Pen and Paper Arguments for SIMON and SIMON-like Designs

Christof Beierle

Horst Görtz Institute for IT Security, Ruhr-Universität Bochum, Germany
christof.beierle@rub.de

Abstract. In this work, we analyze the resistance of SIMON-like ciphers against differential attacks without using computer-aided methods. In this context, we first define the notion of a SIMON-like cipher as a generalization of the SIMON design. For certain instances, we present a method for proving the resistance against differential attacks by upper bounding the probability of a differential characteristic by 2^{-2T+2} where T denotes the number of rounds. Interestingly, if $2n$ denotes the block length, our result is sufficient in order to bound the probability by 2^{-2n} for all full-round variants of SIMON and SIMECK. Thus, it guarantees security in a sense that, even having encryptions of the full codebook, one cannot expect a differential characteristic to hold. The important difference between previous works is that our proof can be verified by hand and thus contributes towards a better understanding of the design. However, it is to mention that we do not analyze the probability of multi-round differentials.
Although there are much better bounds known, especially for a high number of rounds, they are based on experimental search like using SAT/SMT solvers. While those results have already shown that SIMON can be considered resistant against differential cryptanalysis, our argument gives more insights into the design itself. As far as we know, this work presents the first non-experimental security argument for full-round versions of several SIMON-like instances.

Keywords: SIMON · SIMECK · differential cryptanalysis · Feistel

1 Introduction

Once a new cipher is proposed, the designers are expected to provide security arguments, at least against the most important and powerful attack vectors known, that are differential [12] and linear cryptanalysis [22]. Thus, any new design itself should allow for an, if possible simple, security argument. Nowadays, a majority of block ciphers is based on Feistel- and Substitution-Permutation (SP) constructions. As the name already implies, SP designs iterate both substitution and permutation operations. While the latter is a linear function (linear layer), the substitution layer consists of highly non-linear components (e.g. S-boxes). The alternation of those layers is responsible for both offering confusion and diffusion [26].

This separation into linear and non-linear components offers the advantage of analyzing the structure more easily. Two design principles are common, that are the wide-trail strategy [16] and the use of computer-aided methods. In the wide-trail strategy, which was introduced by Daemen and Rijmen, the idea is that the design of the linear layer is related to coding theory, as its construction is based upon a linear code over $GF(2^m)$ with high (and often optimal) minimum distance. Thereby, the parameter m defines the word size of the S-box. As the minimum distance indicates the number of active S-boxes over two consecutive rounds, it contributes to the resistance against differential and linear cryptanalysis in a provable (by pen and paper) way. A more clever choice of the linear layer even allows for arguments on four (resp. eight, sixteen,...) rounds using the so-called superbox (resp. megabox, gigabox, etc.) structure, as for example described in [8,9,17]. In fact, the Rijndael cipher [18], which was standardized as the Advanced Encryption Standard in 2000 [25], was designed according to this principle. The advantage of the wide-trail strategy is one reason why so many AES-like designs occurred in the last years. It also emphasizes that designers prefer well-understood principles. While for AES-like ciphers counting the number of active S-boxes can be somehow done independently of

This article is a preliminary version of the paper to appear in the proceedings of SCN 2016.

the choice of the S-box, some other strategies use specific properties of the non-linear components. For instance, the designers of PRESENT showed that an arbitrary five-round differential characteristic has at least 10 active S-boxes under certain assumptions [14].

The other strategy is measuring the security using computer-aided search methods. For instance, one can model the propagation of differential and linear characteristics as a mixed-integer linear programming problem [8,23,29]. Examples of a design which uses experimental arguments are KECCAK [10] and SERPENT [11]. However, the bounds obtained with this approach are not verifiable without a machine and do not contribute significantly to a better understanding of the design itself.

Basically, in both strategies, (if the non-linear component is not too weak) the design of the linear layer is the crucial step when it comes to providing security against differential and linear attacks. While a single round can often be analyzed quite easily, the analysis of the linear layer w.r.t. diffusion properties usually has to be done using a more complex argument over multiple rounds. Unfortunately, besides the wide-trail strategy, not many constructions are known that guarantee security using pen and paper arguments. Especially, almost every multi-round argument uses some sort of superbox (resp. megabox, etc.) structure. One therefore may seek for alternative design principles. Especially for lots of Feistel designs, the constructions might be less clear and less understood. However, there are some fundamental results on bounding the differential and linear behavior [24]. There are also Feistel designs which consist of SP-type round functions [27,28] combining the advantages of the Feistel construction and the simple arguments of the wide-trail strategy.

In contrast to a scientific design process, the NSA recently presented the SIMON family of lightweight block ciphers [6]. Besides its specification, no arguments on the security are provided. Especially since SIMON is an innovative Feistel cipher, its design is harder to analyze. Besides its non-bijective round function and combining the branches after every round, the difficulties are caused by the bitwise structure. Since the design choice was left unclear, one seeks for a deeper understanding of the cipher[1].

Related Work. The appearance of the SIMON family of block ciphers [6] in the cryptology eprint archive inspired the cryptographic community taking further investigations on the possible design rationale. Therefore, several cryptanalytic results followed. For instance, see [1,2,3,4,5,13,15,20,30,31,32] for a selection. They are mostly based on experimental search.

At CRYPTO 2015, Kölbl, Leander and Tiessen pointed out some interesting properties of SIMON-like round functions [21]. These observations were then used for a further analysis of the differential and linear behavior over multiple rounds. Although the analysis of the round function was done in a mathematical rigorous manner, the multi-round behavior was derived using a computer-aided approach. As one result, the rotation constants of SIMON turned out to be in some sense not optimally chosen. Inspired by the design, Yang et. al. proposed the SIMECK family of lightweight block ciphers at CHES 2015 [33]. It can be seen as a SIMON-like cipher using different rotation constants in its round function and a key schedule inspired by SPECK [6].

Recently, the designers of SIMON published a follow-up paper at the NIST lightweight workshop covering some implementation aspects [7]. However, the authors gave no additional insights into the design choice from a cryptanalytic point of view.

Contribution. After describing a generalization of the SIMON design by decoupling the round function into a linear and a non-linear component, we show that the structure of a SIMON-like

[1] As we only focus on the probabilities of differential characteristics and do not provide a full security analysis, this work should not be seen as a recommendation for using SIMON. Some design choices are still unclear. To mention is the key schedule as one example.

design allows for a proof on the resistance against differential attacks under certain assumptions. The question whether the proof works depends on the interaction between these two components. If the non-linear part ρ is of the form $\rho(\mathbf{x}) = (\mathbf{x} \lll a) \wedge (\mathbf{x} \lll b)$, it can be in general formulated as a property of the linear layer. A sufficient condition is that the linear layer has a branch number of at least 11. Since this is not the case for SIMON and SIMECK, we consider these ciphers separately. In particular, for all instantiations of SIMON and SIMECK, we are able to upper bound the probability of any differential characteristic by 2^{-2n} where $2n$ denotes the block length. We show this in detail for the example of SIMON.

In clear distinction to prior work such as [21], our argument is a formal proof covering multiple rounds and can thus be verified without experimental tools. In our approach, we use the well-known property of the SIMON-like round function that the set of possible output differences U_α defines an affine subspace depending on the input difference $\boldsymbol{\alpha}$ and that the differential probability highly depends on the Hamming weight of $\boldsymbol{\alpha}$. The main idea is that we extend the analysis of the round function to the cases where $\boldsymbol{\alpha}$ has a Hamming weight equal to 2 and consider the propagation of Hamming weights over the Feistel structure.

