A Greater GIFT: Strengthening GIFT against Statistical Cryptanalysis

Ling Sun[1,2,3], Bart Preneel[4], Wei Wang[1,3], and Meiqin Wang(✉)[1,3,5]

[1] Key Laboratory of Cryptologic Technology and Information Security,
Ministry of Education, Shandong University, Jinan, China
[2] State Key Laboratory of Cryptology, P.O.Box 5159, Beijing, 100878, China
[3] School of Cyber Science and Technology, Shandong University, Qingdao, China
[4] Department of Electrical Engineering-ESAT,
KU Leuven and imec, Leuven, Belgium
[5] Quan Cheng Shandong Laboratory, Jinan, China
{lingsun, weiwangsdu, mqwang}@sdu.edu.cn, bart.preneel@kuleuven.be

Abstract. GIFT-64 is a 64-bit block cipher with a 128-bit key that is more lightweight than PRESENT. This paper provides a detailed analysis of GIFT-64 against differential and linear attacks. Our work complements automatic search methods for the best differential and linear characteristics with a careful manual analysis. This hybrid approach leads to new insights. In the differential setting, we theoretically explain the existence of differential characteristics with two active S-boxes per round and derive some novel properties of these characteristics. Furthermore, we prove that all optimal differential characteristics of GIFT-64 covering more than seven rounds must activate two S-boxes per round. We can construct all optimal characteristics by hand. In parallel to the work in the differential setting, we conduct a similar analysis in the linear setting. However, unlike the clear view in differential setting, the optimal linear characteristics of GIFT-64 must have at least one round activating only one S-box. Moreover, with the assistance of automatic searching methods, we identify 24 GIFT-64 variants achieving better resistance against differential attack while maintaining a similar security level against a linear attack. Since the new variants strengthen GIFT-64 against statistical cryptanalysis, we claim that the number of rounds could be reduced from 28 to 26 for the variants. This observation enables us to create a cipher with lower energy consumption than GIFT-64. Similarly to the case in GIFT-64, we do not claim any related-key security for the round-reduced variant as this is not relevant for most applications.

1 Introduction

The expanded deployment of small computing devices that have limited resources (e.g., Radio-Frequency IDentification (RFID) tags, industrial controllers, intra-body sensors) strongly push the evolution of lightweight cryptography. There has been a significant amount of work done by the research community related to this topic. New lightweight algorithms are being proposed on a regular basis. Some lightweight algorithms such as PRESENT [12], PHOTON [19], and

SPONGENT [11] have already been included in ISO standards (ISO/IEC 29192-2:2012 and ISO/IEC 29192-5:2016).

Among the numerous lightweight primitives, PRESENT is probably one of the first candidates particularly designed for efficient hardware implementations. Although the security margin of PRESENT has been reduced by exploiting the clustering effect of linear characteristics [14], it is one of the first generation lightweight ciphers. Notably, NOEKEON [17], which has good hardware performance, was designed in 2000 also before the term lightweight cryptography was widely used.

Ten years after the publication of PRESENT, Banik *et al.* [4,5] revisit the design strategy of PRESENT and propose a new design, named GIFT, that gains much-increased efficiency in hardware and software implementations. In order to avoid some of the potential weaknesses of PRESENT, the designers develop a construction paradigm called "Bad Output must go to Good Input (BOGI)" to guide the selection of bit permutations in PRESENT-like ciphers. GIFT outperforms a vast number of lightweight designs and remains a competitive cipher to date.

Design and cryptanalysis are two inseparable aspects in the development of cryptography and can bring out the best in each other. In the past decade, automatic methods [27,26,32,23,28] gradually develop into powerful tools facilitating the analyses of symmetric-key primitives. This approach has been very successful in developing better attacks and security bounds.

However, tempted by the convenient and fast usage of automatic tools, researchers may spend less attention on a careful study of the primitives themselves. Our research shows that such an analysis can identify new properties and lead to a better understanding of the strength and weaknesses of a design.

This paper studies GIFT-64 with both automatic methods and mathematical analysis; this "hybrid" method uncovers new insights into the security of GIFT-64 and some of its variants.

1.1 Our Results

Motivated by some new observations on differential and linear attacks of GIFT-64, we attempt to explain the results and propose in-depth cryptanalyses of the cipher. The results of this paper can be summarised as follows.

Properties of differential characteristics activating two S-boxes per round. For the crucial role of differential characteristics with two active S-boxes in each round, we try to infer more properties of these characteristics. An alternative description for the round function of GIFT-64 is introduced, where internal states are viewed as 4×4 matrices. With the help of the alternative description, we first show that, for differential characteristics activating two S-boxes per round, the two active S-boxes in one of the first two rounds must be located in the same column of the matrix state. Then, we derive some conditions on the differential propagation for the bit permutation operating on the column of the state, and 26 candidate differential propagations are discovered. After evaluating the compatibilities among these candidates, we prove the existence of differential characteristics with two active S-boxes per round. Beyond that, we also confirm that any differential characteristics covering more than seven rounds and activating two S-boxes in each round must utilise some of the 26 candidate differential propagations.

Explicit formula for the differential probability of the optimal characteristic. We propose an explicit formula for the differential probability of the optimal characteristic. Precisely, the probability $\Pr(r)$ of r-round optimal differential characteristics with $r \geqslant 8$ can be calculated with the following equation

$$-\log_2\big(\Pr(r)\big) = \begin{cases} [(r-3)/2] \cdot 10 + 12 & \text{if } r \bmod 2 \equiv 1, \\ [(r-2)/2] \cdot 10 + 8 & \text{otherwise.} \end{cases}$$

All optimal differential characteristics of `GIFT-64`. All optimal differential characteristics covering more than seven rounds with the maximum probability can be constructed starting from the 26 candidate differential propagations. In other words, all optimal differential characteristics of `GIFT-64` must activate two S-boxes per round. In addition, we show that for the round-reduced variant with an odd number of rounds, the number of optimal characteristics is 288; otherwise, the number of optimal characteristics is 10400.

Properties of linear characteristics with two active S-boxes per round. In parallel to the analysis in the differential setting, we also investigate linear characteristics activating two S-boxes in each round. Moreover, we present some properties for this kind of characteristic, and verify that they can be constructed. However, unlike the clear view in the differential setting, the optimal linear characteristics for `GIFT-64` must contain at least one round with only one active S-box.

Variants with comparable differential and linear properties. Considering the gap between the upper bounds on the differential probability and the linear correlation, we wonder whether we can find a variant of `GIFT-64` with analogous security levels under the differential and linear settings. To facilitate the investigation, we devise a sufficient condition for two `GIFT-64`-like ciphers to be equivalent to each other, enabling us to create an equivalence relation over the set of all `GIFT-64`-like ciphers. Based on the equivalence relation, we identify 168 equivalence classes; the variants in each class share the same cryptographic properties. In other words, it is sufficient to carefully analyse 167 representative variants. After performing an automatic searching method, we recognise one equivalence class, denoted as `GIFT-64[2021]`, with both lengths of the optimal effective differential and linear characteristics equal to 12. In other words, comparing to `GIFT-64`, all the 24 variants in `GIFT-64[2021]` achieve better resistance against differential cryptanalysis while maintaining a similar security level against linear cryptanalysis.

Resistance against other attacks. The security of variants in GIFT-64[2021] w.r.t. the impossible differential attack [9,21], the zero-correlation attack [13], and the integral attack [22] were checked with automatic methods in [15,16,28,33]. Since the new variants strengthen GIFT-64 against statistical cryptanalysis, we claim that 26 rounds could be used rather than 28 rounds for the variants. On this basis, we create a 26-round variant without related-key security, which is more energy-efficient than GIFT-64[6]. Nevertheless, we find that the performance of the 24 variants in the related-key differential attack setting is inferior to that of GIFT-64. This observation suggests that the designers have evaluated the security of the cipher under the related-key differential attacks, although they do not claim security in this setting. For most applications, this security is not required; for the few applications where this is required, the key schedule of the newly proposed variant could be redesigned.

Outline. In Section 2, we review the target cipher GIFT-64 and recall the automatic searching method exploited in this paper. Motivated by some observations on the experimental results, Section 3 presents a series of new differential properties of GIFT-64. In parallel to the search in Section 3, we present in-depth analytic results in the linear setting in Section 4. Section 5 argues why GIFT-64 can indeed be improved by creating a variant. At last, we conclude the paper and list future work in Section 6.

2 Preliminary

In this section, we first review the overall structure and the design philosophy of GIFT-64. Then, an automatic searching method, utilised to assist the following analyses, is briefly recalled.

2.1 Specification of GIFT-64

GIFT [4] is a family of lightweight block ciphers composed of two versions. In this paper, we only focus on GIFT-64, a 64-bit block cipher with a 128-bit key and with 28 rounds.

The cipher initialises the cipher state S with a 64-bit plaintext $b_0 b_1 \cdots b_{63}$, where b_0 stands for the most significant bit. Alternatively, the cipher state can be expressed as sixteen 4-bit nibbles $S = w_0 \| w_1 \| \cdots \| w_{15}$. Apart from the plaintext, the 128-bit key $K = k_0 \| k_1 \| \cdots \| k_7$ acts as the other input of the cipher. After initialising as above, the cipher iteratively uses the round function to update the cipher state. Each round of GIFT-64 consists of three steps.

SubCells(SC). GIFT-64 applies an invertible 4-bit S-box GS to every nibble of the cipher state.

[6] The GIFT designers also did not claim related-key security.

PermBits(`PB`). This operation maps the bit from the position i of the cipher state to the position $P_{64}(i)$ as

$$b_{P_{64}(i)} \leftarrow b_i,\ i \in \{0, 1, \dots, 63\},$$

where $P_{64}(i)$ can be calculated as

$$63 - \left\{4\left\lfloor\frac{63-i}{16}\right\rfloor + 16\left\lfloor 3\left\lfloor\frac{(63-i) \bmod 16}{4}\right\rfloor + (63-i) \bmod 16\right\rfloor + (63-i) \bmod 4\right\} \bmod 64.$$

AddRoundKey($\texttt{ARK}_{RK_r}$). This step adds the round key and the round constant. Since the round constant does not affect the analysis in this paper, we only pay attention to the round key. In the r-th round, a 32-bit round key RK_r is extracted from the key state and is further partitioned into two 16-bit words as $RK_r = U \| V = u_0 u_1 \cdots u_{15} \| v_0 v_1 \cdots v_{15}$. Then, U and V are XORed to the cipher state as

$$b_{4 \cdot i+2} \leftarrow b_{4 \cdot i+2} \oplus u_i,\ b_{4 \cdot i+3} \leftarrow b_{4 \cdot i+3} \oplus v_i,\ i \in \{0, 1, \dots, 15\}.$$

The design of the key schedule realises the goals of minimising the hardware area and supporting efficient software implementation simultaneously. It only involves the key state rotation in blocks of 16-bit and the bit rotation within some 16-bit blocks.

Key schedule. Before the key state updates, a round key is first extracted from it. To be precise, two 16-bit words of the key state are set as the round key $RK = U \| V$, where

$$U \leftarrow k_6,\ V \leftarrow k_7.$$

After generating the round key, `GIFT-64` employs the following transformation to update the key state,

$$k_0 \| k_1 \| \cdots \| k_7 \leftarrow (k_6 \ggg 2) \| (k_7 \ggg 12) \| k_0 \| k_1 \| \cdots \| k_5,$$

where '$\ggg i$' represents an i-bit right rotation within a 16-bit word.

For more details about the cipher, see Banik *et al.* [5].

2.2 Bit Permutation in PermBits Operation of `GIFT-64`

After fixing the overall structure of the cipher as a `PRESENT`-like [12] one, the designers set out a small area goal and manage to use an S-box with a lower implementation cost than that of `RECTANGLE` [34]. However, the S-box GS with low cost cannot reach the differential and linear branching numbers of three. In other words, for GS, *1-1 bit transitions*, which are referred to as differential/linear propagations with input and output differences/masks being unit vectors, are possible in both the differential and linear settings. Note that the 1-1 bit transition may result in long differential and linear characteristics with a single active S-box per round. Hence, to ensure the nonexistence of consecutive 1-1 bit differential and linear transitions in the cipher, the designers propose a new construction paradigm called "*Bad Output must go to Good Input (BOGI)*" to design the bit permutation.

