Observations on the SIMON block cipher family

Stefan Kölbl[1], Gregor Leander[2], and Tyge Tiessen[1]
stek@dtu.dk, gregor.leander@rub.de, tyti@dtu.dk

[1] DTU Compute, Technical University of Denmark, Denmark
[2] Horst Görtz University for IT Security, Ruhr-Universität Bochum, Germany

Abstract. In this paper we analyze the general class of functions underlying the SIMON block cipher. In particular, we derive efficiently computable and easy to implement expressions for the exact differential and linear behavior of SIMON-like round function. Using those expressions we investigate a large set of natural SIMON variants with respect to the most important cryptographic criteria. Interestingly, the NSA's choice for the parameters are not always optimal.
Using a computer aided approach based on SAT/SMT solvers we are able to find both the optimal differential and linear characteristics for variants of SIMON and can also give better estimates on the probability of differentials. As a result of this analysis we propose different sets of rotation constants, which feature better properties on some criteria, and might be interesting for further analysis.

Keywords: SIMON, block cipher, differential cryptanalysis, linear cryptanalysis, boolean functions

1 Introduction

In the last years a large number of new block ciphers has been designed and published. Most of those proposals were lightweight block ciphers, optimized with respect to chip-area, but others performance metrics (such as latency [1], code-size [2] and ease of side-channel protection [3]) has been taken into account as well. In this context it should also be noted that some of those criteria were already treated in NOEKEON [4]. Along with many new proposals, a large variety of papers that analyze and cryptanalyze the new design approaches have been published.

The later results demonstrate that, even so not all proposals are of significant interest in practice, new designs often open up for new fundamental insights within the field of block-ciphers. In this sense the area of lightweight ciphers also increased our fundamental understanding of block ciphers. Thus, the area of lightweight cryptography has been a very fruitful driving force in the field of block cipher design.

The importance of lightweight cryptography and it's applications is also reflected by the NSA publishing the SIMON and SPECK families of lightweight ciphers in 2013 [5]. Given that it is only the third time within four decades that the NSA publishes a block cipher, this is a remarkable situation. Moreover, as NIST started shortly afterwards to investigate the possibilities to standardize lightweight primitives, SIMON and SPECK certainly deserve a careful investigation. This is even more true as only the designs but no analysis or explanation of the design choices were published by NSA. In comparison to what one would expect from publicly available cryptographic primitives nowadays, this lack of openness necessarily gives rise to curiosity and caution.

Our Contribution

In this paper we focus on the SIMON family of ciphers; a very elegant, innovative and extremely efficient set of block ciphers.

There is already a large variety of papers, mainly focusing on linear and differential attacks on SIMON published. Most of the methods used therein are rather ad-hoc and very specific to the particular (and unexplained) parameter set of SIMON. Here we complement those works by a comprehensive studies of the underlying functions in SIMON. In particular, given fast dependency as well as resistance against linear- and differential cryptanalysis as the major criteria, we investigate which parameters would constitute the optimal choices.

As a basis for our goal to understand both the security of SIMON as well as the choice of its parameter set, we rigorously derive formulas for the the differential probabilities and the linear square correlations of the SIMON-like round function that can be evaluated in constant time and time linear in the word size respectively. More precisely, we study differential probabilities and linear correlations of functions of the form

$$S^a(x) \odot S^b(x) + S^c(x)$$

where $S^i(x)$ corresponds to a cyclic left shift of x and $\odot$ denotes the bitwise AND operation.

We achieve this goal by first simplifying this question by considering equivalent descriptions both of the round function as well as the whole cipher (cf. Section 2.4). These simplifications, together with the theory of quadratic boolean functions, result in a clearer analysis of linear and differential properties (cf. Sections 3 and 4). Importantly, the derived simple equations for computing the probabilities of the SIMON round function can be evaluated efficiently and, more importantly maybe, are conceptual very easy. This allows them to be easily used in computer-aided investigations of differential and linear properties over more rounds. It should be noted here that the expression for linear approximations is more complex than the expression for the differential case. However, with respect to the running time of the computer-aided investigations this difference is negligible.

We used this to implement a framework based on SAT and SMT solvers to find the provably best differential and linear characteristics for various instantiations of SIMON (cf. Section 5, in particular Table 1). Furthermore we are able to shed light on how differentials in SIMON profit from the collapse of many differential characteristics by giving exact distributions of the probabilities of these characteristics for chosen differentials. The framework is open source and publicly available to encourage further research[3].

In Section 6 we apply the developed theory and tools to investigate the design space of SIMON-like functions. In particular, using the computer-aided approach, we find that the standard SIMON parameters are not optimal with regard to the best differential and linear characteristics.

As a side result, we improve the probabilities for the best known differentials for several variants and rounds of SIMON. While this might well lead to (slightly) improved attacks, those improved attacks are out of the scope of our work.

Interestingly, at least for SIMON32 our findings indicate that the choices made by the NSA are good but not optimal under our metrics, leaving room for further investigations

[3] In order to respect anonymity, we do not give a link to the source code here.

and questions. To encourage further research, we propose several alternative parameter choices for SIMON32. Here, we are using the parameters that are optimal when restricting the criteria to linear, differential and dependency properties. We encourage further research on those alternative choices to shed more light on the undisclosed design criteria.

We also like to point out that the SIMON key-scheduling was not part of our investigations. It's influence on the security of SIMON is left as an important open questions for further investigations. In line with this, whenever we investigate multi-round properties of SIMON in our work, we implicitly assume independent round keys in the computation of probabilities.

Finally, we note that most of our results can be applied to more general constructions, where the involved operations are restricted to AND, XOR, and rotations.

Related Work

There are various papers published on the cryptanalysis of SIMON [6–12]. The most promising attacks so far are based on differential and linear cryptanalysis, however a clear methodology of how to derive the differential probabilities and square correlations seems to miss in most cases. Biryukov et al.[7] derive a correct, but rather involved method to find the differential probabilities. Abed et al.[12] state an algorithm for the calculation of the differential probabilities but without further explanation. For the calculation of the square correlations an algorithm seems to be missing all together.

Previous work also identifies various properties like the strong differential effect and give estimate of the probability of differentials.

The concept behind our framework was previously also applied on the ARX cipher Salsa20 [13] and the CAESAR candidate NORX [14]. In addition to the applications proposed in previous work we extend it for linear cryptanalysis, examine the influence of rotation constants and use it to compute the distribution of characteristics corresponding to a differential.

2 Preliminaries

In this section, we start by defining our notations and giving a short description of the round function. Afterwards we recall some generalities about suitable notations of equivalence of Boolean functions. We herby focus on equivalences that allow to simplify the investigations on SIMON-like round functions. Those observations will be used in the following sections. Most of the following is generally applicable to any AND-RX construction i.e., a construction that make only use of the operations AND, XOR, and rotations.

2.1 Notation

We denote by $\mathbb{F}_2$ the field with two elements and by $\mathbb{F}_2^n$ the n dimensional vector space over $\mathbb{F}_2$. By $\mathbf{0}$ and $\mathbf{1}$ we denote the vectors of $\mathbb{F}_2^n$ with all 0s and all 1s respectively.

