(P_{\mathrm{L}}) \oplus P_{\mathrm{R}} \oplus \delta rk_0$, and $B = P_{\mathrm{L}} \oplus \delta rk_1$. $X_{4,\mathrm{L}}[i]$ (up to a constant determined only by $rk_0, rk_1, rk_2, rk_3, rk_4$) is a function of the following terms

- $(rk_0 \oplus A)[i-2, i-3, i-4, i-5, i-6, i-7, i-9, i-10, i-11, i-12, i-13, i-14, i-16, i-18, i-19, i-20, i-25, i-26, i-32]$
- $(rk_1 \oplus B)[i-1, i-3, i-4, i-5, i-8, i-10, i-11, i-12, i-17, i-18, i-24]$
- $(rk_2 \oplus \delta rk_2)[i-2, i-3, i-9, i-10, i-16]$

– $(rk_3 \oplus \delta rk_3)[i-1, i-8]$

xored with $A[i]+A[i-4]+A[i-8]+B[i-6]+\delta rk_2[i]+\delta rk_2[i-4]+\delta rk_3[i-2]+\delta rk_4[i]$. We also have $X_{5,\mathrm{R}}[i] = X_{4,\mathrm{L}}[i]$ and the backward computations can be carried out similarly. So when we use a related-key linear approximation represented as (5) with $s = 5$ and $s + r + 5 = R$, we have a compression with $k_O \leq 39\mathrm{wt}(\Gamma_s) + 18\mathrm{wt}(\Gamma_{s+1}) + 18\mathrm{wt}(\Gamma_{R-5}) + 39\mathrm{wt}(\Gamma_{R-4})$ and $d = k_O + 1$.

D Application of FWHT to Related-key Linear Attacks

We will explain how FWHT can be applied in our related-key linear attacks with an example presented in Sect. 5.2. Consider the 8-round attack on PRESENT-L prepending 2 rounds to the 6-round linear trail T-6R-B42 using RKLC-2. Suppose that we have a related-key data D obtained from a base key K^*. We set the compression function H_c as the 21(=16+4+1)-bit valued function defined by

$$(P, C, \Delta K) \mapsto (P \oplus \delta rk_0)[32..47] \| \delta rk_1[40..43] \| (\bigoplus\nolimits_{i=2}^{8} \delta rk_i[42] \oplus C[42]).$$

The round key bits and the parity bit to recover are $rk_0^*[32..47] \| rk_1^*[40..43]$ and $\bigoplus_{i=2}^{8} rk_i^*[42]$, respectively. In Step 1, we perform the data compression to get the compressed set

$$\{(v, n_v) \in \mathbb{F}_2^{21} \times \mathbb{Z} : n_v = |\{(P, C, \Delta K) \in D : H_c(P, C, \Delta K) = v\}|\}.$$

Let h_1 be 4-bit valued function such that

$$h_1(x) = S(x[12..15])[2] \| S(x[8..11]))[2] \| S(x[4..7]))[2] \| S(x[0..3])[2]$$

for each $x \in \mathbb{F}_2^{21}$. Then let h' be the 1-bit valued function such that $h'(x) = S(h_1(x) \oplus x[16..19])[2] \oplus x[20]$ for each $x \in \mathbb{F}_2^{21}$. Then we have to compute $\sum_v n_v (-1)^{h'(z \oplus v)}$ for each $z \in \mathbb{F}_2^{21}$ in Step 2. This can be done by performing FWHTs three times with memory $O(2^{21})$ just as described in [18].

E Additional Results with Present-L

In this section, we present some of the results obtained from the experiments in Sect. 5.2 not provided there due to the lack of space. They are results from the 6-round key recovery using T-4R-B42 (Fig. 11) and the 8-round key recovery using T-6R-B21 (Fig. 12).

F Keys and Data Used in the Experiments

For each cipher, linear trail, and data size, 1,000 keys were used as the targets for recovery. For variants of SIMON, we used a single GTX 1060 GPU and a Core i7-5820K CPU to generate the data and get the results. Experiment with each cipher took just a few days or less.