Figure 1 illustrates the bounds proven with our method and, as a comparison, the bounds obtained from experimental search described in [21, Section 5.2] for two instances of SIMON. It is to mention that, although our bounds are worse than the experimental results, they are still much better than the bounds one obtains by trivially multiplying the worst-case probabilities for every round. Moreover, since the development of the experimental bounds becomes more complex for a high number of rounds, we believe that one cannot expect to significantly improve upon our theoretical result by using a simple argument. Such an argument will likely cover lots of rounds.

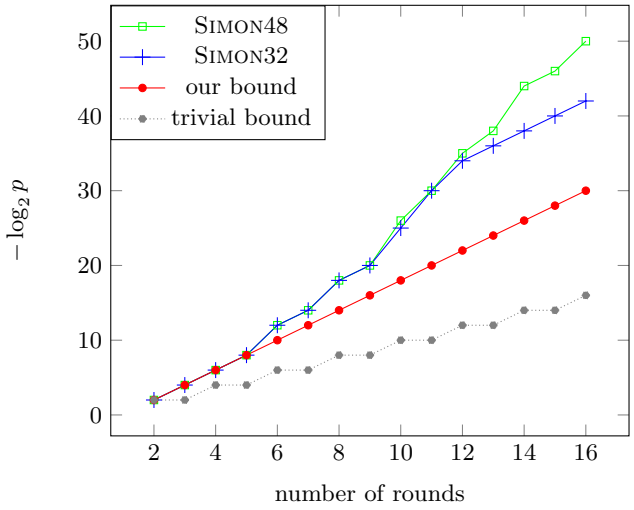

Fig. 1. Comparison of the experimental bounds for SIMON32 and SIMON48 as described in [21, Section 5.2] and our provable bounds.

2 Preliminaries

Elements in the vector space $\mathbb{F}_2^n$ are denoted with bold letters. The all zero vector will be denoted by $\mathbf{0}$ and the all one vector by $\mathbf{1}$, respectively. We use $\mathrm{wt}(\mathbf{x})$ to denote the Hamming weight of a vector $\mathbf{x} = (x_0, x_1, \ldots, x_{n-1})$. Moreover, a superscript notation is used for describing the index of a component. For example, the element $(0, \ldots, 0, y^{(i)}, 0, 0, \ldots, 0)$ denotes the vector $(x_0, x_1, \ldots, x_{n-1})$ with $x_i = y$ and $x_k = 0$ for all $k \neq i$. The Boolean operations, bitwise AND,

OR, NOT and bitwise XOR, are denoted by $\wedge, \vee, \bar{}$ and $\oplus$, respectively. A cyclic rotation (with offset r) is denoted by $\lll r$, if the rotation is to the left, and by $\ggg r$, if the rotation is to the right.

Differential Cryptanalysis. In the following, we recall the basic definitions in differential cryptanalysis. We use the notion of XOR differences in this context.

Definition 1. *For a vectorial function $f : \mathbb{F}_2^n \to \mathbb{F}_2^m$, the* probability of the differential $\boldsymbol{\alpha} \xrightarrow{f} \boldsymbol{\beta}$ *is defined as*

$$P(\boldsymbol{\alpha} \xrightarrow{f} \boldsymbol{\beta}) := \frac{|\Delta_f(\boldsymbol{\alpha}, \boldsymbol{\beta})|}{2^n}$$

where

$$\Delta_f(\boldsymbol{\alpha}, \boldsymbol{\beta}) := \{\mathbf{x} \in \mathbb{F}_2^n \mid \boldsymbol{\beta} = f(\mathbf{x}) \oplus f(\mathbf{x} \oplus \boldsymbol{\alpha})\}.$$

If f_i denotes the i-th round function of an iterated cipher, a *valid T-round differential characteristic* $C : \boldsymbol{\alpha_0} \xrightarrow{f_1} \boldsymbol{\alpha_1} \xrightarrow{f_2} \ldots \xrightarrow{f_T} \boldsymbol{\alpha_T}$ has $|\Delta_{f_i}(\boldsymbol{\alpha_{i-1}}, \boldsymbol{\alpha_i})| \neq 0$ for all $1 \leq i \leq T$. Assuming that the probabilities of all one-round differentials are independent, we compute the probability of the characteristic C as

$$P(C) = \prod_{i=1}^{T} P(\boldsymbol{\alpha_{i-1}} \xrightarrow{f_i} \boldsymbol{\alpha_i}).$$

Note that for a key-alternating cipher, this holds under the assumption of independent round-keys. When designing a block cipher, one would like to avoid the existence of (multi-round) differentials with high probability. Since in general, computing the maximum probability of multi-round differentials is not a trivial task, one concentrates on upper bounding the probability of a characteristic instead. If n denotes the block length, a typical approach is to estimate the number of rounds T' such that $P(C) \leq 2^{-n}$ for any T'-round characteristic C and specify the number of rounds of the primitive as $T = T' + \kappa$ with a reasonable security margin κ.

A Remark on the Feistel Construction. We point out a useful property of the Feistel construction in the following. Recall that for a vectorial function $f : \mathbb{F}_2^n \to \mathbb{F}_2^n$ and $\mathbf{k} \in \mathbb{F}_2^n$, we define a *Feistel round function* as

$$\begin{aligned} F_{\mathbf{k}}^f : \mathbb{F}_2^n \times \mathbb{F}_2^n &\to \mathbb{F}_2^n \times \mathbb{F}_2^n \\ (\mathbf{x}, \mathbf{y}) &\mapsto (f(\mathbf{x}) \oplus \mathbf{y} \oplus \mathbf{k}, \mathbf{x}). \end{aligned}$$

Thereby, f is called the *Feistel function* (or simply *f-function*) and $\mathbf{k}$ is called the *round key*. For simplicity, we will use an identical Feistel function f in every round.

A difference within the Feistel cipher is denoted as $(\boldsymbol{\gamma}, \boldsymbol{\delta})$ describing the left and the right branch, respectively. Lemma 1 presents a general observation on the Feistel construction. It states that, having upper bounds on the probability for all differential characteristics starting with $(\mathbf{0}, \boldsymbol{\alpha})$ and ending with $(\mathbf{0}, \boldsymbol{\beta})$, one can easily bound the probability of any characteristic.

Lemma 1. *For $t \geq 1$, let for all non-zero differences $\boldsymbol{\alpha}, \boldsymbol{\beta}$, the differential probability of any t-round characteristic starting with $(\mathbf{0}, \boldsymbol{\alpha})$ and ending with $(\mathbf{0}, \boldsymbol{\beta})$ be upper bounded by $p(t)$.*

Let further $p(0) := 1$ and $q := \max_{\alpha \neq 0, \beta} P(\boldsymbol{\alpha} \xrightarrow{f} \boldsymbol{\beta})$. Then,

$$P(C) \leq \max_{k \leq T} p(k) q^{T-k-1}$$

for any non-trivial T-round characteristics C with $T > 0$.