The addition in $\mathbb{F}_2^n$ i.e., bit-wise XOR, is denoted by $+$. By $\odot$ we denote the AND operation in $\mathbb{F}_2^n$ i.e., multiplication over $\mathbb{F}_2$ in each coordinate:

$$x \odot y = (x_i y_i)_i.$$

By $\vee$ we denote the bitwise OR operation. By $\overline{x}$ we denote the bitwise negation of x i.e., $\overline{x} := (x + \mathbf{1})$. $\mathrm{wt}(a)$ denotes the Hamming weight of a vector $a \in \mathbb{F}_2^n$. For $x \in \mathbb{F}_2^n$,

$$S^i : \mathbb{F}_2^n \to \mathbb{F}_2^n$$

denotes the left circular shift by i positions. We also note that any arithmetic of bit indices is always done modulo the word size n.

In this paper we are mainly concerned with functions of the form

$$f_{a,b,c}(x) = S^a(x) \odot S^b(x) + S^c(x) \tag{1}$$

and we identify such functions with its triple (a, b, c) of parameters.

Furthermore, $\mathsf{Dom}(f)$ is the domain of a function f, $\mathsf{Img}(f)$ is its image. By $\mathbb{Z}_n$ we denote the integers modulo n.

For a vectorial Boolean function on n bits $f : \mathbb{F}_2^n \to \mathbb{F}_2^n$, we denote by

$$\widehat{f}(\alpha, \beta) = \sum_x \mu\left(\langle \beta, f \rangle + \langle \alpha, x \rangle\right)$$

the Walsh (or Fourier) Coefficient with input mask α and output mask β . Here we use

$$\mu(x) = (-1)^x$$

to simplify notation.

The corresponding squared correlation of f is given by

$$C^2(\alpha \to \beta) = \left(\frac{\widehat{f}(\alpha, \beta)}{2^n}\right)^2.$$

Similarly, for differentials we denote by $P(\alpha \to \beta)$ the probability that a given input difference α results in a given output difference β, i.e.

$$P(\alpha \to \beta) = \frac{|\{x \mid f(x) + f(x + \alpha) = \beta\}|}{2^n}.$$

2.2 Description of SIMON

SIMON is a family of lightweight block ciphers with block sizes 32, 48, 64, 96, and 128 bits. The constructions are Feistel ciphers using a word size n of 16, 24, 32, 48 or 64 bits, respectively. We will denote the variants as SIMON2n. The key size varies between of 2, 3, or 4 n-bit words. The round function of SIMON is composed of AND, rotation, and XOR operations on the complete word (see figure 1). More precisely, the round function if SIMON corresponds to

$$S^8(x) \odot S^1(x) + S^2(x),$$

that is to the parameters $(8, 1, 2)$ for f as given in Equation (1). As we are not only interested in the original SIMON parameters, but in investigating the entire design space of SIMON-like functions, we denote by

$$\mathrm{SIMON}[a, b, c]$$

the variant of SIMON where the original round function is replaced by $f_{a,b,c}$ (cf. Equation (1)).

As it is out of scope for our purpose, we refer to [5] for the description of the key-scheduling.

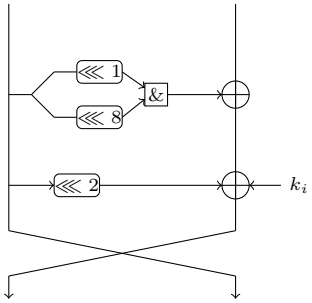

Fig. 1. SIMON round function

2.3 Affine equivalence of Boolean Functions

Given two (vectorial) Boolean functions f_1 and f_2 on $\mathbb{F}_2^n$ related by

$$f_1(x) = (A \circ f_2 \circ B)(x) + C(x)$$

where A and B are affine permutations and C is an arbitrary affine mapping on $\mathbb{F}_2^n$ we say that f_1 and f_2 are *extended affine equivalent* (cf. [15] for a comprehensive survey).

With respect to differential cryptanalysis, if f_1 and f_2 are affine equivalent then the f_1-differential

$$\alpha \xrightarrow{f_1} \beta$$

has probability p_1 if and only if the f_2-differential

$$B(\alpha) \xrightarrow{f_2} A^{-1}(\beta + C(\alpha))$$

has probability p_1.

For linear cryptanalysis, a similar relation holds for the linear correlation. If f_1 and f_2 are related as defined above, it holds that

$$\widehat{f_1}(\alpha, \beta) = \widehat{f_2}\left(\left(C \circ B^{-1}\right)^T \beta + \left(B^{-1}\right)^T \alpha, A^T \beta\right).$$

Thus, up to linear changes we can study f_2 instead of f_1 directly. Note that, for an actual attack, these changes are usually critical and can certainly not be ignored. However, tracing the changes is, again, simple linear algebra.

This means that for differential and linear properties of SIMON-like functions of the form

$$f_{a,b,c}(x) = S^a(x) \odot S^b(x) + S^c(x)$$

it is sufficient to look at the simplified variant

$$f_{0,b-a,0}(x) = x \odot S^{b-a}(x).$$

Using linear algebra the results can simply be transferred to the original function.

2.4 Structural Equivalence Classes in AND-RX Constructions

AND-RX constructions i.e., constructions that make only use of the operations AND ($\odot$), XOR ($+$), and rotations (S^r), exhibit a high degree of symmetry. Not only are they invariant under rotation of all input words, output words and constants, they are furthermore structurally invariant under any linear transformation of the bit-indices. As a consequence of this, several equivalent representations of the SIMON variants exist.

Let T be a permutation of the bits of an n-bit word that corresponds to a linear transformation of the bit-indices. Thus there are $a \in \mathbb{Z}_n^*$ and $b \in \mathbb{Z}_n$ such that bit i is renamed to $a \cdot i + b$. As the AND and XOR operations are bitwise, T clearly commutes with these:

$$Tv \odot Tw = T(v \odot w)$$
$$Tv + Tw = T(v + w)$$

where v and w are n-bit words. A rotation to the left by r can be written bitwise as

$$S^r(v)_i = v_{i-r}.$$

We thus get the following bitwise relation after transformation with T

$$S^r(v)_{a\cdot i+b} = v_{a\cdot(i-r)+b} = v_{a\cdot i+b-a\cdot r}.$$

Substituting $a \cdot i + b$ with j this is the same as

$$S^r(v)_j = v_{j-a\cdot r}.$$

Thus the rotation by r has been changed to a rotation by $a \cdot r$. Thus we can write

$$TS^r v = S^{a\cdot r} Tv.$$

Commuting the linear transformation of the bit-indices with a rotation thus only changes the rotation constant by a factor. In the special case where all input words, output words and constants are rotated, which corresponds to the case $a = 1$, the rotation constant are left untouched.

To summarize the above, when applying such a transformation T to all input words, output words and constants in an AND-RX construction, the structure of the constructions remains untouched apart from a multiplication of the rotation constants by the factor a.