Denote the sixteen S-boxes in the i-th round as GS_0^i, GS_1^i, ..., GS_{15}^i. The S-boxes can be grouped into two different ways:

- ▷ the Quotient group $Q_x^i = \{GS_{4\cdot x}^i, GS_{4\cdot x+1}^i, GS_{4\cdot x+2}^i, GS_{4\cdot x+3}^i\}$, $0 \leqslant x \leqslant 3$;
- ▷ the Remainder group $R_x^i = \{GS_x^i, GS_{x+4}^i, GS_{x+8}^i, GS_{x+12}^i\}$, $0 \leqslant x \leqslant 3$.

With this notation, the design of the 64-bit permutation in PermBits operation boils down to the construction of four independent and identical 16-bit permutations that map the output bits of Q_x^i to the input bits of R_x^{i+1}. In this sense, the BOGI paradigm can be viewed as a guideline for the creation of the 16-bit group mapping. It determines the rule to map the output bits of S-boxes in Q_x^i to the input bits of S-boxes in R_x^{i+1} and is analysed in differential and linear setting parallelly.

In the differential setting, we consider the *1-1 bit DDT* [4], a sub-table of the differential distribution table (DDT) [10], composed of differential transitions with input and output differences being unit vectors (cf. Table 4 in Supplementary Material A of the long version for the 1-1 bit DDT of `GIFT`). Given the 1-1 bit DDT, an input (resp., output) difference $\Delta x = x_0x_1x_2x_3$ (resp., $\Delta y = y_0y_1y_2y_3$) is named as a *good input* (resp., *good output*) if the corresponding row (resp., column) has all zero entries; otherwise, it is called a *bad input* (resp., *bad output*). Denote GI, GO, BI, and BO the sets of positions for the nonzero bits in the good inputs, good outputs, bad inputs, and bad outputs, respectively. Then, based on the 1-1 bit DDT of `GIFT`, we have $GI = \{0, 1, 2\}$, $GO = \{1, 2, 3\}$, $BI = \{3\}$, and $BO = \{0\}$.

Notice that a bad output could come from a 1-1 bit transition through a certain S-box in the current round. The primary purpose of BOGI is to ensure that the existing 1-1 bit transition will not head to another 1-1 bit transition in the succeeding round, which is realised by artificially mapping the active bit of the (potentially) bad output to an active bit of some good inputs in the next round. Concretely, regarding a 1-1 bit DDT with $|BO| \leqslant |GI|$, the *differential BOGI permutation* is defined as a permutation $\pi : BO \cup GO \rightarrow BI \cup GI$ with $\pi(BO) = \{\pi(i) \mid i \in BO\} \subseteq GI$. Likewise, in the linear case, the *linear BOGI permutation* can be derived regarding the *1-1 bit LAT* (cf. Table 5 in Supplementary Material A of the long version), which is the dual notion of 1-1 bit DDT in the linear approximation table (LAT) [24].

For the purpose of increasing the security of the cipher regarding differential and linear cryptanalyses at the same time, the *BOGI permutation* exploited in the cipher should belong to the intersection of the set of differential BOGI permutations and the set of linear BOGI permutations. For `GIFT`, the BOGI permutation is fixed as the identity mapping $\pi(i) = i$. After determining the BOGI permutation, during the construction of the group mapping, the i-th output bits of the S-boxes in Q_x^i must be connected to the $\pi(i)$-th input bits of the S-boxes in R_x^{i+1}. This mandatory requirement breaks the existence of consecutive 1-1 bit transitions. Hence, the cipher assembled with this kind of group mappings does not exhibit long differential and linear characteristics activating a single S-box per round.

Except for the above countermeasure to enhance the security, the group mapping should also validate the following four rules to guarantee the bijectivity of the linear layer and attain an optimal full diffusion[7].

1. The input bits of an S-box in R_x^{i+1} come from 4 distinct S-boxes in Q_x^i.
2. The output bits of an S-box in Q_x^i go to 4 distinct S-boxes in R_x^{i+1}.
3. The input bits of 4 S-boxes from the same Q_x^{i+1} come from 16 different S-boxes.
4. The output bits of 4 S-boxes from the same R_x^i go to 16 different S-boxes.

2.3 Accelerated Automatic Search with the SAT Method

This section briefly reviews the accelerated automatic searching method in [30], which will be used to examine the soundnesses of some theoretical results in the coming sections.

The automatic search is realised via the Boolean satisfiability problem (SAT), which intends to determine if there exists an instantiation that satisfies a given Boolean formula. In practice, we transform cryptanalytic problems into SAT problems and employ the same SAT solver CiDiCaL [8] as in [30] to solve all concerned SAT problems.

To facilitate a SAT solver to detect desired differential and linear characteristics, we should first create Boolean formulas to translate the cryptanalytic properties of the cipher. Due to the concise structure of `GIFT`, descriptions of cryptanalytic properties (e.g., the number of differential/linear active S-boxes, the differential probability, and the linear correlation) are reduced to characterisations of properties for the S-box GS. We refer readers to [30] for a detailed approach to generate differential and linear models of a given S-box.

Because we always target characteristics with good cryptanalytic properties (e.g., a small number of active S-boxes, a relatively high differential probability/linear correlation), a cardinality constraint in the form of $\sum_{j=0}^{\omega-1} x_j \leqslant k$ should be integrated into the SAT problem, where x_j's stand for Boolean variables representing cryptanalytic properties of S-boxes, w is the number of x_j's in the cipher, and k is a predicted value for the cryptanalytic property of the cipher. This cardinality constraint can be viewed as an objective function: it tells the SAT solver what kind of characteristics we want to find. With the sequential encoding method [29], the cardinality constraint can be converted into $\mathcal{O}(\omega \cdot k)$ Boolean formulas by introducing $\mathcal{O}(\omega \cdot k)$ auxiliary variables.

The Boolean expressions specifying the cryptanalytic properties of S-boxes and the objective function constitute a basic SAT problem for searching distinguishers. Next, Sun *et al.* [30] managed to incorporate Matsui's bounding conditions abstracted from the branch-and-bound depth-first searching algorithm

[7] `GIFT-64` achieves full diffusion after three rounds.

[25] into the SAT problem to accelerate the automatic search. The efficiency arises from the manipulation of the knowledge of cryptanalytic properties of short characteristics. For instance, suppose that we are checking the existence of R-round differential characteristics $(\varDelta_0, \varDelta_1, \ldots, \varDelta_R)$ with probability no less than $\Pr_{\texttt{Ini}}(R)$, where $\varDelta_i$ implies the input difference of the i-th round. Given the maximum probability $\Pr_{\texttt{Max}}(i)$ achieved by i-round differential characteristics for all $1 \leqslant i \leqslant R-1$, the bounding condition $\mathcal{C}_{(r_1,r_2)}$, originating from the r_1-th round and terminating with the r_2-th round, forces the SAT solver to concentrate on characteristics validating the following inequality

$$\Pr_{\texttt{Max}}(r_1) \cdot \left[\prod_{i=r_1}^{r_2-1} \Pr\left(\varDelta_i \rightarrow \varDelta_{i+1}\right)\right] \cdot \Pr_{\texttt{Max}}(R - r_2 - 1) \leqslant \Pr_{\texttt{Ini}}(R) \ ,$$

where $\Pr\left(\varDelta_i \rightarrow \varDelta_{i+1}\right)$ stands for the probability of the differential propagation $\varDelta_i \rightarrow \varDelta_{i+1}$ in the i-th round. The adjunction of the bounding condition [30] shrinks the solution space of the basic SAT problem and results in a notable speedup.

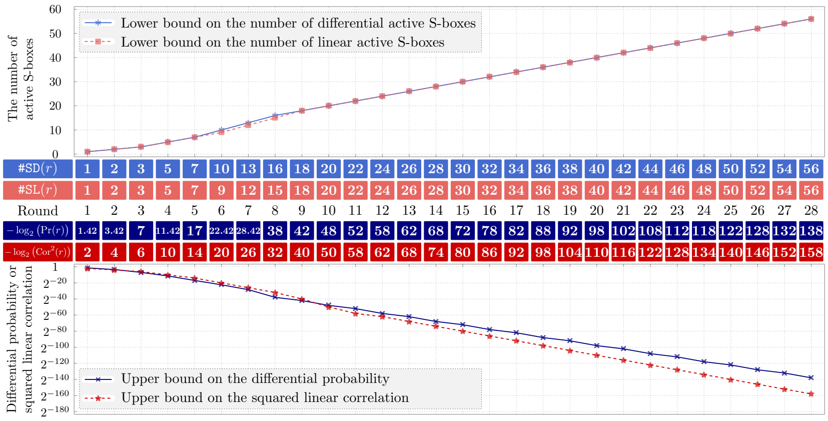
3 Differential Property of GIFT-64

Through analysing the automatic searching results related to differential and linear cryptanalyses of GIFT-64, we attempt to develop an in-depth understanding on the security of the cipher. Therefore, we reimplement the search for GIFT-64 with the publicly available source code provided in [30], even if the authors of [30] have already completed the full picture on the number of active S-boxes, the differential probability, as well as the linear correlation.

Based on the results shown in Fig. 1, this section presents some novel differential properties of GIFT-64. In the following, the minimum numbers of differential and linear active S-boxes for r-round characteristics are denoted as $\texttt{\#SD}(r)$ and $\texttt{\#SL}(r)$, respectively. The maximum differential probability and linear correlation for r-round characteristics are represented as $\Pr(r)$ and $\mathrm{Cor}(r)$.

3.1 Observations on Experimental Results

In Fig. 1, the minimum number of differential active S-boxes $\texttt{\#SD}(r)$ is linearly dependent on r for all $r \geqslant 8$. Starting from the eighth round, $\texttt{\#SD}(r)$ strictly increases by two per round. Further, after decoding the optimal differential characteristic with the maximum probability from the output of the SAT solver, we observe that the optimal characteristics covering more than seven rounds always maintain two active S-boxes in each round. Thus, we wonder whether a characteristic with a single active S-box in some rounds achieving the maximum differential probability exists. The research in this section provides an answer for this issue.



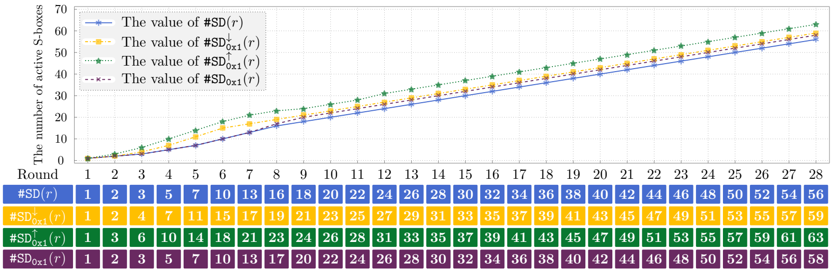
Round	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
#SD(r)	1	2	3	5	7	10	13	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56
#SL(r)	1	2	3	5	7	9	12	15	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56
$-\log_2(\Pr(r))$	1.42	3.42	7	11.42	17	22.42	28.42	38	42	48	52	58	62	68	72	78	82	88	92	98	102	108	112	118	122	128	132	138
$-\log_2(\mathrm{Cor}^2(r))$	2	4	6	10	14	20	26	32	40	50	58	62	68	74	80	86	92	98	104	110	116	122	128	134	140	146	152	158

Fig. 1. Bounds reflecting differential and linear properties of `GIFT-64`.

3.2 Lifted Bounds on the Number of Differential Active S-boxes

Let $\mathbb{D}_{\mathtt{0x1}}$ be the set of differential characteristics with at least one round activating a single S-box, and the value `0x1` equals the input difference of the active S-box. We manage to calculate a lower bound on the number of active S-boxes for characteristics in $\mathbb{D}_{\mathtt{0x1}}$.

The accelerated automatic method reviewed in Section 2.3 is applied to accomplish this task, and we split the search into three steps. To begin with, we explore the lower bound for characteristics with input differences having a single nonzero nibble `0x1`. Then, the characteristics with output differences having a single nonzero nibble `0x1` are considered. Note that the characteristics in $\mathbb{D}_{\mathtt{0x1}}$ can be created with the characteristics in the first two steps. Therefore, the lower bound for characteristics in $\mathbb{D}_{\mathtt{0x1}}$ is derived from the experimental results in the first two steps.



Round	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
#SD(r)	1	2	3	5	7	10	13	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56
$\#\mathtt{SD}^{\downarrow}_{\mathtt{0x1}}(r)$	1	2	4	7	11	15	17	19	21	23	25	27	29	31	33	35	37	39	41	43	45	47	49	51	53	55	57	59
$\#\mathtt{SD}^{\uparrow}_{\mathtt{0x1}}(r)$	1	3	6	10	14	18	21	23	24	26	28	31	33	35	37	39	41	43	45	47	49	51	53	55	57	59	61	63
$\#\mathtt{SD}_{\mathtt{0x1}}(r)$	1	2	3	5	7	10	13	17	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58

Fig. 2. Minimum numbers of differential active S-boxes in different settings.