Proof. For a given T-round characteristic $C = (\boldsymbol{\gamma_0}, \boldsymbol{\delta_0}) \xrightarrow{F^f} \ldots \xrightarrow{F^f} (\boldsymbol{\gamma_T}, \boldsymbol{\delta_T})$, it holds that $\mathrm{P}(C) = \prod_{i=0}^{T-1} P(\boldsymbol{\gamma_i} \xrightarrow{f} \boldsymbol{\gamma_{i+1}})$ assuming independent probabilities. The proof is now split into two cases.

(i) Let's assume that there exist distinct i, j such that $\boldsymbol{\gamma_i} = \boldsymbol{\gamma_j} = \mathbf{0}$. Then one can choose w.l.o.g two distinct indices i', j' such that $\boldsymbol{\gamma_{i'}} = \boldsymbol{\gamma_{j'}} = \mathbf{0}$ and $\boldsymbol{\gamma_k} \neq \mathbf{0}$ for all $k < i'$ and all $k > j'$. Now, by definition

$$\mathrm{P}((\boldsymbol{\gamma_{i'}}, \boldsymbol{\delta_{i'}}) \xrightarrow{F^f} \ldots \xrightarrow{F^f} (\boldsymbol{\gamma_{j'}}, \boldsymbol{\delta_{j'}})) \leq p(j' - i').$$

Since $\boldsymbol{\gamma_{j'}} = \mathbf{0}$ and all other $\boldsymbol{\gamma_k} \neq \mathbf{0}$, we have

$$\begin{aligned}\mathrm{P}(C) &\leq p(j' - i') \prod_{k=0}^{i'-1} P(\boldsymbol{\gamma_k} \xrightarrow{f} \boldsymbol{\gamma_{k+1}}) \prod_{k=j'+1}^{T-1} P(\boldsymbol{\gamma_k} \xrightarrow{f} \boldsymbol{\gamma_{k+1}}) \\ &\leq p(j' - i') q^{i'} q^{T-(j'+1)} = p(j' - i') q^{T-(j'-i')-1}.\end{aligned}$$

(ii) If $\boldsymbol{\gamma_i} = \mathbf{0}$ for at most one i, then

$$\prod_{k<T} P(\boldsymbol{\gamma_k} \xrightarrow{f} \boldsymbol{\gamma_{k+1}}) \leq \prod_{k \neq i} P(\boldsymbol{\gamma_k} \xrightarrow{f} \boldsymbol{\gamma_{k+1}}) \leq q^{T-1} = p(0) q^{T-1}.$$

□

As Lemma 1 is a general statement for all Feistel ciphers, we give a simplified version in Section 3 as Corollary 1. It covers the special case of a SIMON-like round function, which will be defined next.

SIMON and SIMON-like Ciphers. We generalize the design of the SIMON block cipher to the SIMON-like structure. Figure 2 illustrates this construction. For the SIMON-like design, one requires a quadratic, rotational invariant function as the non-linear component. A vectorial function $f : \mathbb{F}_2^n \to \mathbb{F}_2^n$ is called *rotational invariant* iff $f(\mathbf{x} \lll r) = (f(\mathbf{x}) \lll r)$ for all elements $\mathbf{x} \in \mathbb{F}_2^n$ and all offsets r. This leads to the following definition.

Definition 2. *A* SIMON-like f-function *is composed of an $\mathbb{F}_2$-linear function θ and a degree-2 function ρ of the form $\rho(\mathbf{x}) = \vartheta_1(\mathbf{x}) \wedge \vartheta_2(\mathbf{x})$ with $\mathbb{F}_2$-linear and rotational invariant ϑ_i as*

$$f_S : \mathbb{F}_2^n \to \mathbb{F}_2^n, \mathbf{x} \mapsto \rho(\mathbf{x}) \oplus \theta(\mathbf{x}).$$

In this context, a SIMON-like cipher *uses such an f-function in a Feistel construction.*

Note that the rotational invariance is, in this general case, not required for the linear part θ.

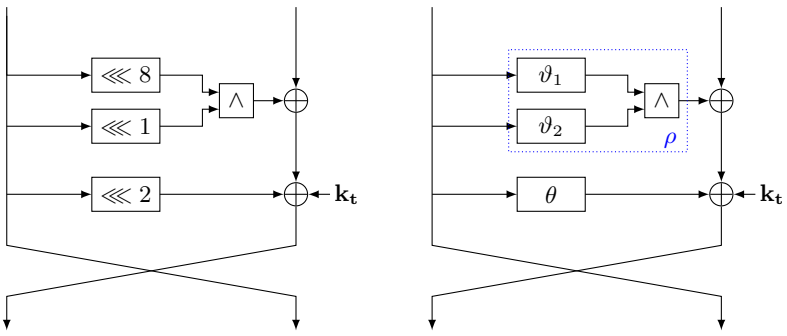

Fig. 2. Illustration of the SIMON and the generalized SIMON-like round function

3 Analysis of Differential Characteristics

In this section, we analyze the propagation characteristics of differences over several rounds under certain assumptions. We rely on the fact that a single SIMON-like round is quite well understood. Let

$$L_\alpha(\mathbf{x}) := (\vartheta_1(\mathbf{x}) \wedge \vartheta_2(\boldsymbol{\alpha})) \oplus (\vartheta_1(\boldsymbol{\alpha}) \wedge \vartheta_2(\mathbf{x})).$$

We first recall the observation that for any input difference $\boldsymbol{\alpha} \in \mathbb{F}_2^n$ into a SIMON-like round function f_S, the output difference lies in the affine subspace $U_\alpha := \operatorname{Im} L_\alpha + f_S(\boldsymbol{\alpha})$. This is formally stated in Theorem 1.

Theorem 1 (Kölbl, Leander, Tiessen [21]). *For an input difference $\boldsymbol{\alpha} \in \mathbb{F}_2^n$ into f_S, the set of possible output differences defines an affine subspace U_α s.t. $P(\boldsymbol{\alpha} \xrightarrow{f_S} \boldsymbol{\beta}) \neq 0$ if and only if $\boldsymbol{\beta} \in U_\alpha$. Defining $d_\alpha := \dim \operatorname{Im} L_\alpha$ it holds*

$$\boldsymbol{\beta} \in U_\alpha \Leftrightarrow \boldsymbol{\beta} \oplus f_S(\boldsymbol{\alpha}) \in \operatorname{Im} L_\alpha$$

and $P(\boldsymbol{\alpha} \xrightarrow{f_S} \boldsymbol{\beta}) = 2^{-d_\alpha}$ for all valid differentials over f_S.

Since the probability is the same for all output differences $\boldsymbol{\beta}$ in this subspace, we simply write p_α for $P(\boldsymbol{\alpha} \xrightarrow{f_S} \boldsymbol{\beta})$ with $\boldsymbol{\beta} \in U_\alpha$. For all output differences which are not elements in this subspace, the probability will be zero.

Because of the rotational invariance, it holds that $\operatorname{Im} L_{(\alpha \lll r)} = (\operatorname{Im} L_\alpha \lll r)$ with $p_{(\alpha \lll r)} = p_\alpha$. One can thus restrict the consideration to a single representative of this equivalence class if only one round function is analyzed.