This means for example for Simon32 that changing the rotation constants from $(8, 1, 2)$ to $(3 \cdot 8, 3 \cdot 1, 3 \cdot 2) = (8, 3, 6)$ and adapting the key schedule accordingly gives us the same cipher apart from a bit permutation. As a has to be coprime to n, all a with $\gcd(a, n) = 1$ are allowed, giving $n/\varphi(a, n)$ equivalent sets of rotation constants in each equivalence class where φ is Euler's phi function.

Together with the result from section 2.3, this implies the following lemma.

Lemma 1. *Any function $f_{a,b,c}$ as defined in Equation (1) is extended affine equivalent to a function*

$$f_{0,d,0} = x \odot S^d(x)$$

where $d|n$ or $d = 0$.

This means that, when looking at differential and square correlations of SIMON-like round functions, it is sufficient to investigate this restricted set of functions. The results from these cases can then simply be transferred to the general case which in turn can then be used to determine the differential probabilities and square correlations of many-rounds characteristics.

3 Differential Probabilities of SIMON-like round functions

In this section, we derive a closed expression for the differential probability for all SIMON-like round functions i.e., all functions described in Equation (1). The main ingredients here are the derived equivalences and the observation that any such function is quadratic. Being quadratic immediately implies that its derivative is linear and thus the computation of differential probabilities basically boils down to linear algebra (cf. Theorem 1). However, to be able to efficiently study multiple round properties and in particular differential characteristics, it is important to have a simple expression for the differential probabilities. Those expressions are given for $f_{0,1,0}$ in Theorem 2 and for the general case in Theorem 3.

3.1 A closed expression for the differential probability

The following statement summarizes the differential properties of the f function.

Theorem 1. *Given an input difference α and an output difference β the probability p of the corresponding differential (characteristic) for the function $f(x) = x \odot S^a(x)$ is given by*

$$p_{\alpha,\beta} = \begin{cases} 2^{-d} & \text{if } \beta + \alpha \odot S^a(\alpha) \in \mathsf{Img}(L_\alpha) \\ 0 & \text{else} \end{cases}$$

where

$$d = \dim \ker(L_\alpha)$$

and

$$L_\alpha(x) = x \odot S^a(\alpha) + \alpha \odot S^a(x)$$

Proof. We have to count the number of solutions to the equation

$$f(x) + f(x + \alpha) = \beta.$$

This simplifies to

$$L_\alpha(x) = x \odot S^a(\alpha) + \alpha \odot S^a(x) = \beta + \alpha \odot S^a(\alpha)$$

As this is an affine equation, it either has zero solutions or the number of solutions equals the kernel size i.e., the number of elements in the subspace

$$\{x \mid x \odot S^a(\alpha) + \alpha \odot S^a(x) = \mathbf{0}\}.$$

Clearly, the equation has solutions if and only if $\beta + \alpha \odot S^a(\alpha)$ is in the image of L_α. □

Next, we present a closed formula to calculate the differential probability in the case where $a = 1$. Furthermore we restrict ourselves to the case where n is even.

Theorem 2. *Let*

$$\texttt{varibits} = S^1(\alpha) \vee \alpha$$

and

$$\texttt{doublebits} = \alpha \odot \overline{S^1(\alpha)} \odot S^2(\alpha).$$

Then the probability that difference α goes to difference β is

$$P(\alpha \to \beta) = \begin{cases} 2^{-n+1} & \text{if } \alpha = \mathbf{1} \text{ and } \mathrm{wt}(\beta) \equiv 0 \mod 2 \\ 2^{-\mathrm{wt}(\texttt{varibits}+\texttt{doublebits})} & \text{if } \alpha \neq \mathbf{1} \text{ and } \beta \odot \overline{\texttt{varibits}} = \mathbf{0} \\ & \text{and } (\beta + S^1(\beta)) \odot \texttt{doublebits} = \mathbf{0} \\ 0 & \text{else} \end{cases}$$

Proof. According to theorem 1, we need to proof two things. Firstly we need to proof that the rank of L_α is $n-1$ when $\alpha = \mathbf{1}$, and $\mathrm{wt}(\texttt{varibits} + \texttt{doublebits}))$ otherwise. Secondly we need to proof that $\beta + \alpha \odot S^a(\alpha) \in \mathsf{Img}(L_\alpha)$ iff $\mathrm{wt}(\beta) \equiv 0 \mod 2$ when $\alpha = \mathbf{1}$, and $\beta \odot \texttt{varibits} = \mathbf{0}$ and $(\beta + S^1(\beta)) \odot \texttt{doublebits} = \mathbf{0}$ when $\alpha \neq \mathbf{1}$.

We first consider the first part. Let us write $L_\alpha(x)$ in matrix form. $S^1(\alpha) \odot x$ can be written as $M_{S^1(\alpha)\odot}x$ with

$$M_{S^1(\alpha)\odot} = \begin{pmatrix} \alpha_{n-1} & \dots & \dots & 0 \\ \vdots & \alpha_0 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & \alpha_{n-2} \end{pmatrix}. \tag{2}$$

Equivalently we can write $\alpha \odot x$ and $S^1(x)$ with matrices as $M_{\alpha\odot}x$ and $M_{S^1}x$ respectively where

$$M_{\alpha\odot} = \begin{pmatrix} \alpha_0 & \dots & \dots & 0 \\ \vdots & \alpha_1 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & \alpha_{n-1} \end{pmatrix} \quad M_{S^1} = \begin{pmatrix} 0_{1,n-1} & I_{1,1} \\ I_{n-1,n-1} & 0_{n-1,1} \end{pmatrix} \tag{3}$$

i.e., M_{S^1} consist of two identity and two zero submatrices. The result of $M_{S^1(\alpha)\odot} + M_{\alpha\odot}M_{S^1}$ can now be written as

$$\begin{pmatrix} \alpha_{n-1} & 0 & 0 & \dots & \alpha_0 \\ \alpha_1 & \alpha_0 & 0 & \dots & 0 \\ 0 & \alpha_2 & \alpha_1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \alpha_{n-1} & \alpha_{n-2} \end{pmatrix} \tag{4}$$

Clearly the rank of the matrix is $n-1$ when all α_i are 1. Suppose now that not all α_i are 1. In that case, a set of non-zero rows is linearly dependent iff there exist two identical

rows in the set. Thus to calculate the rank of the matrix, we need to calculate the number of unique non-zero rows.

By associating the rows in the above matrix with the bits in `varibits`, we can clearly see that the number of non-zero rows in the matrices corresponds to the number of 1s in `varibits`.

To count the number of non-unique rows, first notice that a row can only be identical to the row exactly above or below. Suppose now that a non-zero row i is identical to the row $i-1$ above. Then α_{i-1} has to be 0 while α_i and α_{i-2} have to be 1. But then row α_{i-1} cannot be simultaneously be identical to row $i-2$. Thus it is sufficient to calculate the number of non-zero rows minus the number of rows that are identical to the row above, to find the rank of the matrix. Using the above observations, we need to calculate the number of rows i with

$$\alpha_i\alpha_{i-1} + \alpha_i\overline{\alpha_{i-1}}\alpha_{i-2}.$$

This corresponds to calculating $\mathrm{wt}(\texttt{varibits} + \texttt{doublebits})$.