Step 1: Lower bound for characteristics with input differences having a single nonzero nibble `0x1`. We focus on characteristics with input differences having a single nonzero nibble `0x1`. The set of characteristics satisfying this restriction is denoted as $\mathbb{D}^{\downarrow}_{\texttt{0x1}}$. To obtain a lower bound on the number of active S-boxes for characteristics in $\mathbb{D}^{\downarrow}_{\texttt{0x1}}$, we first convert the restriction on characteristics into Boolean formulas. These formulas are appended to the SAT problem so that the solver ignores unsatisfied characteristics. Besides, the set of bounding conditions terminating with the last round $\mathcal{C}_{(*,R)} = \{\mathcal{C}_{(r,R)} \mid 1 \leqslant r \leqslant R-1\}$ is included in the SAT problem to accelerate the search. Denote the minimum number of active S-boxes for r-round characteristics in $\mathbb{D}^{\downarrow}_{\texttt{0x1}}$ with $\texttt{\#SD}^{\downarrow}_{\texttt{0x1}}(r)$, where $1 \leqslant r \leqslant 28$. Figure 2 shows the results for $\texttt{\#SD}^{\downarrow}_{\texttt{0x1}}(r)$ returned by the solver.

Step 2: Lower bound for characteristics with output differences holding a single nonzero nibble `0x1`. The search space is restricted to characteristics with output differences having a single nonzero nibble `0x1`; the corresponding set of characteristics is denoted with $\mathbb{D}^{\uparrow}_{\texttt{0x1}}$. Also, this constraint is formulated with Boolean expressions, which are added to the basic SAT problem. To simultaneously speed up the search and guarantee the correctness of the test, in this step, we employ the set of bounding conditions starting from the first round $\mathcal{C}_{(0,*)} = \{\mathcal{C}_{(0,r)} \mid 1 \leqslant r \leqslant R-1\}$. Let $\texttt{\#SD}^{\uparrow}_{\texttt{0x1}}(r)$ denote the minimum number of active S-boxes for r-round characteristics in $\mathbb{D}^{\uparrow}_{\texttt{0x1}}$. Figure 2 shows the values for $\texttt{\#SD}^{\uparrow}_{\texttt{0x1}}(r)$.

Step 3: Lower bound for characteristics in $\mathbb{D}_{\texttt{0x1}}$. Let $\texttt{\#SD}_{\texttt{0x1}}(r)$ be the minimum number of active S-boxes for r-round characteristics in $\mathbb{D}_{\texttt{0x1}}$. Since the characteristic in $\mathbb{D}_{\texttt{0x1}}$ can be created with characteristics in $\mathbb{D}^{\uparrow}_{\texttt{0x1}}$ and $\mathbb{D}^{\downarrow}_{\texttt{0x1}}$, a lower bound for the value of $\texttt{\#SD}_{\texttt{0x1}}(r)$ can be calculated with $\texttt{\#SD}^{\downarrow}_{\texttt{0x1}}(*)$ and $\texttt{\#SD}^{\uparrow}_{\texttt{0x1}}(*)$. Specifically, we have

$$\texttt{\#SD}_{\texttt{0x1}}(r) \geqslant \min\left\{\texttt{\#SD}^{\uparrow}_{\texttt{0x1}}(r_1) + \texttt{\#SD}^{\downarrow}_{\texttt{0x1}}(r_2) \;\middle|\; r_1 + r_2 = r,\ r_1 \geqslant 0,\ r_2 \geqslant 0\right\}, \quad (1)$$

and the value of the right-hand side expression is known from the outputs in *Step 1* and *Step 2*. Additionally, as we find the characteristic with the number of active S-boxes exactly matching the lower bound, we ensure that the bound for $\texttt{\#SD}_{\texttt{0x1}}(r)$ in Eqn. (1) is strict. The values of $\texttt{\#SD}_{\texttt{0x1}}(r)$ for all $1 \leqslant r \leqslant 28$ can be found in Fig. 2.

Figure 2 reveals that $\texttt{\#SD}_{\texttt{0x1}}(r) > \texttt{\#SD}(r)$ for all $r \geqslant 8$. Moreover, after exploiting the three-step test to evaluate all sets $\mathbb{D}_{\texttt{i}}$ for $\texttt{i} \in \{\texttt{0x2}, \ldots, \texttt{0xf}\}$ representing the sets of characteristics with at least one round activating a single S-box with the input difference `i`, we find that $\texttt{\#SD}_{\texttt{i}}(r) > \texttt{\#SD}(r)$ for all $r \geqslant 8$ and $\texttt{i} \in \mathbb{F}_2^4 \setminus \{\texttt{0x0}\}$ (cf. Supplementary Material B.1 of the long version). That is, from the eighth round, the optimal differential characteristic with the minimum number of active S-boxes definitely activates more than one S-box in each round. In other words, the optimal characteristic contains at least two active S-boxes per round. Because the characteristics decoded from the solver evidence the existence of differential characteristics with two active S-boxes in each round, we draw the following proposition.

Proposition 1. If $r \geqslant 8$, the optimal r-round differential characteristic of `GIFT-64` with the minimum number of active S-boxes must have two active S-boxes in each round.

3.3 Decreased Upper Bound on the Differential Probability

After obtaining lower bounds on the number of active S-boxes for characteristics in $\mathbb{D}_{\mathtt{i}}$, we attempt to check the differential probability for characteristics in $\mathbb{D}_{\mathtt{i}}$, where `i` traverses all nonzero 4-bit values. The test is also accomplished with three steps as in Section 3.2, and we only accommodate the objective function from the number of active S-boxes to the differential probability. Denote the maximum differential probability for r-round characteristics in $\mathbb{D}_{\mathtt{i}}^{\downarrow}$ (resp., $\mathbb{D}_{\mathtt{i}}^{\uparrow}$, $\mathbb{D}_{\mathtt{i}}$) with $\Pr_{\mathtt{i}}^{\downarrow}(r)$ (resp., $\Pr_{\mathtt{i}}^{\uparrow}(r)$, $\Pr_{\mathtt{i}}(r)$). The results for $\Pr_{\mathtt{i}}^{\downarrow}(r)$, $\Pr_{\mathtt{i}}^{\uparrow}(r)$, and $\Pr_{\mathtt{i}}(r)$ are given in Supplementary Material B.2 of the long version. The following proposition is based on the observation $\Pr_{\mathtt{i}}(r) < \Pr(r)$ for all $\mathtt{i} \in \mathbb{F}_2^4 \setminus \{\mathtt{0x0}\}$ and $r \geqslant 8$.

Proposition 2. If $r \geqslant 8$, the optimal r-round differential characteristic with the maximum probability must activate at least two S-boxes per round.

From Section 3.2 – 3.3, we notice that differential characteristics activating two S-boxes in each round play a crucial role in the security evaluation for `GIFT-64`. Consequently, a natural question is whether one can infer more properties of these characteristics, apart from the quantitative information about active S-boxes. Before looking into these characteristics, we first devise an alternative description for the round function of `GIFT-64`, which facilitates the analyses in the upcoming sections. Note that the designers of `GIFT` proposed a cubic representation of `GIFT-64` [4], which reorganises the 64-bit state as a $4 \times 4 \times 4$ cube. Based on the observation on the cubic representation, Adomnicai *et al.* [1] developed a new `GIFT` representation called *fixslicing* that allows extremely efficient software bitsliced implementations of `GIFT`. The new description in the following is based on a 2-dimensional matrix.

3.4 Alternative Description for the Round Function of `GIFT-64`

In the alternative description, we keep SubCells and AddRoundKey operations and further decompose PermBits operation into two sub-operations. Please find Fig. 3(a) for an illustration.

GroupMaps(`GM`). Denote the 16-bit group mapping utilised in `GIFT-64` as $\mathtt{g}_0$,

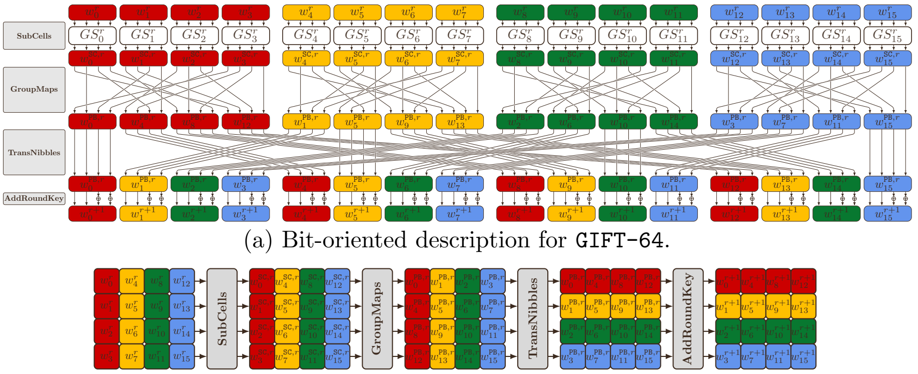
$$\mathtt{g}_0 = (12, 1, 6, 11,\ 8, 13, 2, 7,\ 4, 9, 14, 3,\ 0, 5, 10, 15)\ .$$

It moves the i-th bit of the input to the $\mathtt{g}_0(i)$-th bit for all $0 \leqslant i \leqslant 15$. GroupMaps operation invokes $\mathtt{g}_0$ and independently applies it on each of the 16-bit words $w_{4\cdot j}^{\mathtt{SC},r} \| w_{4\cdot j+1}^{\mathtt{SC},r} \| w_{4\cdot j+2}^{\mathtt{SC},r} \| w_{4\cdot j+3}^{\mathtt{SC},r}$ of the cipher state, where $w_{*}^{\mathtt{SC},r}$ stands for nibbles at the output of the SubCells operation and $0 \leqslant j \leqslant 3$.

TransNibbles(TN). This operation works in nibbles. It shifts the nibble from position i of the cipher state to position $T(i)$ for all $0 \leqslant i \leqslant 15$, and

$$T = (0, 4, 8, 12,\ \ 1, 5, 9, 13,\ \ 2, 6, 10, 14,\ \ 3, 7, 11, 15)\ \ .$$

Equivalently, if we reorganise the cipher state as a 4×4 matrix of nibbles, the bit-oriented description in Fig. 3(a) can be replaced with a nibble-oriented one as in Fig. 3(b), which is a more concise representation. In this description, the 32-bit round key RK_r also should be fitted into a 4×4 matrix of nibbles. In the following, we employ the nibble-oriented description.



(a) Bit-oriented description for `GIFT-64`.

(b) Nibble-oriented description for `GIFT-64`.

Fig. 3. Alternative descriptions for `GIFT-64`.

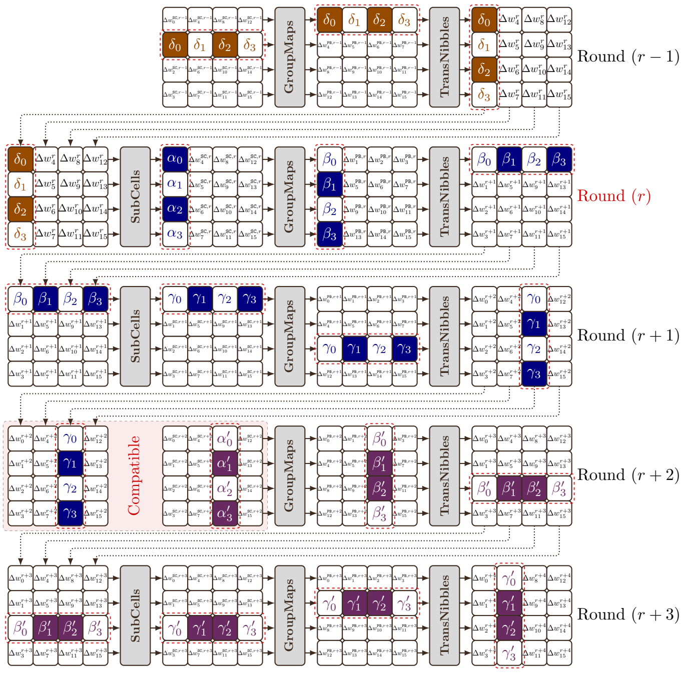
3.5 Differential Characteristics with Two Active S-boxes Per Round

Next, we study the properties of differential characteristics activating two S-boxes per round. Besides, we temporarily omit AddRoundKey operation as it does not influence the differential property in the single-key attack setting.