3.1 Restriction to $\vartheta_1(x) = (x \lll a)$ and $\vartheta_2(x) = (x \lll b)$

This describes the most simple structure of a generalized SIMON-like cipher. For the θ step defined as $\theta(\mathbf{x}) = (\mathbf{x} \lll c)$, one obtains SIMON and SIMECK as a special case using $(8, 1, 2)$, resp. $(5, 0, 1)$, as a choice for the rotation constants (a, b, c). The following lemma states that we can obtain an upper bound on the differential probability over f_S depending on the Hamming weight of the input difference. While a weaker version of Lemma 2 can be deduced from [21, Theorem 3, p. 9], we improved the bound from [21] if the Hamming weight of the input difference equals 2. Although this improvement seems to be of little importance at a first glance, it is exactly this tighter bound which allows us to prove the main result. Thus, Lemma 2, and especially case (2), is one of the core components in our proof of the upper bound on the probability of differential characteristics.

Lemma 2. *Let $\vartheta_1(x) = (x \lll a)$ and $\vartheta_2(x) = (x \lll b)$. Assume that $n \geq 6$ is even and $\gcd(a - b, n) = 1$. Let $\boldsymbol{\alpha}$ be an input difference into f_S. Then, for the differential probability over f_S it holds that*

(1) If $\operatorname{wt}(\boldsymbol{\alpha}) = 1$, then $p_\alpha \leq 2^{-2}$.
(2) If $\operatorname{wt}(\boldsymbol{\alpha}) = 2$, then $p_\alpha \leq 2^{-3}$.
(3) If $\operatorname{wt}(\boldsymbol{\alpha}) \neq n$, then $p_\alpha \leq 2^{-\operatorname{wt}(\alpha)}$.
(4) If $\operatorname{wt}(\boldsymbol{\alpha}) = n$, then $p_\alpha \leq 2^{-n+1}$.

Proof. Without loss of generality one can assume that $b = 0$ and $a < \frac{n}{2}, a \neq 0$ because of the rotational invariance and since $a - b$ and n are coprime. According to [21, Theorem 3, p. 9], it is $p_\alpha = 2^{-d_\alpha}$ with

$$d_\alpha = \begin{cases} \operatorname{wt}\left(((\boldsymbol{\alpha} \lll a) \vee \boldsymbol{\alpha}) \oplus (\boldsymbol{\alpha} \wedge \overline{(\boldsymbol{\alpha} \lll a)} \wedge (\boldsymbol{\alpha} \lll 2a))\right) & \text{iff } \operatorname{wt}(\boldsymbol{\alpha}) \neq n \\ n - 1 & \text{iff } \operatorname{wt}(\boldsymbol{\alpha}) = n \end{cases}.$$

Note that $d_\alpha = \dim \operatorname{Im} L_\alpha$ where

$$L_\alpha(\mathbf{x}) = ((\mathbf{x} \lll a) \wedge \boldsymbol{\alpha}) \oplus ((\boldsymbol{\alpha} \lll a) \wedge \mathbf{x}) .$$

(1), (3) and (4) follow directly from the above formula. In order to show (2), we construct three linearly independent elements in $\operatorname{Im} L_\alpha$.

Let $\operatorname{wt}(\boldsymbol{\alpha}) = 2$ with $\alpha_0 = \alpha_i = 1$. Again, w.l.o.g. let $i \leq \frac{n}{2}, i \neq 0$ since every $\boldsymbol{\alpha}$ with a Hamming weight of two is rotational equivalent to that one assumed. Now, consider the following three elements $\mathbf{x}, \mathbf{y}, \mathbf{z}$:

$$\begin{aligned}
&\mathbf{x} = (0, \ldots, 0, 1^{(a)}, 0, \ldots, 0) && \Rightarrow L_\alpha(\mathbf{x}) = (1^{(0)}, 0, \ldots, 0, \alpha_{2a}^{(a)}, 0, \ldots, 0) \\
&\mathbf{y} = (0, \ldots, 0, 1^{(a+i)}, 0, \ldots, 0) && \Rightarrow L_\alpha(\mathbf{y}) = (0, \ldots, 1^{(i)}, 0, \ldots, 0, \alpha_{i+2a}^{(i+a)}, 0, \ldots, 0) \\
&\mathbf{z} = \mathbf{1} && \Rightarrow L_\alpha(\mathbf{z}) = (\boldsymbol{\alpha} \lll a) \oplus \boldsymbol{\alpha}
\end{aligned}$$

Clearly, $L_\alpha(\mathbf{x})$ and $L_\alpha(\mathbf{y})$ are linearly independent. To show that $L_\alpha(\mathbf{z}) \notin \operatorname{span}\{L_\alpha(\mathbf{x}), L_\alpha(\mathbf{y})\}$, consider the two cases

(i) $\alpha_{i+2a} = 0$: Then $L_\alpha(\mathbf{y})_{i+a} = 0$. Since $L_\alpha(\mathbf{z})_{n-a} = 1$ and $n - a \notin \{0, i, a\}$, the linear independence follows.

(ii) $\alpha_{i+2a} = 1$: Then $i + 2a \mod n \in \{0, i\}$ because of the construction of $\boldsymbol{\alpha}$. However, since $2a \neq 0 \mod n$, it follows that $i + 2a = 0 \mod n$. Hence, $2a = n - i$. Now $2a \neq i$, because otherwise $n = 4a$ which is contradictory to $\gcd(a, n) = 1$ (since $n \geq 6$). Thus $L_\alpha(\mathbf{x})_a = 0$. In addition, $i \neq a$ because otherwise $3a = 0 \mod n$ which is also contradictory to $\gcd(a, n) = 1$. Now, $L_\alpha(\mathbf{z})_{i-a \mod n} = 1$ and $i - a \notin \{0, i, i + a\}$. □

In all cases, we thus have $p_\alpha \leq 2^{-2}$ if $\boldsymbol{\alpha} \neq \mathbf{0}$ and $p_0 = 1$. The interesting property is the fact that $p_\alpha \leq 2^{-\operatorname{wt}(\alpha)-1}$ if $\boldsymbol{\alpha}$ has a Hamming weight of 2. This is what we make use of in the following arguments. The basic idea is to guarantee enough transitions with a probability $\leq 2^{-3}$ before a zero input difference into f_S occurs (then $p_0 = 1$). This allows us to catch up the factor 2^{-2} that we lose for the zero input difference. Otherwise, if we were not able to guarantee the tighter bound described in Lemma 2 (2), the input difference into f_S of every second round might be equal to zero in the worst case and our argument would only provide the trivial bound of 2^{-T} over T rounds. See also Figure 1 for an illustration. For the formal proof, we give Corollary 1 at first. It is an implication of Lemma 1 for the SIMON-like f function.

Corollary 1. *Let for all non-zero differences $\boldsymbol{\alpha}, \boldsymbol{\beta}$ and all $t \geq 1$ the differential probability of any t-round characteristic starting with $(\mathbf{0}, \boldsymbol{\alpha})$ and ending with $(\mathbf{0}, \boldsymbol{\beta})$ be upper bounded by 2^{-2t}. Let further $p_\alpha \leq 2^{-2}$. Then,*

$$P(C) \leq 2^{-2T+2}$$

for any non-trivial T-round characteristics C with $T > 0$.