For the second part of the proof, first notice that $\alpha \odot S^1(\alpha)$ is in the image of L_α (consider for x the vector with bits alternately set to 0 and 1). Thus it is sufficient to test whether β is in $\mathsf{Img}L_\alpha$. Let $y = L_\alpha(x)$. In the case of $\alpha = \mathbf{1}$, we can deduce from bit y_i whether $x_i = x_{i-1}$ or $x_i \neq x_{i-1}$. Thus the bits in y create a chain of equations/inequations for x which can be fulfilled iff there is an even number of inequations. Hence $\beta \in \mathsf{Img}L_\alpha$ iff $\mathrm{wt}(\beta) \equiv 0 \mod 2$.

For the case that $\alpha \neq \mathbf{1}$, we first note that y_i has to be zero if row zero in equation (4) is zero. Thus following our discussion of the matrix earlier, we see that y_i is independent of the rest of y if the corresponding row is linearly independent and $y_i = y_j$ if the corresponding rows are identical. Thus y is in the image of L_α iff $y \odot \overline{\texttt{varibits}} = \mathbf{0}$ and $(y + S^1(y)) \odot \texttt{doublebits} = \mathbf{0}$. □

3.2 The full formula for differentials.

Above we treated only the case for $a = 1$, $b = 0$, and $c = 0$. As mentioned earlier, the general case where $\gcd(a-b, n) = 1$ can be deduced from this with linear algebra. When $\gcd(d, n) \neq 1$ though, the function $f(x) = x \odot S^d(x)$ partitions the output bits into independent classes. This not only raises differential probabilities (worst case $d = 0$), it also makes the the notation for the formulas more complex and cumbersome, though not difficult. We thus restrict ourselves to the most important case when $\gcd(a-b, n) = 1$. The general formulas are then

Theorem 3. *Let $f(x) = S^a(x) \odot S^b(x) + S^c(x)$, where $\gcd(n, a-b) = 1$, n even, and $a \geq b$ and let α and β be an input and an output difference where not all bits in α are set. Then with*

$$\texttt{varibits} = S^a(\alpha) \vee S^b(\alpha)$$

and

$$\texttt{doublebits} = S^b(\alpha) \odot \overline{S^a(\alpha)} \odot S^{2a-b}(\alpha)$$

and

$$\gamma = \beta + S^c(\alpha)$$

we have that the probability that difference α goes to difference β is

$$P(\alpha \to \beta) = \begin{cases} 2^{-n+1} & \text{if } \alpha = \mathbf{1} \text{ and } \mathrm{wt}(\gamma) \equiv 0 \mod 2 \\ 2^{-\mathrm{wt}(\mathtt{varibits}+\mathtt{doublebits})} & \text{if } \alpha \neq \mathbf{1} \text{ and } \gamma \odot \overline{\mathtt{varibits}} = \mathbf{0} \\ & \text{and } (\gamma + S^{a-b}(\gamma)) \odot \mathtt{doublebits} = \mathbf{0} \\ 0 & \text{else} \end{cases}$$

4 Linear Correlations of SIMON-like round functions

As in the differential case, for the study of linear approximations, we also build up on the results from subsections 2.3 and 2.4. We will thus start with studying linear approximations for the function $f(x) = x \odot S^a(x)$. Again, the key point here is that all those functions are quadratic and thus their Fourier coefficient, or - equivalently - their correlation, can be computed by linear algebra (cf Theorem 4). Theorem 5 is then, in analogy to the differential case, the explicit expression for the linear correlations. It basically corresponds to an explicit formula for the dimension of the involved subspace.

The first result is the following:

Theorem 4.

$$\widehat{f}(\alpha,\beta)^2 = \begin{cases} 2^{n+d} & \text{if } \alpha \in U_\beta^\perp \\ 0 & \text{else} \end{cases}$$

where

$$d = \dim U_\beta$$

and

$$U_\beta = \{y \mid \beta \odot S^a(y) + S^{-a}(\beta \odot y) = \mathbf{0}\}$$

Proof. We compute

$$\begin{aligned} \widehat{f}(\alpha,\beta)^2 &= \sum_{x,y} \mu\left(\langle \beta, f(x) + f(y)\rangle + \langle \alpha, x+y\rangle\right) \\ &= \sum_{x,y} \mu\left(\langle \beta, f(x) + f(x+y)\rangle + \langle \alpha, y\rangle\right) \\ &= \sum_{x,y} \mu\left(\langle \beta, x \odot S^a(x) + (x+y) \odot S^a(x+y)\rangle + \langle \alpha, y\rangle\right) \\ &= \sum_{y} \mu\left(\langle \beta, f(y)\rangle + \langle \alpha, y\rangle\right) \sum_{x} \mu\left(\langle \beta, x \odot S^a(y) + y \odot S^a(x)\rangle\right) \\ &= \sum_{y} \mu\left(\langle \beta, f(y)\rangle + \langle \alpha, y\rangle\right) \sum_{x} \mu\left(\langle x, \beta \odot S^a(y) + S^{-a}(\beta \odot y)\rangle\right) \end{aligned}$$

Now, for the sum over x only two outcomes are possible, 2^n or zero. More precisely, it holds that

$$\sum_x \mu\left(\langle x, \beta \odot S^a(y) + S^{-a}(\beta \odot y)\rangle\right) = \begin{cases} 2^n & \text{if } \beta \odot S^a(y) + S^{-a}(\beta \odot y) = 0 \\ 0 & \text{else} \quad . \end{cases}$$

Thus, defining

$$U_\beta = \{y \mid \beta \odot S^a(y) + S^{-a}(\beta \odot y) = 0\}$$

we get

$$\widehat{f}(\alpha,\beta)^2 = 2^n \sum_{y \in U_\beta} \mu\left(\langle \beta, f(y)\rangle + \langle \alpha, y\rangle\right).$$

Now, as U is the radical of f_β, the function f_β restricted to U is linear. Moreover, as f_β is unbalanced for all β, it follows that actually f_β is constant zero on U_β. We thus conclude that

$$\widehat{f}(\alpha,\beta)^2 = 2^n \sum_{y \in U_\beta} \mu\left(\langle \alpha, y\rangle\right).$$

With a similar argument as above, it follows that $\widehat{f}(\alpha,\beta)^2$ is non-zero if and only if α is contained in $U_\beta^\perp$. □

Let us now restrict ourselves to the case where $f(x) = x \odot S^1(x)$. The general case can be deduced analogously to the differential probabilities. For simplicity we also restrict ourselves to the case where n is even.

First we need to introduce some notation. Let $x \in \mathbb{F}_2^n$ with not all bits equal to 1. We now look at blocks of consecutive 1s in x, including potentially a block that "wraps around" the ends of x. Let the lengths of these blocks, measured in bits, be denoted as $c_0, \ldots, c_m$. For example, the bitstring 100101111011 has blocks of length 1, 3, and 4. With this notation define $\theta(x) := \sum_{i=0}^{m} \lceil \frac{c_i}{2} \rceil$.