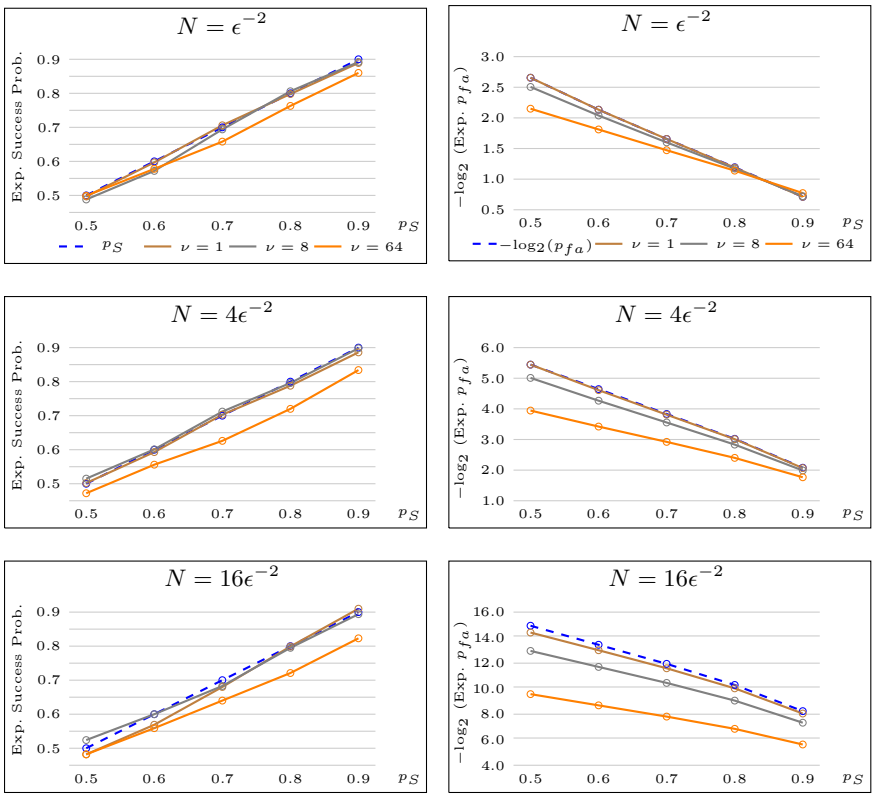

Fig. 11: Experimental results for 6-round key recovery using T-4R-B42

F.1 Experiments with the Small-scale Simon

For experiments with 22-round SIMON24 and 16-round SIMON32, the number ν of data entries per key difference was fixed at 2^{16} regardless of N. The following keys and data were used for 22-round SIMON24.

- base keys K^l $(l = 0, \ldots, 999)$:
  - $K^l[36..47] = \texttt{0x012} + l \pmod{2^{12}}$, $K^l[24..35] = \texttt{0x345} + l \pmod{2^{12}}$
  - $K^l[12..23] = \texttt{0x678} + l \pmod{2^{12}}$, $K^l[0..11] = \texttt{0x9ab} + l \pmod{2^{12}}$
- $\Delta K_i^l = \texttt{0xfedcba987654321} \times (Nl/\nu + i) \pmod{2^{48}}$ $(i = 0, \ldots, N/\nu - 1)$
- $P_{i,j}^l = j$, $j = 0, \ldots, \nu - 1$

Similar keys and data were used for 16-round SIMON32:

- base keys K^l $(l = 0, \ldots, 999)$:
  - $K^l[48..63] = \texttt{0x0123} + l \pmod{2^{16}}$, $K^l[32..47] = \texttt{0x4567} + l \pmod{2^{16}}$
  - $K^l[16..31] = \texttt{0x89ab} + l \pmod{2^{16}}$, $K^l[0..11] = \texttt{0xcdef} + l \pmod{2^{16}}$
- $\Delta K_i^l = \texttt{0xfedcba987654321} \times (Nl/\nu + i) \pmod{2^{64}}$ $(i = 0, \ldots, N/\nu - 1)$
- $P_{i,j}^l = j$, $j = 0, \ldots, \nu - 1$

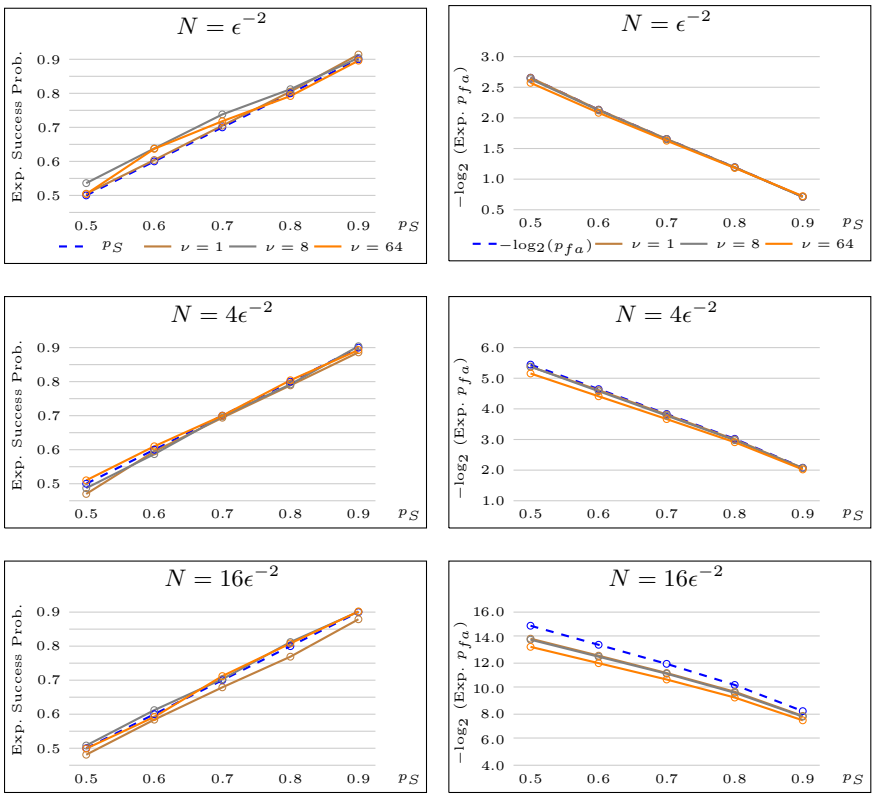

Fig. 12: Experimental results for 8-round key recovery using T-6R-B21

F.2 Experiments with Present-L

The keys and data used are as follows regardless of the number of rounds or the linear trail:

- base keys K^l $(l = 0, \ldots, 999)$:
  - $K^l[64..127] = \texttt{0x123456789abcdef} \times l \pmod{2^{64}}$
  - $K^l[0..63] = \texttt{0x123456789abcdef} \times (l+1) \pmod{2^{64}}$
- key differences ΔK_i^l $(i = 0, \ldots, N/\nu - 1)$:
  - $\Delta K_i^l[64..127] = \texttt{0xfedcba987654321} \times i \times \texttt{0x321321321321321} \pmod{2^{64}}$
  - $\Delta K_i^l[0..63] = \texttt{0xfedcba987654321} \times i \times \texttt{0x432405887348} \pmod{2^{64}}$
- plaintexts $P_{i,j}^l = j$, $j = 0, \ldots, \nu - 1$