Proof. With the notation in Lemma 1, it is $p(t) = 2^{-2t}$ and $q = 2^{-2}$. Thus,

$$P(C) \leq \max_{k \leq T} p(k) q^{T-k-1} = \max_{k \leq T} 2^{-2k} 2^{-2T+2k+2} = 2^{-2T+2}.$$

□

Thus, in order to prove an upper bound on the probability of a differential characteristic of 2^{-2T+2} we only have to concentrate on t-round characteristics of the form $(\mathbf{0}, \boldsymbol{\alpha}) \to \cdots \to (\mathbf{0}, \boldsymbol{\beta})$ and prove an upper bound of 2^{-2t} for all of these. We further can restrict ourselves to the shortest characteristics of this form, e.g. $\boldsymbol{\gamma_i} \neq \mathbf{0}$ for all intermediate $\boldsymbol{\gamma_i}$. The reason is that one can easily concatenate these short characteristics to longer ones for which the property holds as well.

We have to do the analysis for a specific choice of the linear mapping θ. As a more general case, Theorem 2 formulates a sufficient condition for the argument to work. For a linear mapping $\theta : \mathbb{F}_2^n \to \mathbb{F}_2^n$, the *differential branch number* is defined as the minimum number of active bits in the differential $(\boldsymbol{\alpha} \xrightarrow{\theta} \theta(\boldsymbol{\alpha}))$, formally

$$\mathcal{B}_\theta := \min_{\alpha \neq 0} \{\mathrm{wt}(\boldsymbol{\alpha}) + \mathrm{wt}\left(\theta\left(\boldsymbol{\alpha}\right)\right)\}.$$

Theorem 2. *Let $\mathcal{B}_\theta \geq 11$. Then for any distinct a, b and any n fulfilling the properties of Lemma 2, the probability of a T-round differential characteristic is upper bounded by 2^{-2T+2}.*

Proof. Fix a t-round characteristic of the form

$$(\mathbf{0}, \boldsymbol{\alpha}) \to (\boldsymbol{\gamma_1} = \boldsymbol{\alpha}, \mathbf{0}) \to (\boldsymbol{\gamma_2}, \boldsymbol{\delta_2}) \to \cdots \to (\boldsymbol{\gamma_{t-1}}, \boldsymbol{\delta_{t-1}}) \to (\mathbf{0}, \boldsymbol{\beta})$$

with $\boldsymbol{\gamma_i} \neq \mathbf{0}$ for all $i \in \{1, \ldots, t-1\}$. Thus, we have $p_{\gamma_i} \leq 2^{-2}$ for all i. Since $\boldsymbol{\gamma_1} = \boldsymbol{\alpha}$ and $(\mathbf{0}, \boldsymbol{\alpha}) \xrightarrow{1} (\boldsymbol{\alpha}, \mathbf{0})$ holds with certainty ($p_0 = 1$), we have to show that either $p_{\gamma_i} \leq 2^{-4}$ for at least one i or that $p_{\gamma_i}, p_{\gamma_j} \leq 2^{-3}$ for at least two distinct indices i, j. In other words, one has to make sure to gain a factor of 2^{-2} within the characteristic. In order to show this, we make use of Lemma 2. If $\mathrm{wt}(\boldsymbol{\alpha}) \geq 4$, we are clearly done since $p_{\gamma_1} = p_\alpha \leq 2^{-\mathrm{wt}(\alpha)}$. We thus have to distinguish 3 cases.

(i) $\mathrm{wt}(\boldsymbol{\alpha}) = 1$: Because of the branch number, it is $\mathrm{wt}(\theta(\mathbf{x}) \oplus \theta(\mathbf{x} \oplus \boldsymbol{\alpha})) \geq 10$. Since further $\mathrm{wt}(\rho(\mathbf{x}) \oplus \rho(\mathbf{x} \oplus \boldsymbol{\alpha})) \leq 2$, we have $\mathrm{wt}(\boldsymbol{\gamma}_2) \geq 8$ and $p_{\gamma_2} \leq 2^{-4}$.
(ii) $\mathrm{wt}(\boldsymbol{\alpha}) = 2$: It is $\mathrm{wt}(\theta(\mathbf{x}) \oplus \theta(\mathbf{x} \oplus \boldsymbol{\alpha})) \geq 9$ and $\mathrm{wt}(\rho(\mathbf{x}) \oplus \rho(\mathbf{x} \oplus \boldsymbol{\alpha})) \leq 4$. Thus, $\mathrm{wt}(\boldsymbol{\gamma}_2) \geq 5$ and therefore $p_{\gamma_2} \leq 2^{-4}$.
(iii) $\mathrm{wt}(\boldsymbol{\alpha}) = 3$: We already have $p_\alpha \leq 2^{-3}$. Since $\mathrm{wt}(\theta(\mathbf{x}) \oplus \theta(\mathbf{x} \oplus \boldsymbol{\alpha})) \geq 8$ and $\mathrm{wt}(\rho(\mathbf{x}) \oplus \rho(\mathbf{x} \oplus \boldsymbol{\alpha})) \leq 6$, it is $\mathrm{wt}(\boldsymbol{\gamma}_2) \geq 2$ and therefore $p_{\gamma_2} \leq 2^{-3}$.

See also Figure 3 for the propagation of the differential Hamming weights. □

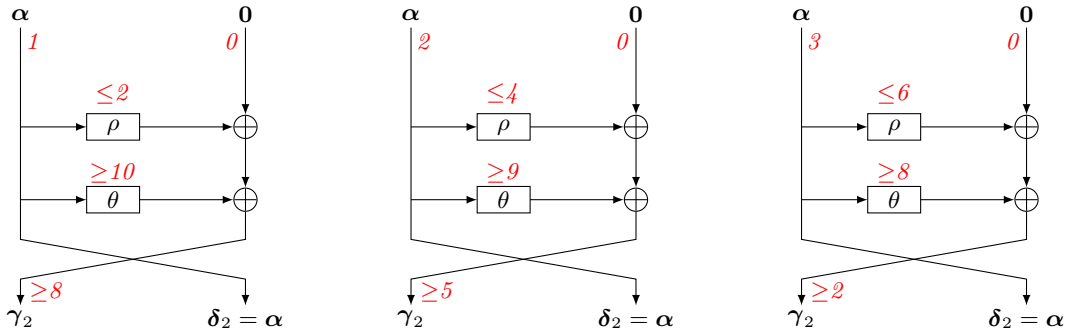

Fig. 3. Propagation of the differential Hamming weight for $\mathrm{wt}(\boldsymbol{\alpha}) \in \{1, 2, 3\}$.

We recall that θ does not have to be rotational invariant. Nevertheless, having a branch number of at least 11 is a quite restrictive property on a linear layer and in fact, for $n = 16$, there does not exist such a linear mapping. The reason is that the minimum distance d of any $[32, 16, d]$ code over $\mathbb{F}_2$ is at most 8 [19]. However, for $n \in \{24, 32, 48, 64\}$, such a linear mapping θ exists as one can also deduce from [19]. As the previous argument is more generic, we investigate the linear part of SIMON in more detail in the rest of the paper.