Lemma 1. For `GIFT-64`, if a differential characteristic activates two S-boxes per round, then the two active S-boxes in one of the first two rounds must be located in the same column of the matrix state.

For the proof of Lemma 1, see Supplementary Material B.3 of the long version.

Now, given a differential characteristic with two active S-boxes per round; we assume that the two active S-boxes in the r-th round are located in the same column. Without loss of generality, the column is set as the first one. Denote the differential propagation of the group mapping g_0 in the r-th round operating on the first column as $\alpha_0\|\alpha_1\|\alpha_2\|\alpha_3 \xrightarrow{\mathsf{g}_0} \beta_0\|\beta_1\|\beta_2\|\beta_3$, where two nibbles in



GroupMaps
TransNibbles
SubCells
Round $(r-1)$
Round (r)
Round $(r+1)$
Round $(r+2)$
Round $(r+3)$
Compatible
δ_0 δ_1 δ_2 δ_3
α_0 α_1 α_2 α_3
β_0 β_1 β_2 β_3
γ_0 γ_1 γ_2 γ_3
α'_0 α'_1 α'_2 α'_3
β'_0 β'_1 β'_2 β'_3
γ'_0 γ'_1 γ'_2 γ'_3

Fig. 4. Illustration for conditions on the group mapping $\mathtt{g}_0$.

$\alpha_0\|\alpha_1\|\alpha_2\|\alpha_3$ are nonzero. In the following, we will see that this propagation should meet some conditions so that the differential characteristic based on it can sustain two active S-boxes in rounds $(r-1)$ and $(r+1)$.

Condition 1 The output difference $\beta_0\|\beta_1\|\beta_2\|\beta_3$ of $\mathtt{g}_0$ has two nonzero nibbles.

Proof. As in Fig. 4, the cipher structure guarantees that $\beta_0\|\beta_1\|\beta_2\|\beta_3$ equals the first row of input difference for the $(r+1)$-th round, which is the composition of input differences for four S-boxes. Because we are analysing characteristics with two active S-boxes in each round, two nibbles among β_0, β_1, β_2, and β_3 have to be nonzero. ■

Condition 2 Two nonzero nibbles in $\beta_0\|\beta_1\|\beta_2\|\beta_3$ cannot take values from the set $\{\mathtt{0x2}, \mathtt{0x4}, \mathtt{0x8}\}$.

Proof. Without loss of generality, suppose that the two nonzero nibbles are β_1 and β_3. As in Fig. 4, β_0, β_1, β_2, and β_3 are input differences of four S-boxes in the $(r+1)$-th round, and we denote the corresponding output differences as γ_0, γ_1, γ_2, and γ_3. Based on the diffusion property of GroupMaps operation, to maintain two active S-boxes in the $(r+2)$-nd round, γ_1 and γ_3 should be unit vectors. Accordingly, the input differences β_1 and β_3 regarding γ_1 and γ_3 must be different from 0x2, 0x4, or 0x8, for these input differences cannot perform the 1-1 bit transition. The proof is complete. ■

Condition 3 Two nonzero nibbles in $\alpha_0\|\alpha_1\|\alpha_2\|\alpha_3$ cannot take values from the set $\{\texttt{0x1}, \texttt{0x2}, \texttt{0x4}\}$.

Proof. Likewise, without loss of generality, suppose that the two nonzero nibbles are α_0 and α_2. We propagate the difference $\alpha_0\|\alpha_1\|\alpha_2\|\alpha_3$ in the backward direction and utilise $\delta_0\|\delta_1\|\delta_2\|\delta_3$ to stand for the input difference of SubCells operation regarding $\alpha_0\|\alpha_1\|\alpha_2\|\alpha_3$. As in Fig. 4, at the output of GroupMaps operation in the $(r-1)$-th round, δ_0 and δ_2 are located in different columns. Thus, they must originate from the two active S-boxes in the $(r-1)$-th round. By the diffusion property of GroupMaps operation, δ_0 and δ_2 should be unit vectors. As δ_0 and δ_2 also act as input differences of two active S-boxes in the r-th round, the corresponding output differences α_0 and α_2 have to take values from the complementary set of $\{\texttt{0x1}, \texttt{0x2}, \texttt{0x4}\} \subset \mathbb{F}_2^4$. ■

Condition 4 Denote β_i and β_j the two nonzero nibbles in $\beta_0\|\beta_1\|\beta_2\|\beta_3$, where $i, j \in \{0, 1, 2, 3\}$ and $i \neq j$. Let $\mathcal{S}_i^D$ and $\mathcal{S}_j^D$ be the sets of 1-bit output differences that can be propagated from β_i and β_j, respectively, i.e.,

$$\mathcal{S}_i^D = \{\gamma_i \mid \beta_i \xrightarrow{GS} \gamma_i \text{ is a possible propagation, and } \gamma_i \text{ is a unit vector}\} \enspace ,$$

$$\mathcal{S}_j^D = \{\gamma_j \mid \beta_j \xrightarrow{GS} \gamma_j \text{ is a possible propagation, and } \gamma_j \text{ is a unit vector}\} \enspace .$$

Then, $\mathcal{S}_i^D \cap \mathcal{S}_j^D \neq \emptyset$ must hold.

Proof. Without loss of generality, suppose that $i = 1$ and $j = 3$. We have already proved in Condition 2 that the output differences γ_1 and γ_3 corresponding to β_1 and β_3 must be unit vectors. As in Fig. 4, γ_1 and γ_3 at the input of GroupMaps operation in the $(r+1)$-th round are located in different columns but the same row. If $\gamma_1 \neq \gamma_3$, then at least one of them differs from `0x1`. Since the following GroupMaps operation shifts the two nonzero bits in γ_1 and γ_3 to different rows, the inequality incurs at least three active S-boxes in the $(r+3)$-rd round for sure. So, the preset condition on the characteristic determines that the output differences corresponding to β_1 and β_3 must be an identical unit vector. ■

Summarising all analyses in the proofs for Condition 1 – 4, we derive the following proposition.

Proposition 3. For an R-round differential characteristic activating two S-boxes per round, if the two active S-boxes in the r-th round are located in the same column, then, for all i with $0 \leqslant r + 2 \cdot i < R$, the two active S-boxes in the $(r + 2 \cdot i)$-th round are also located in the same column.

Based on Lemma 1 and Proposition 3, we conclude that all differential characteristics with two active S-boxes per round can be decomposed into several pieces of 2-round characteristics, for which the two active S-boxes in the first round are located in the same column. Furthermore, the differential propagations of the form $\alpha_0\|\alpha_1\|\alpha_2\|\alpha_3 \xrightarrow{\mathtt{g}_0} \beta_0\|\beta_1\|\beta_2\|\beta_3 \xrightarrow{GS} \gamma_0\|\gamma_1\|\gamma_2\|\gamma_3$ abstracted from these 2-round characteristics fulfil Condition 1 – 4.

On the contrary, consider two differential propagations validating Condition 1 – 4,

$$\alpha_0\Vert\alpha_1\Vert\alpha_2\Vert\alpha_3 \xrightarrow{\mathtt{g}_0} \beta_0\Vert\beta_1\Vert\beta_2\Vert\beta_3 \xrightarrow{GS} \gamma_0\Vert\gamma_1\Vert\gamma_2\Vert\gamma_3 \ ,$$
$$\alpha'_0\Vert\alpha'_1\Vert\alpha'_2\Vert\alpha'_3 \xrightarrow{\mathtt{g}_0} \beta'_0\Vert\beta'_1\Vert\beta'_2\Vert\beta'_3 \xrightarrow{GS} \gamma'_0\Vert\gamma'_1\Vert\gamma'_2\Vert\gamma'_3 \ ,$$

if the positions of nonzero nibbles in $\gamma_0\Vert\gamma_1\Vert\gamma_2\Vert\gamma_3$ and $\alpha'_0\Vert\alpha'_1\Vert\alpha'_2\Vert\alpha'_3$ are the same, and $\gamma_i \xrightarrow{GS} \alpha'_i$ are possible transitions for all $0 \leqslant i \leqslant 3$, then the two propagations are said to be *compatible* with each other. As shown in Fig. 4, we can craft long differential characteristics that activate two S-boxes per round with compatible propagations.

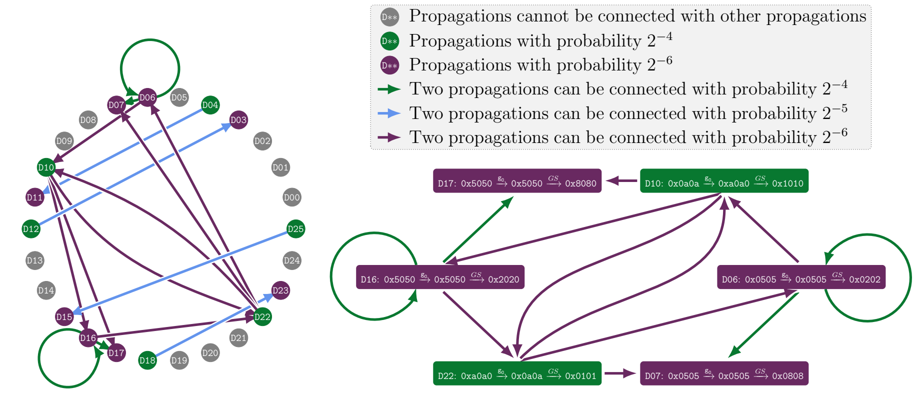
As a result, to figure out the structure of the differential characteristic activating two S-boxes in each round, we should find out all possible propagations of the form $\alpha_0\Vert\alpha_1\Vert\alpha_2\Vert\alpha_3 \xrightarrow{\mathtt{g}_0} \beta_0\Vert\beta_1\Vert\beta_2\Vert\beta_3 \xrightarrow{GS} \gamma_0\Vert\gamma_1\Vert\gamma_2\Vert\gamma_3$. We implement a test and find that 26 propagations validate Condition 1 – 4 simultaneously (cf. Table 1). Then, we evaluate the compatibilities among them and illustrate the result in Fig. 5(a).

The graph in Fig. 5(a) contains some isolated nodes and short paths, and the corresponding propagations cannot be manipulated to create long differential characteristics. Thus, we remove these nodes and picture a more succinct graph as in Fig. 5(b), which manifests several cycles. On the one hand, these cycles theoretically explain the existence of long differential characteristics with two active S-boxes per round. On the other hand, accompanied by the preceding analyses, we conclude that any differential characteristics covering more than seven rounds with two active S-boxes per round must utilise certain paths in Fig. 5(b).

In particular, from Fig. 5(b), we identify three categories of 4-round iterative differential characteristics with probability 2^{-20}, which are demonstrated in Fig. 6. Note that the three categories cover the eight 4-round iterative differential characteristics with probability 2^{-20} proposed in [35].

Table 1. Candidate propagations $\alpha_0\Vert\alpha_1\Vert\alpha_2\Vert\alpha_3 \xrightarrow{\mathtt{g}_0} \beta_0\Vert\beta_1\Vert\beta_2\Vert\beta_3 \xrightarrow{GS} \gamma_0\Vert\gamma_1\Vert\gamma_2\Vert\gamma_3$.