Noting that the linear square correlation of f is $\frac{\widehat{f}(\alpha,\beta)^2}{2^{2n}}$, we then have the following theorem:

Theorem 5. *With the notation from above it holds that the linear square correlation of $\alpha \xrightarrow{f} \beta$ can be calculated as*

$$C(\alpha \to \beta) = \begin{cases} 2^{-n+2} & \text{if } \beta = \mathbf{1} \text{ and } \alpha \in U_\beta^\perp \\ 2^{-\theta(\beta))} & \text{if } \beta \neq \mathbf{1} \text{ and } \alpha \in U_\beta^\perp \\ 0 & \text{else.} \end{cases}$$

Proof. Define $L_\beta(x) := \beta \odot S^1(x) + S^{-1}(\beta \odot x)$. Clearly L_β is linear. Also $U_\beta = \ker L_\beta(x)$. Let us determine the rank of this mapping. Define the matrices $M_{\beta\cdot}$, M_{S^1}, and $M_{S^{-1}}$ as

$$M_{\beta\cdot} = \begin{pmatrix} \beta_0 & \cdots & \cdots & 0 \\ \vdots & \beta_1 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \beta_{n-1} \end{pmatrix} \quad M_{S^1} = \begin{pmatrix} 0_{1,n-1} & I_{1,1} \\ I_{n-1,n-1} & 0_{n-1,1} \end{pmatrix} \quad M_{S^{-1}} = \begin{pmatrix} 0_{n-1,1} & I_{n-1,n-1} \\ I_{1,1} & 0_{1,n-1} \end{pmatrix} \tag{5}$$

We can then write L_β in matrix form as

$$\begin{pmatrix} 0 & \beta_1 & 0 & \dots & 0 & \beta_0 \\ \beta_1 & 0 & \beta_2 & 0 & \dots & 0 \\ 0 & \beta_2 & 0 & \beta_3 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \ddots & 0 & \beta_{n-1} \\ \beta_0 & 0 & \dots & 0 & \beta_{n-1} & 0 \end{pmatrix} \tag{6}$$

Clearly, if β is all 1s, the rank of the matrix is $n-2$ as n is even.[4] Let us therefore now assume that β is not all 1s. When we look at a block of 1s in β e.g., $\beta_{i-1} = 0$, $\beta_i, \beta_{i+1}, \dots, \beta_{i+l-1} = 1$, and $\beta_l = 0$. Then clearly the l rows are linearly independent when l is even. When l is odd though, the sum of rows i, $i+2$, $i+4$, up to row $i+l-3$ will equal row $i+l-1$. In that case there are thus only $l-1$ linearly independent rows. As the blocks of 1s in β generate independent blocks of rows, we can summarize that the rank of the matrix is exactly $\theta(\beta)$. □

Analogously to the differential probabilities, the linear probabilities in the general case can be derived from this. It is likewise straightforward to derive how to determine whether $\alpha \in U_\beta^+$. As an explicit formulations this is rather tedious, we instead refer to the implementation in Python given in the Appendix A where both is achieved in the case where $\gcd(a-b, n) = 1$ and n is even.

5 Finding Optimal Differential and Linear Characteristics

While there are various methods to find good characteristics, there has been little progress on finding optimal differential characteristics. The formulas derived for both differential and linear probabilities enable us to apply an algebraic approach to finding the best characteristics. A similar technique has been applied to the ARX cipher Salsa20 [13] and the CAESAR candidate NORX [14]. For finding the optimal characteristics for SIMON we implemented an open source tool based on SAT/SMT solvers.

In the next section we will show how SIMON can be modeled to find both the best differential and linear characteristics in this framework and how this can be used to solve cryptanalytic problems.

5.1 Model for Differential Cryptanalysis of SIMON

First, we define the variables used in the model of SIMON. We use two n-bit variables x_i, y_i to represent the XOR-difference in the left and right half of the state for each round and an additional variable z_i to store the XOR-difference of the output of the AND operation.

For computing the probability of the characteristic we introduce an additional variable w_i for each round. The sum over all weights w_i then gives the absolute value of the logarithm of the probability of a differential characteristic. The individual w_i are computed according to theorem 3 as

[4] The rank is $n-1$ when n is odd.

$$w_i = \mathrm{wt}(\texttt{varibits} + \texttt{doublebits}) \tag{7}$$

where $\texttt{wt}(x)$ is the Hamming weight of x and

$$\begin{aligned} \texttt{varibits} &= (x_i \lll a) \vee (x_i \lll b) \\ \texttt{doublebits} &= (x_i \lll b) \wedge \neg(x_i \lll a) \wedge (x_i \lll (2a - b)) \end{aligned}$$

Therefore, for one round of `SIMON` we get the following set of constraints:

$$\begin{aligned} y_{i+1} &= x_i \\ 0 &= (z_i \wedge \texttt{varibits}) \\ 0 &= (z_i + (z_i \lll (a - b))) \wedge \texttt{doublebits} \\ x_{i+1} &= y_i + z_i + (x_i \lll c) \\ w_i &= \mathrm{wt}(\texttt{varibits} + \texttt{doublebits}) \end{aligned} \tag{8}$$

A model for linear characteristics, though slightly more complex, can be implemented in a similar way. A description of this model can be found in the implementation of our framework. Despite the increase in complexity, we could not observe any significant impact on the solving time for the linear model.

5.2 Finding Optimal Characteristics

We can now use the previous model for SIMON to search for optimal differential characteristics. This is done by formulating the problem of finding a valid characteristic, with respect to our constraints, for a given sum of weights w_i. This is important to limit the search space and also we are more interested in differential characteristics with a low weight resp. high probability as they are more promising to lead to attacks with a lower complexity. Therefore, we start with a low weight and check if a characteristic with the respective probability exists. If not we increase the weight. The procedure can be described in the following way:

- For each round of the cipher add the corresponding constraints as defined in (8). This system of constraints then exactly describes the form of a valid characteristic for the given parameters.
- Add a condition which accumulates the weights of each round as defined in (7) and check if it is equal to our target weight w_i.
- Query if there exists an assignment of variables which is satisfiable under the constraints.
- Increment the weight w_i and repeat the procedure.

One of the main advantages compared to other approaches is that we can proof a lower bound on the weight of characteristics for a given cipher and number of rounds. If the solvers determines the set of conditions unsatisfiable, we know that no characteristic with the specified weight exists. We used this approach to determine the characteristics with minimal weight for different variants of SIMON. The results are given in Table 1.

Table 1. Overview of the optimal differential (on top) and linear characteristics for different variants of SIMON. The probabilities are given as $\log_2(p)$.