3.2 Obtaining the upper bound for SIMON and Simeck

In the following, we consider the linear layer $\theta(\mathbf{x}) = (\mathbf{x} \lll c)$ which has a branch number of only 2. Choosing $(8, 1, 2)$ for the rotation constants (a, b, c), we obtain the round function of SIMON. Theorem 3 states the same bound as above for all variants of SIMON. Note that the results are dependent on the specific choice of the rotation constants, but can be proven for other choices in a similar way. Of course, it does not hold for all possible a, b and c. For example, if $c = a$ or $c = b$, one obtains the trivial bound of 2^{-t} since

$$\big((1, 0, \ldots, 0)\ \mathbf{0}\big) \to \big(\mathbf{0}\ (1, 0, \ldots, 0)\big) \to \big((1, 0, \ldots, 0)\ \mathbf{0}\big)$$

would be a valid two-round iterative characteristic with probability 2^{-2}.

Theorem 3 (Bounds for Simon). *Let $n \in \{16, 24, 32, 48, 64\}$ and let $\theta(\mathbf{x}) = (\mathbf{x} \lll 2)$. For the rotation constants $a = 8, b = 1$, the probability of any T-round differential characteristic is upper bounded by 2^{-2T+2}.*

Proof. Again, fix a t-round characteristic of the form

$$(\mathbf{0}, \boldsymbol{\alpha}) \to (\boldsymbol{\gamma_1} = \boldsymbol{\alpha}, \mathbf{0}) \to (\boldsymbol{\gamma_2}, \boldsymbol{\delta_2}) \to \cdots \to (\boldsymbol{\gamma_{t-1}}, \boldsymbol{\delta_{t-1}}) \to (\mathbf{0}, \boldsymbol{\beta})$$

with $\boldsymbol{\gamma_i} \neq \mathbf{0}$ for all $i \in \{1, \ldots, t-1\}$. We have to show that either $p_{\gamma_i} \leq 2^{-4}$ for at least one i or that $p_{\gamma_i}, p_{\gamma_j} \leq 2^{-3}$ for at least two distinct indices i, j. In order to show this, Lemma 2 is used several times within this proof. Again, we have to distinguish 3 cases. Note that for simplicity with indices, we assume rotations to the right in the following. We use the $*$ symbol to indicate an unknown bit.

(i) $\mathrm{wt}(\boldsymbol{\alpha}) = 1$: Considering the rotational equivalence, let w.l.o.g.

$$\boldsymbol{\alpha} = (1, 0, \ldots, 0).$$

Recall that we get $U_\alpha = \operatorname{Im} L_\alpha \oplus f_S(\boldsymbol{\alpha})$. Since we assume

$$f_S : \mathbf{x} \mapsto (\mathbf{x} \ggg 8) \wedge (\mathbf{x} \ggg 1) \oplus (\mathbf{x} \ggg 2),$$

we obtain

$$\boldsymbol{\gamma_2} = (0, *_1, 1, 0, \quad 0, 0, 0, 0, \quad *_2, 0, 0, 0, \quad 0, 0, 0, 0 \quad \ldots) \in U_\alpha \oplus \mathbf{0}.$$

Case 1 ($*_2 = 0$): Then,

$$\boldsymbol{\gamma_3} = (1, 0, *, *, \quad 1, 0, 0, 0, \quad 0, *, *, 0, \quad 0, 0, 0, 0 \quad \ldots) \in U_{\gamma_2} \oplus \boldsymbol{\alpha},$$
$$\boldsymbol{\gamma_4} = (0, *, *^{\dagger}, *, \quad *, *, 1, 0, \quad *, 0, *, *, \quad *, 0, 0, 0 \quad \ldots) \in U_{\gamma_3} \oplus \boldsymbol{\gamma_2}.$$

If now the weight of $\boldsymbol{\gamma_4}$ is higher than 1, then $p_{\gamma_3}, p_{\gamma_4} \leq 2^{-3}$. Thus, let $\mathrm{wt}(\boldsymbol{\gamma_4}) = 1$. It follows that

$$\boldsymbol{\gamma_5} = (1, 0, *, *, \quad 1, 0, 0, *, \quad 1, *, *, 0, \quad 0, 0, *, 0 \quad \ldots) \in U_{\gamma_4} \oplus \boldsymbol{\gamma_3}$$

and thus $p_{\gamma_5} \leq 2^{-3}$.

Case 2 ($*_2 = 1$): Then $p_{\gamma_2} \leq 2^{-3}$ already holds and

$$\boldsymbol{\gamma_3} = (*^{\ddagger}, 0, *, *, \quad 1, 0, 0, 0, \quad 0, *, *, 0, \quad 0, 0, 0, 0 \quad \ldots) \in U_{\gamma_2} \oplus \boldsymbol{\alpha}.$$

Again, let w.l.o.g $\mathrm{wt}(\boldsymbol{\gamma_3}) = 1$. It follows that

$$\boldsymbol{\gamma_4} = (0, *, 1, 0, \quad 0, *, 1, 0, \quad 1, 0, 0, 0, \quad *, 0, 0, 0 \quad \ldots) \in U_{\gamma_3} \oplus \boldsymbol{\gamma_2}$$

and thus $p_{\gamma_4} \leq 2^{-3}$.

---

† This bit is only unknown if the bitlength is 16 bit ($n = 16$). Therefore, w.l.o.g. we assume this bit to be unknown. In the following, we may also consider certain bits to be unknown if the actual value does not matter for the proof.

‡ Of course, this bit is already equal to 1 if the bitlength n is greater than 16.

(ii) $\mathrm{wt}(\boldsymbol{\alpha}) = 2$: Considering the rotational equivalence, let w.l.o.g.

$$\boldsymbol{\alpha} = (1, 0, \ldots, 0, 1^{(i)}, 0, \ldots, 0)$$

with $i \leq \frac{n}{2}$. It follows that already $p_\alpha \leq 2^{-3}$.
Case 1 ($i = 1$): Then,

$$\boldsymbol{\gamma_2} = (0, *, *, 1, \quad 0, 0, 0, 0, \quad *, *, 0, 0, \quad 0, 0, 0, 0 \quad \ldots) \in U_\alpha \oplus \mathbf{0}.$$

Again, let w.l.o.g. $\mathrm{wt}(\boldsymbol{\gamma_2}) = 1$. Then,

$$\boldsymbol{\gamma_3} = (1, 1, 0, 0, \quad *, 1, 0, 0, \quad 0, 0, 0, *, \quad 0, 0, 0, 0 \quad \ldots) \in U_{\gamma_2} \oplus \boldsymbol{\alpha}$$

and thus $p_{\gamma_3} \leq 2^{-3}$.
Case 2 ($i = 4$): Then,

$$\boldsymbol{\gamma_2} = (0, *, 1, 0, \quad 0, *, 1, 0, \quad *, 0, 0, 0, \quad *, 0, 0, 0 \quad \ldots) \in U_\alpha \oplus \mathbf{0}$$

and $p_{\gamma_2} \leq 2^{-3}$.
Case 3 ($i \neq 1, i \neq 4$): Then,

$$\boldsymbol{\gamma_2} = (*, *, 1, *, \quad *, *, *, *, \quad *, *, *, *, \quad *, *, *, * \quad \ldots) \in U_\alpha \oplus \mathbf{0}.$$

Again, let w.l.o.g. $\mathrm{wt}(\boldsymbol{\gamma_2}) = 1$. Then,

$$\boldsymbol{\gamma_3} = (1, *, *, *, \quad 1, *, *, *, \quad *, *, *, *, \quad *, *, *, * \quad \ldots) \in U_{\gamma_2} \oplus \boldsymbol{\alpha}$$

and thus $p_{\gamma_3} \leq 2^{-3}$.