Index	$\alpha_0\Vert\alpha_1\Vert\alpha_2\Vert\alpha_3 \xrightarrow{\mathtt{g}_0} \beta_0\Vert\beta_1\Vert\beta_2\Vert\beta_3 \xrightarrow{GS} \gamma_0\Vert\gamma_1\Vert\gamma_2\Vert\gamma_3$	Probability	Index	$\alpha_0\Vert\alpha_1\Vert\alpha_2\Vert\alpha_3 \xrightarrow{\mathtt{g}_0} \beta_0\Vert\beta_1\Vert\beta_2\Vert\beta_3 \xrightarrow{GS} \gamma_0\Vert\gamma_1\Vert\gamma_2\Vert\gamma_3$	Probability
D00	0x0039 $\xrightarrow{\mathtt{g}_0}$ 0x9003 $\xrightarrow{GS}$ 0x8008	2^{-6}	D13	0x3900 $\xrightarrow{\mathtt{g}_0}$ 0x0390 $\xrightarrow{GS}$ 0x0880	2^{-6}
D01	0x0085 $\xrightarrow{\mathtt{g}_0}$ 0x0c01 $\xrightarrow{GS}$ 0x0808	2^{-6}	D14	0x5008 $\xrightarrow{\mathtt{g}_0}$ 0xc010 $\xrightarrow{GS}$ 0x8080	2^{-6}
D02	0x009c $\xrightarrow{\mathtt{g}_0}$ 0x9c00 $\xrightarrow{GS}$ 0x8800	2^{-6}	D15	0x500a $\xrightarrow{\mathtt{g}_0}$ 0xc030 $\xrightarrow{GS}$ 0x8080	2^{-6}
D03	0x00a5 $\xrightarrow{\mathtt{g}_0}$ 0x0c03 $\xrightarrow{GS}$ 0x0808	2^{-6}	D16	0x5050 $\xrightarrow{\mathtt{g}_0}$ 0x5050 $\xrightarrow{GS}$ 0x2020	2^{-6}
D04	0x00c6 $\xrightarrow{\mathtt{g}_0}$ 0x0c60 $\xrightarrow{GS}$ 0x0220	2^{-4}	D17	0x5050 $\xrightarrow{\mathtt{g}_0}$ 0x5050 $\xrightarrow{GS}$ 0x8080	2^{-6}
D05	0x0390 $\xrightarrow{\mathtt{g}_0}$ 0x3900 $\xrightarrow{GS}$ 0x8800	2^{-6}	D18	0x600c $\xrightarrow{\mathtt{g}_0}$ 0xc600 $\xrightarrow{GS}$ 0x2200	2^{-4}
D06	0x0505 $\xrightarrow{\mathtt{g}_0}$ 0x0505 $\xrightarrow{GS}$ 0x0202	2^{-6}	D19	0x8500 $\xrightarrow{\mathtt{g}_0}$ 0x010c $\xrightarrow{GS}$ 0x0808	2^{-6}
D07	0x0505 $\xrightarrow{\mathtt{g}_0}$ 0x0505 $\xrightarrow{GS}$ 0x0808	2^{-6}	D20	0x9003 $\xrightarrow{\mathtt{g}_0}$ 0x0039 $\xrightarrow{GS}$ 0x0088	2^{-6}
D08	0x0850 $\xrightarrow{\mathtt{g}_0}$ 0x10c0 $\xrightarrow{GS}$ 0x8080	2^{-6}	D21	0x9c00 $\xrightarrow{\mathtt{g}_0}$ 0x009c $\xrightarrow{GS}$ 0x0088	2^{-6}
D09	0x09c0 $\xrightarrow{\mathtt{g}_0}$ 0x09c0 $\xrightarrow{GS}$ 0x0880	2^{-6}	D22	0xa0a0 $\xrightarrow{\mathtt{g}_0}$ 0x0a0a $\xrightarrow{GS}$ 0x0101	2^{-4}
D10	0x0a0a $\xrightarrow{\mathtt{g}_0}$ 0xa0a0 $\xrightarrow{GS}$ 0x1010	2^{-4}	D23	0xa500 $\xrightarrow{\mathtt{g}_0}$ 0x030c $\xrightarrow{GS}$ 0x0808	2^{-6}
D11	0x0a50 $\xrightarrow{\mathtt{g}_0}$ 0x30c0 $\xrightarrow{GS}$ 0x8080	2^{-6}	D24	0xc009 $\xrightarrow{\mathtt{g}_0}$ 0xc009 $\xrightarrow{GS}$ 0x8008	2^{-6}
D12	0x0c60 $\xrightarrow{\mathtt{g}_0}$ 0x00c6 $\xrightarrow{GS}$ 0x0022	2^{-4}	D25	0xc600 $\xrightarrow{\mathtt{g}_0}$ 0x600c $\xrightarrow{GS}$ 0x2002	2^{-4}


D** Propagations cannot be connected with other propagations
D** Propagations with probability 2^{-4}
D** Propagations with probability 2^{-6}
Two propagations can be connected with probability 2^{-4}
Two propagations can be connected with probability 2^{-5}
Two propagations can be connected with probability 2^{-6}
D00
D01
D02
D03
D04
D05
D06
D07
D08
D09
D10
D11
D12
D13
D14
D15
D16
D17
D18
D19
D20
D21
D22
D23
D24
D25
D17: 0x5050 $\xrightarrow{g_0}$ 0x5050 $\xrightarrow{GS}$ 0x8080
D10: 0x0a0a $\xrightarrow{g_0}$ 0xa0a0 $\xrightarrow{GS}$ 0x1010
D16: 0x5050 $\xrightarrow{g_0}$ 0x5050 $\xrightarrow{GS}$ 0x2020
D06: 0x0505 $\xrightarrow{g_0}$ 0x0505 $\xrightarrow{GS}$ 0x0202
D22: 0xa0a0 $\xrightarrow{g_0}$ 0x0a0a $\xrightarrow{GS}$ 0x0101
D07: 0x0505 $\xrightarrow{g_0}$ 0x0505 $\xrightarrow{GS}$ 0x0808

(a) Graph for all 26 propagations.

(b) Six propagations participating in cycles.

Fig. 5. Compatibilities among 26 candidate differential propagations.

3.6 Enumerating All Optimal Differential Characteristics

This section reveals that all optimal r-round differential characteristics ($r \geqslant 8$) with the maximum probabilities can be created with the two cycles '`D06` $\to$ `D06` $\to$ `D06`' and '`D16` $\to$ `D16` $\to$ `D16`' in Fig. 5(b).

In Fig. 1, we note that the probability of r-round optimal differential characteristics with $r \geqslant 8$ satisfies the following equation

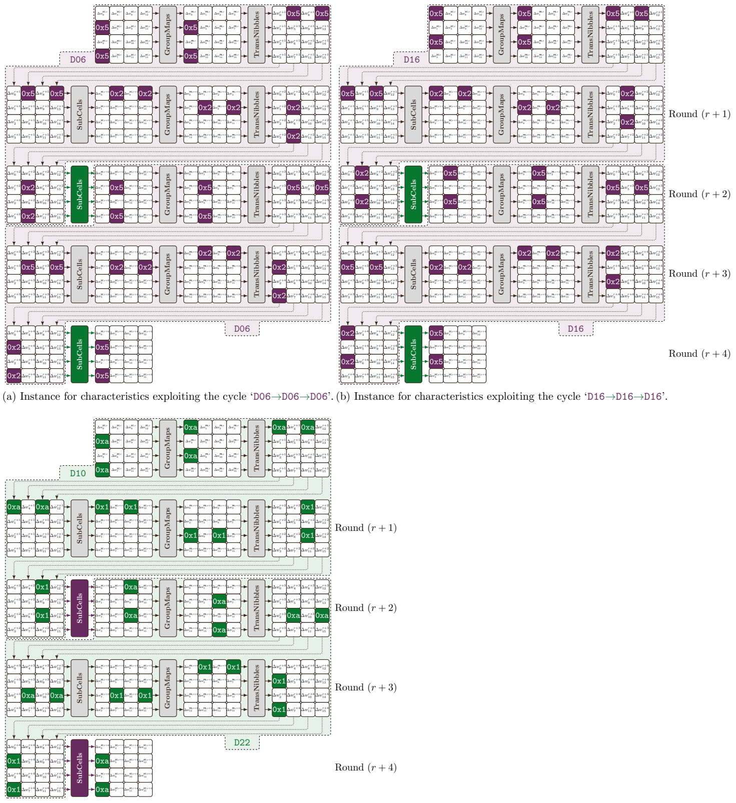
$$-\log_2\left(\Pr(r)\right) = \begin{cases} [(r-3)/2]\cdot 10 + 12 & \text{if } r \bmod 2 \equiv 1, \\ [(r-2)/2]\cdot 10 + 8 & \text{otherwise.} \end{cases}$$

which is a linear function when the independent variable r is restricted to even or odd numbers. The two restrictions of the function have a slope of 5 ($= 10/2$). Meanwhile, for all 4-round iterative differential characteristics in Fig. 6, the probabilities of any two consecutive rounds of characteristics are 2^{-10}. Prompted by these two observations, we attempt to construct optimal differential characteristics with cycles of propagations in Fig. 5.

If we only apply the six differential propagations in Fig. 5(b) to compose characteristics, the maximum probability $\underline{\Pr}(r)$ obtained in this case can be calculated via the following formula

$$-\log_2\left(\underline{\Pr}(r)\right) = \begin{cases} [(r-1)/2]\cdot 10 + 4 & \text{if } r \bmod 2 \equiv 1, \\ (r/2)\cdot 10 & \text{otherwise.} \end{cases}$$

It can be verified that $\underline{\Pr}(r) = \Pr(r) \cdot 2^{-2}$ for all $r \geqslant 8$. To rectify this gap, we fine-tune the head and(or) the tail of the characteristics generated with the cycles and devise numerous characteristics achieving the optimal probability. The adjustment differs depending on the number r of rounds.



(a) Instance for characteristics exploiting the cycle 'D06→D06→D06'. (b) Instance for characteristics exploiting the cycle 'D16→D16→D16'.

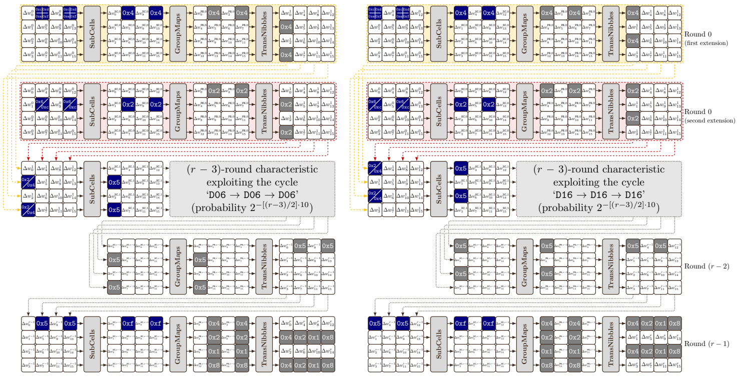
(c) Instance for characteristics exploiting the cycle 'D10→D22→D10'.

Fig. 6. Three categories of 4-round iterative differential characteristics with probability 2^{-20}. In each category, more characteristics can be created by cyclically shifting the columns/rows of the differences for the internal states.

288 optimal characteristics with an odd number of rounds. If $r \bmod 2 \equiv 1$, we can formulate two categories of optimal differential characteristics with the probability being $2^{-\{[(r-3)/2]\cdot 10+12\}}$. As in Fig. 7, the first category is based on the cycle '`D06` → `D06` → `D06`', while the second category iteratively utilises the cycle '`D16` → `D16` → `D16`'. In both categories, to lift the differential probability in the last round, the differential propagations of the two active S-boxes are replaced from `0x5` $\xrightarrow{GS}$ `0x2` to `0x5` $\xrightarrow{GS}$ `0xf`. Also, at the head of the characteristic, we devise two kinds of extensions and ensure that the probabilities of the four active S-boxes in the first two rounds are all equal to 2^{-2}. Each category is composed of 144 characteristics. Thus, in total, we manually identify 288 optimal characteristics.

10400 optimal characteristics with an even number of rounds. If $r \bmod 2 \equiv 0$, we construct four categories of optimal differential characteristics with probability

$2^{-\{\lceil (r-2)/2\rceil \cdot 10+8\}}$. The number of characteristics is 10400. For more details, see Supplementary Material B.4 of the long version.



(a) Instance for characteristics in the first category. (b) Instance for characteristics in the second category.

Fig. 7. 288 optimal characteristics with an odd number of rounds. In each category, more characteristics can be created by cyclically shifting the columns/rows of the differences for the internal states.

We utilise the automatic searching method to find all optimal characteristics with the maximum probability. The experimental results reflect that the manually created characteristics constitute all the optimal characteristics for `GIFT-64`. That is, we know the looks of all optimal differential characteristics for `GIFT-64`.

Last but not least, the cycle '`D10` → `D22` → `D10`' cannot be used to construct optimal characteristics, although it can be employed to create 4-round iterative characteristics. We explain this with the case illustrated in Fig. 7(a). Note that the extension at the head of the characteristic should ensure that the four active S-boxes in the first two rounds have a differential probability being equal to 2^{-2}. On the other side, the two nonzero nibbles in the input difference of `D10` are equal to `0xa`. It can be observed from the DDT of GS that the probabilities of all possible transitions with `0xa` as the output difference are equal to 2^{-3}, which explains why we cannot create optimal characteristics with this cycle.

4 Linear Property of `GIFT-64`

In parallel to the case of differential setting investigated in Section 3, we derive some in-depth analytic results in the linear setting.