Rounds:	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Differential															
SIMON32	-2	-4	-6	-8	-12	-14	-18	-20	-25	-30	-34	-36	-38	-40	-42
SIMON48	-2	-4	-6	-8	-12	-14	-18	-20	-26	-30	-35	-38	-44	-46	-50
SIMON64	-2	-4	-6	-8	-12	-14	-18	-20	-26	-30	-36	-38	-44	-48	-54
Linear															
SIMON32	-2	-4	-6	-8	-12	-14	-18	-20	-26	-30	-34	-36	-38	-40	-42
SIMON48	-2	-4	-6	-8	-12	-14	-18	-20	-26	-30	-36	-38	-44	-46	-50
SIMON64	-2	-4	-6	-8	-12	-14	-18	-20	-26	-30	-36	-38	-44	-48	-54

5.3 Computing the Probability of a Differential

Given a differential characteristic it is of interest to determine the probability of the associated differential ($\Delta\text{in} \xrightarrow{f^r} \Delta\text{out}$) as it might potentially have a much higher probability then the characteristic by itself. Often it is assumed that the probability of the best differential characteristic can be used to estimate the probability of the best differential. However, this assumption only gives an inaccurate estimate in the case of SIMON.

Similarly to the previous approach for finding the characteristic, we can formalize the problem of finding the probability of a given differential in the following way:

- Add the same system of constraints which were used for finding the characteristic.
- Add a constraint fixing the variables (x_0, y_0) to Δin and (x_r, y_r) to Δout.
- Use a SAT solver to find **all** solutions s_i for the weight w_i.
- Increment the weight w_i and repeat the procedure.

The probability of the differential is then given by

$$P(\Delta in \xrightarrow{f^r} \Delta out) = \sum_{i=w_{\min}}^{w_{\max}} s_i \cdot 2^{-i} \tag{9}$$

where s_i is the number of characteristics of weight i.

We used this approach to compute the probability for various differentials (see Table 2). As one example we've chosen the 16-round SIMON48 differential used in [16]. Enumerating all characteristics up to probability 2^{-60} takes less than five minutes on single cpu core and we continued up to a probability of 2^{-68}.

Additionally we looked at differentials which can cover an additional round compared to previous attacks and might have potential to improve attacks. For SIMON48 we also looked more closely how the distribution of characteristics behaves for a consecutive number of rounds (see Figure 3). The main advantage of our method here is that we get all characteristics with a specific probability and the performance seems to be very competitive compared to dedicated approaches like in [17].

Still the approach is limited by the available computing power and in general it seems to be infeasible to count all characteristics for weights in $[w_{\min}, w_{\max}]$, as the number of characteristics seems to grow exponential in the weight.

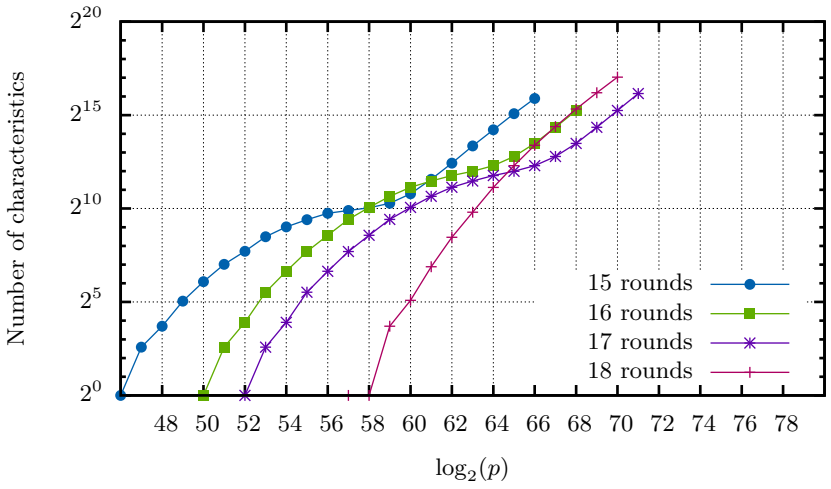

Fig. 2. The total number of characteristics for a specific weight for SIMON48 is shown.

Upper Bound for the Characteristics. During our experiments we observed that it seems to be an easy problem for the SMT/SAT solver to proof the absence of differential characteristics above $w_{\max}$. This can be used to get an upper bound on the weight of the characteristics contributing to the differential. The procedure is similar to finding to finding the optimal characteristics.

- Start with a high initial weight w_i.
- Add the same system of constraints which were used for finding the characteristic.
- Add a constraint fixing the variables (x_0, y_0) to Δin and (x_r, y_r) to Δout.
- Query if there is a solution for this weight.
- Decrease the weight w_i and repeat the procedure until a solution is found.

Table 2. Overview of the differentials analysed with regard to their probability and the range of the weight of contributing characteristics.

Cipher	Rounds	Δin	Δout	$w_{\min}$	$w_{\max}$	$\log_2(p)$
SIMON32	13	$(0, 40)$	$(4000, 0)$	36	91	≥ -30.35
SIMON48	15	$(20, 800088)$	$(800208, 2)$	46	219	≥ -41.02
SIMON48	16	$(800000, 220082)$	$(800000, 220000)$	50	256	≥ -44.33
SIMON48	17	$(80, 222)$	$(222, 80)$	52	269	≥ -46.33
SIMON64	22	$(440, 1880)$	$(440, 100)$	72	502	≥ -61.48

6 Analysis of the Parameter Choices

Due to the absence of any criteria and security analysis of the choice of rotation constants we applied our methods to find good sets of parameters and also compare them with the

chosen parameters by the designers. We considered all possible sets of rotation constants a, b, c[5] and checked them for diffusion properties and the optimal differential and linear characteristics.

6.1 Diffusion

As a very simple measure to estimate the quality of the rotation constants, we measure the number of rounds that are needed to reach full diffusion. Full diffusion is reached when every state bit principally depends on all input bits. Compared to computing linear and differntial properties, computing the dependency is trivially computed.

In Table 3 we give a comparison to how well the standard SIMON rotation parameters fare within the distribution of all possible parameter sets. The exact distributions for all SIMON variants can be found in the appendix in Table 10.

Table 3. The number of rounds after which full diffusion is reached for the standard Simon parameters in comparison to the whole possible set of parameters.

Block size	32	48	64	96	128
Standard parameters	7	8	9	11	13
Median	8	10	11	13	14
First quartile	7	9	9	11	12
Best possible	6	7	8	9	10
Rank	2nd	2nd	2nd	3rd	4th

6.2 Differential and Linear

As a second criteria for our parameters, we computed for all $a > b$ and $\gcd(a - b, n) = 1$ the optimal differential and linear characteristics for 10 rounds SIMON32, SIMON48 and SIMON64. A list of the parameters which are optimal for all three variants of SIMON can be found in Appendix C.

It is important here to note that there are also many parameters, including the standard choice, for which the 10-round characteristics for SIMON32 have a probability of 2^{-25} compared to the optimum of 2^{-26}. However, his effect does not occur for more than 10 rounds and also not for larger variants of SIMON.

6.3 Interesting Alternative Parameter Sets

Our investigation resulted in particular in three sets of parameters that deserve further attention. Those variants, SIMON$[12, 5, 3]$, SIMON$[7, 0, 2]$ and SIMON$[1, 0, 2]$ seem very promising alternatives to the standard parameters.

SIMON$[12, 5, 3]$ has the best diffusion amongst the parameters which have optimal differential and linear characteristics for 10 rounds. The two other choices are both restricted by setting $b = 0$ as this would allow a more efficient implementation in software.

[5] Without lack of generality, we assume though that $a \geq b$.