(iii) $\mathrm{wt}(\boldsymbol{\alpha}) = 3$: Let w.l.o.g. $\boldsymbol{\alpha} = (1, 0, \ldots, 1^{(i)}, 0, \ldots, 1^{(j)}, 0, \ldots, 0)$ with $i \geq \frac{n}{3}$ because of the rotational invariance. Again, $p_\alpha \leq 2^{-3}$. Since $n \geq 16$, it is $i \geq 6$. We distinguish the following cases:
Case 1 ($j \neq n-6, i \neq n-6$): Then,

$$\boldsymbol{\gamma_2} = (*, *, 1, *, \quad *, *, *, *, \quad \ldots \quad *, *, *, *, \quad *, *, *, *) \in U_\alpha \oplus \mathbf{0}$$

and for $\mathrm{wt}(\boldsymbol{\gamma_2}) = 1$ we obtain

$$\boldsymbol{\gamma_3} = (1, 0, 0, *, \quad 1, 0, *, *, \quad \ldots \quad *, *, *, *, \quad *, *, *, *) \in U_{\gamma_2} \oplus \boldsymbol{\alpha}$$

such that $p_{\gamma_3} \leq 2^{-3}$.
Case 2 ($i = n-6$): Then,

$$\boldsymbol{\gamma_2} = (*, *, *, *, \quad *, *, *, *, \quad \ldots \quad *, *, *, *, \quad 1, *, *, *) \in U_\alpha \oplus \mathbf{0}$$

if $j \neq n-5$ and

$$\boldsymbol{\gamma_2} = (*, *, *, *, \quad *, *, *, *, \quad \ldots \quad *, *, *, *, \quad *, 1, *, *) \in U_\alpha \oplus \mathbf{0}$$

if $j = n-5$. In both cases, for $\mathrm{wt}(\boldsymbol{\gamma_2}) = 1$ we obtain

$$\boldsymbol{\gamma_3} = (1^{(0)}, 0, 0, 0, \quad *, *, 0, 0, \quad \ldots \quad 0, 0, 1^{(i)}, *, \quad *, *, *, *) \in U_{\gamma_2} \oplus \boldsymbol{\alpha}$$

such that $p_{\gamma_3} \leq 2^{-3}$.
Case 3 ($j = n-6$): Now, we still have to consider the two possibilities $j - i \neq 6$ and $j - i = 6$. For the first case, one gets

$$\boldsymbol{\gamma_2} = (*, *, *, *, \quad *, *, *, *, \quad \ldots \quad *, *, *, *, \quad 1, *, *, *) \in U_\alpha \oplus \mathbf{0}$$

and for $\mathrm{wt}(\boldsymbol{\gamma_2}) = 1$,

$$\boldsymbol{\gamma_3} = (1, *, *, *, \quad *, *, *, *, \quad \ldots \quad *, *, *, *, \quad *, *, 1, *) \in U_{\gamma_2} \oplus \boldsymbol{\alpha}.$$

If $j - i = 6$, then,

$$\boldsymbol{\gamma_2} = (*, *, *, *, \quad \ldots \quad *, *, 1^{(i+2)}, *, \quad *, *, *, *, \quad *, *, *, *) \in U_{\alpha} \oplus \mathbf{0}$$

and for $\mathrm{wt}(\boldsymbol{\gamma_2}) = 1$,

$$\boldsymbol{\gamma_3} = (1^{(1)}, *, *, *, \quad \ldots \quad 1^{(i)}, *, *, *, \quad 1, *, 1^{(j)}, *, \quad *, *, *, *) \in U_{\gamma_2} \oplus \boldsymbol{\alpha}.$$

□

Using a similar argument, one obtains the bounds for SIMECK as the following theorem states.

Theorem 4 (Bounds for Simeck). *Let $n \in \{16, 24, 32\}$ and $\theta(\mathbf{x}) = (\mathbf{x} \lll 1)$. For the rotation constants $a = 5, b = 0$, the probability of any T-round differential characteristic is upper bounded by 2^{-2T+2}.*

Interestingly, for every instance of SIMON and SIMECK, it turns out that our approach is sufficient in order to bound the probability of differential characteristics below 2^{-2n} where n denotes the bit length of one Feistel branch. For n up to 32, the security margin κ of the corresponding primitive(s) can be considered as reasonable. See Table 1 for a comparison.

Table 1. Number of rounds needed for bounding the differential probability of a characteristic by 2^{-2n} for all instances of SIMON and SIMECK. The $\star$ symbol indicates that there is an appropriate instance of SIMECK with the same number of rounds.

	rounds	**rounds needed**	**margin κ**
SIMON32/64*	32	17	15
SIMON48/72	36	25	11
SIMON48/96*	36	25	11
SIMON64/96	42	33	9
SIMON64/128*	44	33	11
SIMON96/96	52	49	3
SIMON96/144	54	49	5
SIMON128/128	68	65	3
SIMON128/192	69	65	4
SIMON128/256	72	65	7

4 Conclusion

We presented a more general description of SIMON-like designs by separating the round function into a linear and a non-linear component and proved upper bounds on the probability of differential characteristics for specific instances. In fact, we developed a non-experimental security argument on full-round versions of SIMON that can be verified by pen and paper. We hope that this work encourages to further research on analyzing SIMON-like designs. An open question is whether our approach can be generalized in order to obtain better bounds over multiple rounds. However, as described earlier, we believe that such an argument would be much more complex. Furthermore, it would be favorable to avoid the consideration of every special case individually. This is related to the question of how to design the linear part θ in this set-up.

Acknowledgements. The author's work was supported by DFG Research Training Group GRK 1817 Ubicrypt. Special thanks go to Gregor Leander for his valuable suggestions and comments.