4.1 Fluctuant Bounds in Linear Cryptanalysis Setting

Since we wonder about the performances of linear characteristics with a single active S-box in some rounds, we apply the same method in Section 3.2 to determine the lower bound on the number of linear active S-boxes of these characteristics. Denote $\texttt{\#SL}_{\texttt{i}}(r)$ the minimum number of active S-boxes for r-round linear characteristics with at least one round activating a single S-box with the input mask $\texttt{i}$, where $\texttt{i} \in \mathbb{F}_2^4 \setminus \{\texttt{0x0}\}$. The corresponding test results are given in Supplementary Material C.1 of the long version. It can be noticed that for all $\texttt{i} \in \mathbb{F}_2^4 \setminus \{\texttt{0x0}\}$, $\texttt{\#SL}_{\texttt{i}}(r)$ diverges from the initial bound $\texttt{\#SL}(r)$ from the tenth round. So, we introduce the following proposition.

Proposition 4. If $r \geqslant 10$, then the optimal r-round linear characteristic of GIFT-64 with the minimum number of active S-boxes must activate two S-boxes per round.

Next, the linear correlation bound is studied. Denote $\mathrm{Cor}_{\texttt{i}}(r)$ the maximum linear correlation for r-round characteristics with at least one round activating a single S-box with the input mask $\texttt{i}$, where $\texttt{i} \in \mathbb{F}_2^4 \setminus \{\texttt{0x0}\}$. The test result about $\mathrm{Cor}_{\texttt{i}}(r)$ can be found in Supplementary Material C.2 of the long version. It can be noticed that some points of curves for $\mathrm{Cor}_{\texttt{0x1}}(r)$, $\mathrm{Cor}_{\texttt{0x2}}(r)$, $\mathrm{Cor}_{\texttt{0x8}}(r)$, $\mathrm{Cor}_{\texttt{0xa}}(r)$, $\mathrm{Cor}_{\texttt{0xc}}(r)$ overlap with those of the curve for $\mathrm{Cor}(r)$ when $r \geqslant 10$. Thus, unlike the case in differential setting, the optimal linear characteristic with the maximum correlation can contain characteristics with a single active S-box in some rounds.

4.2 Linear Characteristics with Two Active S-boxes Per Round

It can be observed from Fig. 1 that the minimum number of linear active S-boxes $\texttt{\#SL}(r)$ is also linearly dependent on r for all $r \geqslant 9$. Hence, we adjust the approach in Section 3.5 to the linear setting and look into properties of linear characteristics with two active S-boxes in each round. The ideas to prove lemmas and conditions in this section are similar to those in Section 3.5, and we omit proofs.

Lemma 2. For GIFT-64, if a linear characteristic has two active S-boxes per round, then the two active S-boxes in one of the first two rounds must be located in the same column of the matrix state.

Given a linear characteristic activating two S-boxes per round, suppose that the two active S-boxes in the r-th round are located in the same column. Also, without loss of generality, the column is fixed as the first one. Let $\zeta_0\|\zeta_1\|\zeta_2\|\zeta_3 \xrightarrow{\mathsf{g}_0} \eta_0\|\eta_1\|\eta_2\|\eta_3$ be the linear propagation of the group mapping g_0 in the r-th round operating on the first column. Two nibbles in the vector $\zeta_0\|\zeta_1\|\zeta_2\|\zeta_3$ are nonzero. Then, the propagation should satisfy the following conditions so that the linear characteristic exploiting it keeps two active S-boxes in rounds $(r-1)$ and $(r+1)$.

Condition 5 The output mask $\eta_0\|\eta_1\|\eta_2\|\eta_3$ of g_0 has two nonzero nibbles.

Condition 6 Two nonzero nibbles in $\eta_0\|\eta_1\|\eta_2\|\eta_3$ cannot take values from the set $\{\texttt{0x4}, \texttt{0x8}\}$.

Condition 7 Two nonzero nibbles in $\zeta_0\|\zeta_1\|\zeta_2\|\zeta_3$ cannot take values from the set $\{\texttt{0x1}, \texttt{0x2}\}$.

Condition 8 Let η_i and η_j be the two nonzero nibbles in $\eta_0\|\eta_1\|\eta_2\|\eta_3$, where $i, j \in \{0, 1, 2, 3\}$ and $i \neq j$. Define two sets $\mathcal{S}_i^L$ and $\mathcal{S}_j^L$ as

$$\mathcal{S}_i^L = \{\lambda_i \mid \eta_i \xrightarrow{GS} \lambda_i \text{ is a possible propagation , and } \lambda_i \text{ is a unit vector}\} \ ,$$
$$\mathcal{S}_j^L = \{\lambda_j \mid \eta_j \xrightarrow{GS} \lambda_j \text{ is a possible propagation , and } \lambda_j \text{ is a unit vector}\} \ .$$

Then, $\mathcal{S}_i^L \cap \mathcal{S}_j^L \neq \emptyset$ must hold.

A dual proposition of Proposition 3 can now be formulated.

Proposition 5. For an R-round linear characteristic with two active S-boxes per round, if the two active S-boxes in the r-th round are located in the same column, the two active S-boxes in the $(r + 2 \cdot i)$-th round are also located in the same column for all i with $0 \leqslant r + 2 \cdot i < R$.

We find that 46 propagations of the form $\zeta_0\|\zeta_1\|\zeta_2\|\zeta_3 \xrightarrow{\mathtt{g}_0} \eta_0\|\eta_1\|\eta_2\|\eta_3 \xrightarrow{GS} \lambda_0\|\lambda_1\|\lambda_2\|\lambda_3$ satisfy Condition 5 – 8, which are listed in Table 6 of Supplementary Material C.3 of the long version. The compatibilities among the 46 candidates are demonstrated in Fig. 17 of Supplementary Material C.4 of the long version. Based on the cycle in the graph, we also theoretically explain the existence of long linear characteristics with two active S-boxes per round.

5 Can We Improve GIFT-64?

The results in Fig. 1 reflect that the differential and linear properties of GIFT-64 are comparable if we only consider the number of differential and linear active S-boxes. However, when it comes to the differential probability and the linear correlation, the resistance of the cipher regarding these two cryptanalytic methods is inconsistent. In particular, the longest effective differential characteristics with probability greater than 2^{-64} covers 13 rounds, while the longest effective linear characteristics covers 12 rounds. We also notice that besides the group mapping $\mathtt{g}_0$ applied in GIFT-64, numerous candidates validate BOGI requirement and the four rules in Section 2.2. So, we wonder whether we can find a variant of GIFT-64 constructed with a new group mapping that possesses comparable upper bounds on the differential probability and the linear correlation. The content in this section constitutes our answer to this question.

5.1 Candidate Variants

Among the 24 permutations over the set $\{0, 1, 2, 3\}$, four permutations are BOGI permutations. After taking the four rules in Section 2.2 into consideration, we can generate 2304 group mappings (including $\mathtt{g}_0$ in GIFT-64) meeting all requirements for the one in GIFT-64. We call the corresponding 2303 candidate variants, constructed with the 2303 group mappings, GIFT-64-like ciphers. In theory, we should evaluate the differential and linear properties of 2303 variants.

Note that the nibble-oriented description in Section 3.4 for GIFT-64 can be used to represent GIFT-64-like ciphers, and the unique modification lies in the group mapping exploited in GroupMaps operation. The GroupMaps operation based on the group mapping $\mathtt{g}$ is denoted as $\mathtt{GM}_{\mathtt{g}}$, and we use $\mathbb{GM}$ to stand for the set of 2304 GroupMaps operations in GIFT-64-like ciphers.

5.2 Classifying the Variants of GIFT-64

With the alternative description in Section 3.4, we are able to create a sufficient condition for two GIFT-64-like ciphers to be equivalent to each other. We start by introducing two special categories of linear transformations over the 4×4 matrix of nibbles, which will be utilised to derive the sufficient condition.

Definition 1 (Row Transformation). Let $\mathbb{P}$ be the set of all permutations over the set $\{0, 1, 2, 3\}$. Given ϱ in $\mathbb{P}$, the *row transformation generated with* ϱ, denoted by $\mathtt{RT}_\varrho$, is a permutation over the 4×4 matrix that transfers the i-th row of the input to the $\varrho(i)$-th row for all $0 \leqslant i \leqslant 3$.

Definition 2 (Column Transformation). Given ϱ in $\mathbb{P}$, the *column transformation generated with* ϱ, denoted by $\mathtt{CT}_\varrho$, is a permutation over the 4×4 matrix that shifts the i-th column of the input to the $\varrho(i)$-th column for all $0 \leqslant i \leqslant 3$.

With the simple definitions of the two kinds of transformations, we can quickly write their inverse operations.

Lemma 3. If $\varrho \in \mathbb{P}$ and ϱ^{-1} is the inverse permutation of ϱ, then the inverse operation of $\mathtt{RT}_\varrho$ is $\mathtt{RT}_{\varrho^{-1}}$. In symbols, $(\mathtt{RT}_\varrho)^{-1} = \mathtt{RT}_{\varrho^{-1}}$. Likewise, $\mathtt{CT}_\varrho$ and $\mathtt{CT}_{\varrho^{-1}}$ are inverse of each other.

Because the row and column transformations only involve permutations over rows and columns of the input matrix, the composition of these two categories of transformations is commutative.

Lemma 4. If ϱ_1 and $\varrho_2 \in \mathbb{P}$, then $\mathtt{RT}_{\varrho_1} \circ \mathtt{CT}_{\varrho_2} = \mathtt{CT}_{\varrho_2} \circ \mathtt{RT}_{\varrho_1}$.

To establish the equivalence among GIFT-64-like ciphers, we also investigate the commutativity of the composition between these artificial transformations and the operations in the round function of the GIFT-64-like cipher.

As $\mathtt{RT}_\varrho$ and $\mathtt{CT}_\varrho$ do not change the values of the entries in the input matrix, the composition between $\mathtt{RT}_\varrho/\mathtt{CT}_\varrho$ and SubCells operation is commutative.

Lemma 5. If $\varrho \in \mathbb{P}$, then $\mathtt{RT}_\varrho \circ \mathtt{SC} = \mathtt{SC} \circ \mathtt{RT}_\varrho$ and $\mathtt{CT}_\varrho \circ \mathtt{SC} = \mathtt{SC} \circ \mathtt{CT}_\varrho$.

Since the column transformation $\mathtt{CT}_\varrho$ only alter the positions of the columns and do not touch on any permutations within columns, the composition between $\mathtt{CT}_\varrho$ and GroupMaps operation is commutative.

Lemma 6. If $\varrho \in \mathbb{P}$ and $\mathtt{GM}_{\mathsf{g}} \in \mathbb{GM}$, then $\mathtt{CT}_{\varrho} \circ \mathtt{GM}_{\mathsf{g}} = \mathtt{GM}_{\mathsf{g}} \circ \mathtt{CT}_{\varrho}$.

Note that $\mathtt{RT}_{\varrho}$ and $\mathtt{CT}_{\varrho}$ apply the same permutation ϱ to realise the diffusion of the input matrix in the vertical and horizontal directions, respectively. Recall that TransNibbles operation over the input matrix works like a transposition. Taken together, we obtain the following lemma.

Lemma 7. If $\varrho \in \mathbb{P}$, then $\mathtt{RT}_{\varrho} \circ \mathtt{TN} = \mathtt{TN} \circ \mathtt{CT}_{\varrho}$ and $\mathtt{CT}_{\varrho} \circ \mathtt{TN} = \mathtt{TN} \circ \mathtt{RT}_{\varrho}$.

Under a permutation of the round keys, the commutativity of the composition between $\mathtt{RT}_{\varrho}/\mathtt{CT}_{\varrho}$ and AddRoundKey operation can be constructed.

Lemma 8. If $\varrho \in \mathbb{P}$ and $k \in \left(\mathbb{F}_2^4\right)^{4\times 4}$, then $\mathtt{RT}_{\varrho} \circ \mathtt{ARK}_k = \mathtt{ARK}_{\mathtt{RT}_{\varrho}(k)} \circ \mathtt{RT}_{\varrho}$ and $\mathtt{CT}_{\varrho} \circ \mathtt{ARK}_k = \mathtt{ARK}_{\mathtt{CT}_{\varrho}(k)} \circ \mathtt{CT}_{\varrho}$.

For simplicity, denote the r-th round function of a `GIFT-64`-like cipher with the group mapping g as $\mathcal{F}(\mathsf{g}, k_r)$, i.e., $\mathcal{F}(\mathsf{g}, k_r) = \mathtt{ARK}_{k_r} \circ \mathtt{TN} \circ \mathtt{GM}_{\mathsf{g}} \circ \mathtt{SC}$. Note that the following result is an easy consequence by combining all properties of row and column transformations in Lemma 4 – 8.