Among those SIMON$[7,0,2]$ has the best diffusion and the characteristics behave similar to the standard parameters. Ignoring the diffusion SIMON$[1,0,2]$ seems also an interesting choice as it is optimal for the differential and linear characteristics.

If we look at the differential corresponding to the best differential characteristic of SIMON$[7,0,2]$ and SIMON$[1,0,2]$, then we can see the number of characteristics is significant higher than for the standard parameters 4. However, for SIMON$[12,5,3]$ the differential shows a surprisingly different behavior and the probability of the differential is much closer to the probability of the characteristic. The characteristics seem to be worse for the larger variants as can be seen in Table 5.

7 Conclusion and Future Work

In this work we analyzed SIMON-like round functions with the aim of understanding the possible design criteria for SIMON better. We gave explicit formulas for calculating the differential probability and square correlation of the generalized round function. We hope that this will ease future cryptanalysis and help to rigorize attacks and security proofs alike.

Clearly, this work opens up for further investigations. In particular, the choice and reasoning of the NSA parameters for SIMON remains unclear. The results of our study assist in determining the quality of the original parameters regarding differential and linear properties and do not hint towards any serious flaw in the choice. However, we identified three alternative set of parameters (SIMON$[12,5,3]$, SIMON$[7,0,2]$ and SIMON$[1,0,2]$) that, from our perspective, might be worth considering as they compare favorable for some metrics to the original ones and we would like to encourage further studies on those variants. Besides our progress concerning the round function, the design of the key-scheduling remains largely unclear and further investigation is needed here.

References

1. Borghoff, J., Canteaut, A., Güneysu, T., Kavun, E.B., Knežević, M., Knudsen, L.R., Leander, G., Nikov, V., Paar, C., Rechberger, C., Rombouts, P., Thomsen, S.S., Yalçın, T.: PRINCE - A Low-Latency Block Cipher for Pervasive Computing Applications - Extended Abstract. In: ASIACRYPT. Volume 7658 of LNCS., Springer (2012) 208–225
2. Albrecht, M.R., Driessen, B., Kavun, E.B., Leander, G., Paar, C., Yalçin, T.: Block ciphers - focus on the linear layer (feat. PRIDE). In Garay, J.A., Gennaro, R., eds.: Advances in Cryptology - CRYPTO 2014 - 34th Annual Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2014, Proceedings, Part I. Volume 8616 of Lecture Notes in Computer Science., Springer (2014) 57–76
3. Grosso, V., Leurent, G., Standaert, F.X., Varıcı, K.: LS-Designs: Bitslice Encryption for Efficient Masked Software Implementations. In: Fast Software Encryption (FSE). LNCS, Springer (2014, to appear)
4. Daemen, J., Peeters, M., Assche, G.V., Rijmen, V.: The NOEKEON Block Cipher. Submission to the NESSIE project (2000)
5. 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/.
6. Alizadeh, J., Alkhzaimi, H.A., Aref, M.R., Bagheri, N., Gauravaram, P., Kumar, A., Lauridsen, M.M., Sanadhya, S.K.: Cryptanalysis of simon variants with connections. In: Radio Frequency Identification: Security and Privacy Issues. Springer (2014) 90–107

7. Biryukov, A., Velichkov, V.: Automatic search for differential trails in arx ciphers. In: Topics in Cryptology–CT-RSA 2014. Springer (2014) 227–250
8. Dinur, I.: Improved differential cryptanalysis of round-reduced speck. In: Selected Areas in Cryptography–SAC 2014. Springer (2014) 147–164
9. Wang, Q., Liu, Z., Varıcı, K., Sasaki, Y., Rijmen, V., Todo, Y.: Cryptanalysis of reduced-round simon32 and simon48. In: Progress in Cryptology–INDOCRYPT 2014. Springer (2014) 143–160
10. Sun, S., Hu, L., Wang, P., Qiao, K., Ma, X., Song, L.: Automatic security evaluation and (related-key) differential characteristic search: Application to simon, present, lblock, des (l) and other bit-oriented block ciphers. In: Advances in Cryptology–ASIACRYPT 2014. Springer (2014) 158–178
11. Sun, S., Hu, L., Wang, M., Wang, P., Qiao, K., Ma, X., Shi, D., Song, L., Fu, K.: Constructing mixed-integer programming models whose feasible region is exactly the set of all valid differential characteristics of simon. Cryptology ePrint Archive, Report 2015/122 (2015) http://eprint.iacr.org/.
12. Abed, F., List, E., Lucks, S., Wenzel, J.: Differential Cryptanalysis of round-reduced Simon and Speck. In: International Workshop on Fast Software Encryption-FSE. (2014)
13. Mouha, N., Preneel, B.: Towards finding optimal differential characteristics for arx: Application to salsa20. Cryptology ePrint Archive, Report 2013/328 (2013) http://eprint.iacr.org/.
14. Aumasson, J.P., Jovanovic, P., Neves, S.: Analysis of norx: Investigating differential and rotational properties. Cryptology ePrint Archive, Report 2014/317 (2014) http://eprint.iacr.org/.
15. Carlet, C. In: Vectorial (multi-output) Boolean Functions for Cryptography. Cambridge University Press (to appear)
16. Sun, S., Hu, L., Wang, M., Wang, P., Qiao, K., Ma, X., Shi, D., Song, L., Fu, K.: Towards finding the best characteristics of some bit-oriented block ciphers and automatic enumeration of (related-key) differential and linear characteristics with predefined properties. Cryptology ePrint Archive, Report 2014/747 (2014) http://eprint.iacr.org/.
17. Biryukov, A., Roy, A., Velichkov, V.: Differential analysis of block ciphers simon and speck. In: International Workshop on Fast Software Encryption-FSE. (2014)

A Python code to calculate linear and differential probabilities

In the following, code for calculating the differential and linear probabilities are given in Python. Restrictions are that the constants need to fulfil $\gcd(a-b,n)=1$ and n has to be even. We assume that the functions $S^a(x)$ and$\mathrm{wt}(x)$ have been implemented as well as a function that `parity` that calculates the parity $\mathrm{wt}(x) \mod 2$ of a bit vector x. a, b, and c have to be defined in the program as well.