References

1. M. A. Abdelraheem, J. Alizadeh, H. A. Alkhzaimi, M. R. Aref, N. Bagheri, and P. Gauravaram. Improved linear cryptanalysis of reduced-round SIMON-32 and SIMON-48. In A. Biryukov and V. Goyal, editors, *Progress in Cryptology – INDOCRYPT 2015*, volume 9462 of *LNCS*, pages 153–179. Springer International Publishing, 2015.
2. F. Abed, E. List, S. Lucks, and J. Wenzel. Differential cryptanalysis of round-reduced SIMON and SPECK. In C. Cid and C. Rechberger, editors, *Fast Software Encryption*, volume 8540 of *LNCS*, pages 525–545. Springer Berlin Heidelberg, 2015.
3. J. Alizadeh, N. Bagheri, P. Gauravaram, A. Kumar, and S. K. Sanadhya. Linear cryptanalysis of round reduced SIMON. Cryptology ePrint Archive, Report 2013/663, 2013. `http://eprint.iacr.org/2013/663`.
4. H. A. Alkhzaimi and M. M. Lauridsen. Cryptanalysis of the SIMON family of block ciphers. Cryptology ePrint Archive, Report 2013/543, 2013. `http://eprint.iacr.org/2013/543`.
5. T. Ashur. Improved linear trails for the block cipher Simon. Cryptology ePrint Archive, Report 2015/285, 2015. `http://eprint.iacr.org/`.
6. R. Beaulieu, D. Shors, J. Smith, S. Treatman-Clark, B. Weeks, and L. Wingers. The SIMON and SPECK families of lightweight block ciphers. Cryptology ePrint Archive, Report 2013/404, 2013. `http://eprint.iacr.org/2013/404`.
7. R. Beaulieu, D. Shors, J. Smith, S. Treatman-Clark, B. Weeks, and L. Wingers. SIMON and SPECK: Block ciphers for the internet of things. In *NIST Lightweight Cryptography Workshop*, volume 2015, 2015.
8. C. Beierle, P. Jovanovic, M. Lauridsen, G. Leander, and C. Rechberger. Analyzing permutations for AES-like ciphers: Understanding ShiftRows. In K. Nyberg, editor, *Topics in Cryptology — CT-RSA 2015*, volume 9048 of *LNCS*, pages 37–58. Springer International Publishing, 2015.
9. R. Benadjila, O. Billet, H. Gilbert, G. Macario-Rat, T. Peyrin, M. Robshaw, and Y. Seurin. SHA-3 Proposal: ECHO, 2010. `http://crypto.rd.francetelecom.com/ECHO/`.
10. G. Bertoni, J. Daemen, M. Peeters, and G. Assche. The Keccak reference. Submission to NIST (Round 3), 13, 2011.
11. E. Biham, R. Anderson, and L. Knudsen. Serpent: A new block cipher proposal. In S. Vaudenay, editor, *Fast Software Encryption*, volume 1372 of *LNCS*, pages 222–238. Springer Berlin Heidelberg, 1998.
12. E. Biham and A. Shamir. Differential cryptanalysis of DES-like cryptosystems. In A. Menezes and S. Vanstone, editors, *Advances in Cryptology-CRYPT0' 90*, volume 537 of *LNCS*, pages 2–21. Springer Berlin Heidelberg, 1991.
13. A. Biryukov, A. Roy, and V. Velichkov. Differential analysis of block ciphers SIMON and SPECK. In C. Cid and C. Rechberger, editors, *Fast Software Encryption*, volume 8540 of *LNCS*, pages 546–570. Springer Berlin Heidelberg, 2015.
14. A. Bogdanov, L. Knudsen, G. Leander, C. Paar, A. Poschmann, M. Robshaw, Y. Seurin, and C. Vikkelsoe. PRESENT: An ultra-lightweight block cipher. In P. Paillier and I. Verbauwhede, editors, *Cryptographic Hardware and Embedded Systems - CHES 2007*, volume 4727 of *LNCS*, pages 450–466. Springer Berlin Heidelberg, 2007.
15. H. Chen and X. Wang. Improved linear hull attack on round-reduced SIMON with dynamic key-guessing techniques. Fast Software Encryption. LNCS. Springer (to appear), 2016.
16. J. Daemen. *Cipher and hash function design strategies based on linear and differential cryptanalysis*. PhD thesis, Doctoral Dissertation, March 1995, KU Leuven, 1995.
17. J. Daemen, M. Lamberger, N. Pramstaller, V. Rijmen, and F. Vercauteren. Computational aspects of the expected differential probability of 4-round AES and AES-like ciphers. *Computing*, 85(1-2):85–104, 2009.
18. J. Daemen and V. Rijmen. AES Proposal: Rjindael, 1998. `http://csrc.nist.gov/archive/aes/rijndael/Rijndael-ammended.pdf`.
19. M. Grassl. Bounds on the minimum distance of linear codes and quantum codes. Online available at `http://www.codetables.de`, 2007. Accessed on 2016-02-15.
20. K. Kondo, Y. Sasaki, and T. Iwata. On the design rationale of SIMON block cipher: Integral attacks and impossible differential attacks against SIMON variants. In M. Manulis, A. Sadeghi, and S. Schneider, editors, *Applied Cryptography and Network Security - 14th International Conference, ACNS 2016*, volume 9696 of *LNCS*, pages 518–536. Springer, 2016.
21. S. Kölbl, G. Leander, and T. Tiessen. Observations on the SIMON block cipher family. In R. Gennaro and M. Robshaw, editors, *Advances in Cryptology – CRYPTO 2015*, volume 9215 of *LNCS*, pages 161–185. Springer Berlin Heidelberg, 2015.
22. M. Matsui. Linear cryptanalysis method for DES cipher. In T. Helleseth, editor, *Advances in Cryptology — EUROCRYPT '93*, volume 765 of *LNCS*, pages 386–397. Springer Berlin Heidelberg, 1994.

23. N. Mouha, Q. Wang, D. Gu, and B. Preneel. Differential and linear cryptanalysis using mixed-integer linear programming. In C.-K. Wu, M. Yung, and D. Lin, editors, *Information Security and Cryptology*, volume 7537 of *LNCS*, pages 57–76. Springer Berlin Heidelberg, 2012.
24. K. Nyberg and L. Knudsen. Provable security against a differential attack. *Journal of Cryptology*, 8(1):27–37, 1995.
25. PUB FIPS. 197: Advanced encryption standard (AES). *National Institute of Standards and Technology*, 2001. Available online at `http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf`.
26. C. E. Shannon. Communication theory of secrecy systems. *Bell system technical journal*, 28(4):656–715, 1949.
27. T. Shirai and B. Preneel. On Feistel ciphers using optimal diffusion mappings across multiple rounds. In P. Lee, editor, *Advances in Cryptology - ASIACRYPT 2004*, volume 3329 of *LNCS*, pages 1–15. Springer Berlin Heidelberg, 2004.
28. T. Shirai and K. Shibutani. Improving immunity of Feistel ciphers against differential cryptanalysis by using multiple MDS matrices. In B. Roy and W. Meier, editors, *Fast Software Encryption*, volume 3017 of *LNCS*, pages 260–278. Springer Berlin Heidelberg, 2004.
29. S. Sun, L. Hu, P. Wang, K. Qiao, X. Ma, and L. Song. Automatic security evaluation and (related-key) differential characteristic search: Application to SIMON, PRESENT, LBlock, DES(L) and other bit-oriented block ciphers. In P. Sarkar and T. Iwata, editors, *Advances in Cryptology – ASIACRYPT 2014*, volume 8873 of *LNCS*, pages 158–178. Springer Berlin Heidelberg, 2014.
30. Y. Todo and M. Morii. Bit-based division property and application to Simon family. Fast Software Encryption. LNCS. Springer (to appear), 2016.
31. N. Wang, X. Wang, K. Jia, and J. Zhao. Differential attacks on reduced SIMON versions with dynamic key-guessing techniques. Cryptology ePrint Archive, Report 2014/448, 2014. `http://eprint.iacr.org/2014/448`.
32. Q. Wang, Z. Liu, K. Varıcı, Y. Sasaki, V. Rijmen, and Y. Todo. Cryptanalysis of reduced-round SIMON32 and SIMON48. In W. Meier and D. Mukhopadhyay, editors, *Progress in Cryptology – INDOCRYPT 2014*, volume 8885 of *LNCS*, pages 143–160. Springer International Publishing, 2014.
33. G. Yang, B. Zhu, V. Suder, M. Aagaard, and G. Gong. The Simeck family of lightweight block ciphers. In T. Güneysu and H. Handschuh, editors, *Cryptographic Hardware and Embedded Systems – CHES 2015*, volume 9293 of *LNCS*, pages 307–329. Springer Berlin Heidelberg, 2015.