Proposition 6. If $\varrho \in \mathbb{P}$, then $\mathtt{RT}_{\varrho} \circ \mathcal{F}(\mathsf{g}, k_r) = \mathcal{F}\big(\mathsf{g}, \mathtt{RT}_{\varrho}(k_r)\big) \circ \mathtt{CT}_{\varrho}$.

The following proposition points out a sufficient condition for two `GIFT-64`-like ciphers being equivalent to each other.

Proposition 7. Let $\mathtt{GIFT\text{-}64}[\mathsf{g}_1]$ and $\mathtt{GIFT\text{-}64}[\mathsf{g}_2]$ be two `GIFT-64`-like ciphers respectively instantiated with group mappings g_1 and g_2. If there exists an element $\varrho \in \mathbb{P}$ such that $\mathtt{GM}_{\mathsf{g}_2} = \mathtt{RT}_{\varrho} \circ \mathtt{GM}_{\mathsf{g}_1} \circ \mathtt{RT}_{\varrho^{-1}}$, then $\mathtt{GIFT\text{-}64}[\mathsf{g}_1]$ and $\mathtt{GIFT\text{-}64}[\mathsf{g}_2]$ differ only by a permutation on the plaintext and ciphertext and a corresponding permutation of the round keys.

For the proofs of Proposition 6 – 7, see Supplementary Material D of the long version.

Definition 3 ($\mathbb{GM}$-equivalence). Given two elements $\mathtt{GM}_{\mathsf{g}_1}$ and $\mathtt{GM}_{\mathsf{g}_2}$ of the set $\mathbb{GM}$, $\mathtt{GM}_{\mathsf{g}_1}$ and $\mathtt{GM}_{\mathsf{g}_2}$ are called $\mathbb{GM}$-equivalence, if there exists a $\varrho \in \mathbb{P}$ such that $\mathtt{GM}_{\mathsf{g}_2} = \mathtt{RT}_{\varrho} \circ \mathtt{GM}_{\mathsf{g}_1} \circ \mathtt{RT}_{\varrho^{-1}}$. In symbols, $\mathtt{GM}_{\mathsf{g}_1} \sim \mathtt{GM}_{\mathsf{g}_2}$.

It can be verified that the binary relation '$\sim$' on the set $\mathbb{GM}$ is reflexive, symmetric and transitive. Hence, '$\sim$' is an equivalence relation on $\mathbb{GM}$. Because of the conclusion in Proposition 7, if $\mathtt{GM}_{\mathsf{g}_1}$ and $\mathtt{GM}_{\mathsf{g}_2}$ are $\mathbb{GM}$-equivalent permutations, the two `GIFT-64`-like ciphers implemented with $\mathtt{GM}_{\mathsf{g}_1}$ and $\mathtt{GM}_{\mathsf{g}_2}$ share the same cryptographic properties. In particular, this fact holds for the case of differential and linear cryptanalyses.

We classify all permutations in $\mathbb{GM}$ up to $\mathbb{GM}$-equivalence and split the set $\mathbb{GM}$ into 168 distinct equivalence classes. Accordingly, the set of 2304 `GIFT-64`-like ciphers is partitioned into 168 equivalence classes. Therefore, we only need to check the property of one representative in each possible equivalence class, and the number of candidates is reduced from 2303 to 167. Note that we do not count in the equivalence class containing `GIFT-64`.

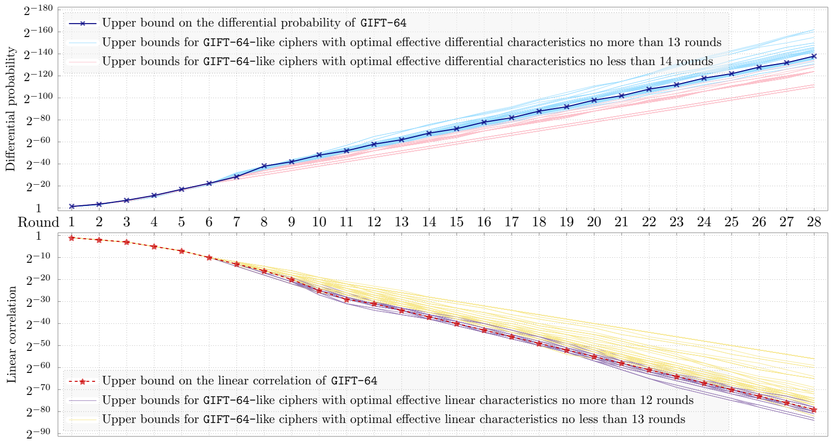


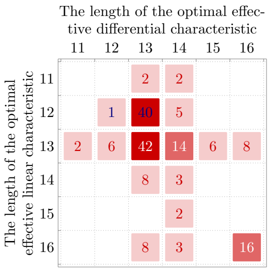
Fig. 8. Test results for 167 representatives.

5.3 Differential and Linear Properties of GIFT-64-like Ciphers

We apply the accelerated automatic method to search for upper bounds on differential probabilities and linear correlations of 167 representative variants. The test results are illustrated in Fig. 8. As the longest effective differential characteristics of GIFT-64 achieve 13 rounds, we split all the 168 representatives into two groups, according to whether the length of the optimal effective differential characteristic is longer than 13 rounds. To make a distinction, in Fig. 8, we use blue curves to exhibit variants with optimal effective differential characteristics no more than 13 rounds. For variants with effective differential characteristics covering more than 13 rounds, the differential probability curves are coloured in red. Since a cipher with short effective differential characteristics is more likely to withstand a differential attack, then, we conclude from Fig. 8 that the security of GIFT-64 against the differential cryptanalysis is moderate among all the 168 representatives.

Similarly, in the linear setting, the 168 representatives are classified according to whether the optimal effective linear characteristic goes beyond 12 rounds. In Fig. 8, the purple curves correspond to GIFT-64-like ciphers with optimal effective linear characteristics no more than 12 rounds, while the variants with yellow curves have longer effective linear characteristics than that of GIFT-64. Unlike the case in differential cryptanalysis, the capability of GIFT-64 against linear cryptanalysis is almost among the best of candidates.

Then, we consider the combination of differential and linear properties. According to the lengths of the optimal effective differential and linear characteristics, the 168 representatives can be divided into 17 groups, and the results can be found in Fig. 9. It can be notified that the performance of GIFT-64 resisting differential and linear attacks is good, and 40 representatives achieve similar security levels to GIFT-64. Moreover, we identify that one representative may



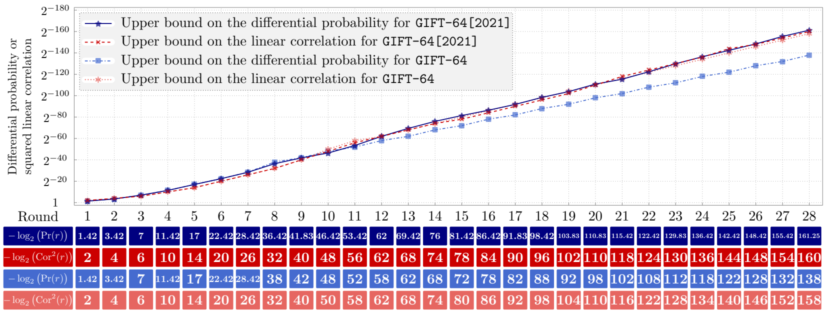
	11	12	13	14	15	16
11			2	2		
12		1	40	5		
13	2	6	42	14	6	8
14			8	3		
15				2		
16			8	3		16

Fig. 9. Heatmap for the number of representatives with different properties.

possess comparable security levels against differential and linear cryptanalyses, and its optimal effective differential and linear characteristics achieve 12 rounds. For simplicity, the equivalence class containing this representative is denoted as `GIFT-64[2021]`. Next, we discuss the cryptanalytic properties of `GIFT-64`-like ciphers in `GIFT-64[2021]`.

5.4 Properties of Variants in `GIFT-64[2021]`

The equivalence class `GIFT-64[2021]` contains 24 elements, and 24 underlying group mappings can be found in Table 7 of Supplementary Material E.1 of the long version. All variants belonging to `GIFT-64[2021]` share the same differential and linear properties, which are illustrated in Fig. 10. The clustering effects of differential and linear characteristics are evaluated (cf. Supplementary Material E.3 – E.4 of the long version). Similarly to the case of `GIFT-64`, the differential and linear hull properties of `GIFT-64[2021]` are not significant.



Round	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
$-\log_2(\Pr(r))$	1.42	3.42	7	11.42	17	22.42	28.42	36.42	41.83	46.42	53.42	62	69.42	76	81.42	86.42	91.83	98.42	103.83	110.83	115.42	122.42	129.83	136.42	142.42	148.42	155.42	161.25
$-\log_2(\mathrm{Cor}^2(r))$	2	4	6	10	14	20	26	32	40	48	56	62	68	74	78	84	90	96	102	110	118	124	130	136	144	148	154	160
$-\log_2(\Pr(r))$	1.42	3.42	7	11.42	17	22.42	28.42	38	42	48	52	58	62	68	72	78	82	88	92	98	102	108	112	118	122	128	132	138
$-\log_2(\mathrm{Cor}^2(r))$	2	4	6	10	14	20	26	32	40	50	58	62	68	74	80	86	92	98	104	110	116	122	128	134	140	146	152	158

Fig. 10. Cryptanalytic properties of `GIFT-64[2021]`. A detailed comparison between `GIFT-64[2021]` and `GIFT-64` is given in Supplementary Material E.2 of the long version.

Beyond that, we implement the automatic search of impossible differential distinguishers [28], zero-correlation linear distinguishers [16,15], and integral distinguishers [33] for the variants belonging to GIFT-64[2021]. The experimental results indicate that the security levels of the variants in GIFT-64[2021] withstanding impossible differential (ID) attack, zero-correlation linear attack, and integral attack are similar to those of GIFT-64.

Table 2. Attack results on GIFT-64 and GIFT-64[$\mathrm{g}_0^{\mathrm{c}}$].

Method	GIFT-64					GIFT-64[$\mathrm{g}_0^{\mathrm{c}}$]			
	Round	Time	Data	Memory	Ref.	Round	Time	Data	Memory
Differential	20	$2^{125.50}$	$2^{62.58}$	$2^{62.58}$	[31]	18[†]	$2^{125.16}$	$2^{62.27}$	$2^{62.27}$
Linear	19	$2^{127.11}$	$2^{62.96}$	$2^{60.00}$	[31]	18[‡]	$2^{126.60}$	$2^{62.96}$	$2^{53.00}$
Integral	14	$2^{97.00}$	$2^{63.00}$	-	[5]	14	$2^{97.00}$	$2^{63.00}$	-
ID[*]	6	-	-	-	[5]	6	-	-	-

[†]: The differential attack is realised with the 12-round differential in Table ??.

[‡]: The linear attack is realised with the 12-round linear hull in Table ??.

[*]: The number of rounds is the length of the distinguisher.

In Table 2, we compare the attack results in the single-key attack setting on GIFT-64 and GIFT-64[$\mathrm{g}_0^{\mathrm{c}}$], which is the representative of GIFT-64[2021] instantiated with the group mapping $\mathrm{g}_0^{\mathrm{c}}$ in Table 7 of the long version. Note that the best attack on GIFT-64[$\mathrm{g}_0^{\mathrm{c}}$] achieves 18 rounds, which is two rounds less than the length of the best attack on GIFT-64. Furthermore, in the design document, the designer of GIFT-64 expected that the differential probability of 14-round differential would be lower than 2^{-63}. For this reason, they believed 28-round GIFT-64 is enough to resist differential cryptanalysis. Taken these observations together, we claim that for the variant GIFT-64[$\mathrm{g}_0^{\mathrm{c}}$], if the security in the related-key attack setting is not required, 26 rounds could be used rather than 28 rounds.

As mentioned by the designers, for the simple and clean design strategy, GIFT offers extremely good performances and even surpasses both SKINNY [7] and SIMON [6] for round-based implementations. On this basis, 26-round GIFT-64[$\mathrm{g}_0^{\mathrm{c}}$] may become one of the most energy-efficient ciphers as of today and is probably more suitable for the low-energy consumption use cases than GIFT-64. In Table 3, we compare the hardware performance of 26-round GIFT-64[$\mathrm{g}_0^{\mathrm{c}}$] with other lightweight ciphers. The new variant achieves higher throughput and requires a lower energy consumption than GIFT-64.