The differential probability of $\alpha \xrightarrow{f} \beta$ can then be calculated with the following function:

```
def pdiff (alpha,beta):
    gamma = beta ^ S(alpha,2)
    if alpha == 2**n-1:
        if hw(tmp)%2 == 0:
            return 2**(n-1)
        else:
            return 0
    varibits = S(alpha, 8) | S(alpha,1)
    if gamma & ~varibits != 0:
        return 0
    doublebits = S(alpha,-6) & ~S(alpha,1) & S(alpha,8)
    if (gamma ^ S(gamma,-7)) & doublebits != 0:
        return 0
    return 2**(-hw(varibits^doublebits))
```

The squared correlation of $\alpha \xrightarrow{f} \beta$ can be calculated with the following function:

```
def plin (alpha,beta):
    alpha ^= S(beta,-c)
    if ((S(beta,-a) | S(beta,-b)) ^ alpha) & alpha != 0:
        return 0
    if beta == 2**n-1:
           t, v = lin, 0
           while t != 0:
               v ^= t & 3
               t >>= 2
           if v != 0:
               return 0
           else:
               return 2**(-n+2)
    tmp = beta
    abits = beta
    while tmp != 0:
        tmp = beta & S(tmp, -(a-b))
        abits ^= tmp
    sbits = S(beta, -(a-b)) & ~beta & ~S(abits, -(a-b))
    sbits = S(sbits, -b)
    pbits = 0
```

```
while sbits != 0:
    pbits ^= sbits & alpha
    sbits = S(sbits, (a-b)) & S(beta,-b)
    sbits = S(sbits, (a-b))
    pbits = S(pbits, 2*(a-b))
if pbits != 0:
    return 0
return 2**(-2*hw(abits))
```

B Additional Differential Bounds

Table 4. Analysis of the 13 rounds differentials for SIMON32

$\log_2(p)$	$[8,1,2]$	$[12,5,3]$	$[7,0,2]$	$[1,0,2]$
36	1	1	4	1
37	4	2	16	6
38	15	3	56	27
39	46	2	144	88
40	124	1	336	283
41	288	0	744	822
42	673	0	1644	2297
43	1426	0	3420	6006
44	2973	0	6933	14954
45	5962	0	13270	34524
46	11661	1	24436	73972
47	21916	3	43784	150272
48	40226	14	76261	292118
49	72246	32	130068	-
50	126574	54	218832	-
51	218516	83	362284	-

Table 5. Overview of the optimal differential characteristics for different variants of SIMON with $a = 12$, $b = 5$, $c = 3$.

Rounds:	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Differential															
SIMON32	-2	-4	-6	-8	-12	-14	-18	-20	-26	-28	-34	-36	-42	-44	-47
SIMON48	-2	-4	-6	-8	-12	-14	-18	-20	-26	-30	-36	-36	-38	-40	-42
SIMON64	-2	-4	-6	-8	-12	-14	-18	-20	-26	-30	-35	-37	-43	-47	-

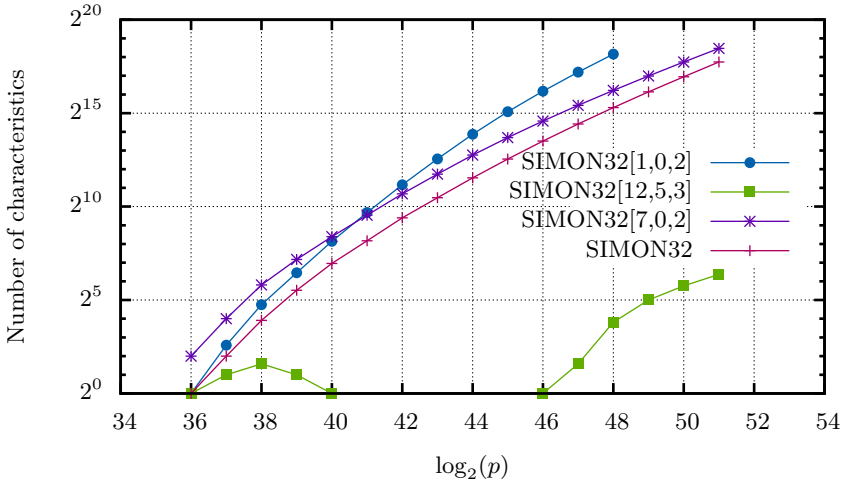

Fig. 3. Distribution of characteristics for 13-round differentials for different variants of SIMON32[a, b, c].

Table 6. Overview of the optimal differential characteristics for different variants of SIMON with $a = 7$, $b = 0$, $c = 2$.

Rounds:	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Differential															
SIMON32	-2	-4	-6	-8	-12	-14	-18	-20	-25	-30	-35	-36	-38	-40	-42
SIMON48	-2	-4	-6	-8	-12	-14	-18	-20	-26	-30	-35	-38	-44	-48	-53
SIMON64	-2	-4	-6	-8	-12	-14	-18	-20	-26	-30	-36	-38	-44	-48	-

Table 7. Overview of the optimal differential characteristics for different variants of SIMON with $a = 1$, $b = 0$, $c = 2$.

Rounds:	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Differential															
SIMON32	-2	-4	-6	-8	-12	-14	-18	-20	-26	-30	-36	-36	-38	-40	-42
SIMON48	-2	-4	-6	-8	-12	-14	-18	-20	-26	-30	-36	-38	-44	-48	-54
SIMON64	-2	-4	-6	-8	-12	-14	-18	-20	-26	-30	-36	-38	-44	-48	-54

Table 8. Number of differential characteristics for the differential $(0, 40) \xrightarrow{f^{13}} (4000, 0)$ for SIMON32.

$\log_2(p)$	#Characteristics
36	1
37	4
38	15
39	46
40	124
41	288
42	673
43	1426
44	2973
45	5962
46	11661
47	21916
48	40226
49	72246
50	126574
51	218516

Table 9. Number of differential characteristics for the differential $(440, 1880) \xrightarrow{f^{22}} (440, 100)$ for SIMON64.

$\log_2(p)$	#Characteristics
72	2
73	14
74	74
75	306
76	1105
77	3502
78	10279
79	27773
80	71337
81	173431
82	228685

C Optimal parameters for differential characteristics

The parameter sets that are optimal regarding differential characteristics in SIMON32, SIMON48, and SIMON64 are given here:

```
(1, 0, 2), (1, 0, 3), (2, 1, 3), (4, 3, 5), (5, 0, 10),
(5, 0, 15), (5, 4, 3), (7, 0, 14), (7, 6, 5), (8, 1, 3),
(8, 3, 14), (8, 7, 5), (10, 5, 15), (11, 6, 1), (12, 1, 7),
(12, 5, 3), (12, 7, 1), (13, 0, 10), (13, 0, 7), (13, 8, 2)
```

D Distribution for Diffusion

Table 10. For each Simon variant and each possible number of rounds, the number of possible combinations of rotation constants (a, b, c) with $a \geq b$ is given that reaches full diffusion.

SIMON32

Rounds	6	7	8	9	10	11	17	∞
$\#(a,b,c)$	48	600	528	88	144	128	64	576

SIMON48

Rounds	7	8	9	10	11	13	14	15	25	∞
$\#(a,b,c)$	48	1392	1680	792	528	344	144	128	64	2080

SIMON64

Rounds	8	9	10	11	12	13	15	17	18	19	33	∞
$\#(a,b,c)$	384	4800	2112	2256	1152	608	512	48	288	256	128	4352

SIMON96

Rounds	9	10	11	12	13	14	15	16	17	19	21	25	26	27	49	∞
$\#(a,b,c)$	336	4272	13920	7104	5568	3456	912	1152	800	1568	640	48	288	256	128	16000

SIMON128

Rounds	10	11	12	13	14	15	16	17	18	19	20
$\#(a,b,c)$	768	10944	26112	25536	9024	6912	7488	2496	192	1824	2304

21	23	24	25	33	34	35	65	∞
1792	1024	960	512	96	576	512	256	33792