Although GIFT designer did not claim any related-key security, the security of the cipher in the related-key attack setting was investigated in recent years [20,18,31]. We also check the security of the 24 variants in GIFT-64[2021] in the related-key attack setting. The key schedule remains the same as the one in GIFT-64. We test the lower bound on the number of active S-boxes for up to 18 rounds with the accelerated automatic method. Figure 19 of Supplementary Material E.5 of the long version contains the experimental results. For all 24 variants, the number of active S-boxes in the related-key differential attack set-

Table 3. Comparison of performance metrics for round-based implementations synthesised with TSMC 90nm standard cell library.

	Area (GE)	Delay (ns)	Cycle	$\mathrm{TP}_{\mathrm{MAX}}$ (MBit/s)	Power (μW)	Energy (pJ)
GIFT-64[$\mathrm{g}_0^{\mathrm{c}}$]	1769	0.55	26	4475.5	36.7	95.4
GIFT-64	1770	0.56	28	4081.6	36.7	102.7
SKINNY-64-128	1804	0.86	36	2067.2	36.8	132.5
SIMON-64-128	1829	0.81	44	1795.7	36.5	160.5

ting is always lower than that of GIFT-64. Thus, we believe GIFT-64 maintains a relatively good performance against the related-key differential attack, even though the designers do not claim its security in the related-key attack setting.

To sum up, we find a greater GIFT-64, which strengthens GIFT-64 against statistical cryptanalysis. In this sense, a variant GIFT-64[$\mathrm{g}_0^{\mathrm{c}}$] with 26 rounds is created and achieves better performance than GIFT-64. Likewise, we do not claim any related-key security for the new variant since most applications do not need related-key security. For the few applications where this security is required, the key schedule of the variant could be redesigned.

A probable explanation for the improved resistance against the differential cryptanalysis of GIFT-64[$\mathrm{g}_0^{\mathrm{c}}$] is provided in Supplementary Material E.6 of the long version. As we prepare the paper, we notice that Baek et al. [2] also created a variant for GIFT-64. The distinction between [2] and this paper is explained in Supplementary Material E.7 of the long version.

6 Conclusion and Future Work

6.1 Conclusion

This paper targets the cryptanalysis of GIFT-64 and combines automatic and manual methods to evaluate its security. In the differential setting, we theoretically explain the existence of differential characteristics with two active S-boxes per round and derive some properties of these characteristics, apart from the quantitative information about active S-boxes. Furthermore, all optimal differential characteristics covering more than seven rounds are identified. Parallel work is conducted in the linear setting. Considering the gap between the upper bounds on the differential probability and the linear correlation, we study a variant of GIFT-64 with comparable security levels in the differential and linear settings. With the support of automatic searching methods, we identify 24 variants achieving better resistance against differential cryptanalysis than GIFT-64 while maintaining a similar security level against linear cryptanalysis. As the new variants strengthen GIFT-64 against statistical cryptanalysis, we claim that for the variant GIFT-64[$\mathrm{g}_0^{\mathrm{c}}$], if the security in the related-key attack setting is not required, 26 rounds could be used rather than 28 rounds. This observation results in a cipher more suitable for the low-energy consumption use cases than GIFT-64. The performance of the 24 variants in the related-key differential attack setting is inferior to that of GIFT-64. However, most applications do not need related-key security.

6.2 Future Work

If one is concerned with related-key attacks, we conjecture that the resistance of variants in GIFT-64[2021] regarding related-key differential attack can be lifted by carefully crafting the key schedule. However, many parameters should be fine-tuned. Thus, we left it as future work.

Secondly, in the construction of GIFT-64-like cipher, we apply the same 16-bit group mapping to each column of the state. How to efficiently evaluate the cases where the group mappings operating on different columns are distinct is an open problem.

Lastly, for GIFT-128, the security levels regarding differential and linear cryptanalyses are also not comparable. We attempt to create an equivalence relation among all variants for GIFT-128. Nevertheless, the number of equivalence classes is 1344. In addition, due to the considerable state size, investigating the security of the variants for GIFT-128 is much more complicated than that of GIFT-64. Still, considering the significant status of GIFT-128 among the lightweight block ciphers and its supporting role in a series of Authenticated Encryptions with Associated Data (AEADs), especially in one of the finalists GIFT-COFB[3] of NIST Lightweight Cryptography project[8], we believe checking the existence of a balanced variant for GIFT-128 will be interesting future work. For more details about the test of GIFT-128, see Supplementary Material F of the long version.

Acknowledgements. The authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the quality of the paper. The research leading to these results has received funding from the National Natural Science Foundation of China (Grant No. 62002201, Grant No. 62032014), the National Key Research and Development Program of China (Grant No. 2018YFA0704702), and the Major Basic Research Project of Natural Science Foundation of Shandong Province, China (Grant No. ZR202010220025). Bart Preneel was supported by CyberSecurity Research Flanders with reference number VR20192203.

References

1. Adomnicai, A., Najm, Z., Peyrin, T.: Fixslicing: A new GIFT representation fast constant-time implementations of GIFT and GIFT-COFB on ARM cortex-m. IACR Trans. Cryptogr. Hardw. Embed. Syst. **2020**(3), 402–427 (2020). https://doi.org/10.13154/tches.v2020.i3.402-427

[8] https://csrc.nist.gov/Projects/Lightweight-Cryptography

2. Baek, S., Kim, H., Kim, J.: Development and security analysis of GIFT-64-variant that can be efficiently implemented by bit-slice technique. Journal of the Korea Institute of Information Security & Cryptology **30**(3), 349–356 (06 2020)
3. Banik, S., Chakraborti, A., Iwata, T., Minematsu, K., Nandi, M., Peyrin, T., Sasaki, Y., Sim, S.M., Todo, Y.: GIFT-COFB. IACR Cryptol. ePrint Arch. **2020**, 738 (2020), https://eprint.iacr.org/2020/738
4. Banik, S., Pandey, S.K., Peyrin, T., Sasaki, Y., Sim, S.M., Todo, Y.: GIFT: A small present - towards reaching the limit of lightweight encryption. In: Cryptographic Hardware and Embedded Systems - CHES 2017 - 19th International Conference, Taipei, Taiwan, September 25-28, 2017, Proceedings. pp. 321–345 (2017). https://doi.org/10.1007/978-3-319-66787-4_16
5. Banik, S., Pandey, S.K., Peyrin, T., Sim, S.M., Todo, Y., Sasaki, Y.: GIFT: A small present. IACR Cryptol. ePrint Arch. **2017**, 622 (2017), http://eprint.iacr.org/2017/622
6. Beaulieu, R., Shors, D., Smith, J., Treatman-Clark, S., Weeks, B., Wingers, L.: The SIMON and SPECK families of lightweight block ciphers. IACR Cryptol. ePrint Arch. **2013**, 404 (2013)
7. Beierle, C., Jean, J., Kölbl, S., Leander, G., Moradi, A., Peyrin, T., Sasaki, Y., Sasdrich, P., Sim, S.M.: The SKINNY family of block ciphers and its low-latency variant MANTIS. In: Advances in Cryptology - CRYPTO 2016 - 36th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 14-18, 2016, Proceedings, Part II. pp. 123–153 (2016). https://doi.org/10.1007/978-3-662-53008-5_5
8. Biere, A.: CaDiCaL at the SAT Race 2019. In: Heule, M., Järvisalo, M., Suda, M. (eds.) Proc. of SAT Race 2019 – Solver and Benchmark Descriptions. Department of Computer Science Series of Publications B, vol. B-2019-1, pp. 8–9. University of Helsinki (2019)
9. Biham, E., Biryukov, A., Shamir, A.: Cryptanalysis of Skipjack reduced to 31 rounds using impossible differentials. In: Advances in Cryptology - EUROCRYPT '99, International Conference on the Theory and Application of Cryptographic Techniques, Prague, Czech Republic, May 2-6, 1999, Proceeding. pp. 12–23 (1999). https://doi.org/10.1007/3-540-48910-X_2
10. Biham, E., Shamir, A.: Differential cryptanalysis of DES-like cryptosystems. In: Advances in Cryptology - CRYPTO '90, 10th Annual International Cryptology Conference, Santa Barbara, California, USA, August 11-15, 1990, Proceedings. pp. 2–21 (1990). https://doi.org/10.1007/3-540-38424-3_1
11. Bogdanov, A., Knezevic, M., Leander, G., Toz, D., Varici, K., Verbauwhede, I.: SPONGENT: A lightweight hash function. In: Preneel, B., Takagi, T. (eds.) Cryptographic Hardware and Embedded Systems - CHES 2011 - 13th International Workshop, Nara, Japan, September 28 - October 1, 2011. Proceedings. Lecture Notes in Computer Science, vol. 6917, pp. 312–325. Springer (2011). https://doi.org/10.1007/978-3-642-23951-9_21
12. Bogdanov, A., Knudsen, L.R., Leander, G., Paar, C., Poschmann, A., Robshaw, M.J.B., Seurin, Y., Vikkelsoe, C.: PRESENT: An ultra-lightweight block cipher. In: Cryptographic Hardware and Embedded Systems - CHES 2007, 9th International Workshop, Vienna, Austria, September 10-13, 2007, Proceedings. pp. 450–466 (2007). https://doi.org/10.1007/978-3-540-74735-2_31
13. Bogdanov, A., Rijmen, V.: Linear hulls with correlation zero and linear cryptanalysis of block ciphers. Des. Codes Cryptography **70**(3), 369–383 (2014)

14. Cho, J.Y.: Linear cryptanalysis of reduced-round PRESENT. In: Pieprzyk, J. (ed.) Topics in Cryptology - CT-RSA 2010, The Cryptographers' Track at the RSA Conference 2010, San Francisco, CA, USA, March 1-5, 2010. Proceedings. Lecture Notes in Computer Science, vol. 5985, pp. 302–317. Springer (2010). https://doi.org/10.1007/978-3-642-11925-5_21
15. Cui, T., Chen, S., Fu, K., Wang, M., Jia, K.: New automatic tool for finding impossible differentials and zero-correlation linear approximations. Sci. China Inf. Sci. **64**(2) (2021). https://doi.org/10.1007/s11432-018-1506-4
16. Cui, T., Jia, K., Fu, K., Chen, S., Wang, M.: New automatic search tool for impossible differentials and zero-correlation linear approximations. IACR Cryptol. ePrint Arch. **2016**, 689 (2016), http://eprint.iacr.org/2016/689
17. Daemen, J., Peeters, M., Van Assche, G., Rijmen, V.: Nessie proposal: NOEKEON. In: First Open NESSIE Workshop. pp. 213–230 (2000)
18. Dong, X., Qin, L., Sun, S., Wang, X.: Key guessing strategies for linear key-schedule algorithms in rectangle attacks. Cryptology ePrint Archive, Report 2021/856 (2021), https://ia.cr/2021/856
19. Guo, J., Peyrin, T., Poschmann, A.: The PHOTON family of lightweight hash functions. In: Rogaway, P. (ed.) Advances in Cryptology - CRYPTO 2011 - 31st Annual Cryptology Conference, Santa Barbara, CA, USA, August 14-18, 2011. Proceedings. Lecture Notes in Computer Science, vol. 6841, pp. 222–239. Springer (2011). https://doi.org/10.1007/978-3-642-22792-9_13
20. Ji, F., Zhang, W., Zhou, C., Ding, T.: Improved (related-key) differential cryptanalysis on GIFT. IACR Cryptol. ePrint Arch. **2020**, 1242 (2020), https://eprint.iacr.org/2020/1242
21. Knudsen, L.: DEAL-A 128-bit block cipher. complexity **258**(2), 216 (1998)
22. Knudsen, L.R., Wagner, D.A.: Integral cryptanalysis. In: Daemen, J., Rijmen, V. (eds.) Fast Software Encryption, 9th International Workshop, FSE 2002, Leuven, Belgium, February 4-6, 2002, Revised Papers. Lecture Notes in Computer Science, vol. 2365, pp. 112–127. Springer (2002). https://doi.org/10.1007/3-540-45661-9_9
23. Kölbl, S., Leander, G., Tiessen, T.: Observations on the SIMON block cipher family. In: Advances in Cryptology - CRYPTO 2015 - 35th Annual Cryptology Conference, Santa Barbara, CA, USA, August 16-20, 2015, Proceedings, Part I. pp. 161–185 (2015). https://doi.org/10.1007/978-3-662-47989-6_8
24. Matsui, M.: Linear cryptanalysis method for DES cipher. In: Advances in Cryptology - EUROCRYPT '93, Workshop on the Theory and Application of of Cryptographic Techniques, Lofthus, Norway, May 23-27, 1993, Proceedings. pp. 386–397 (1993). https://doi.org/10.1007/3-540-48285-7_33
25. Matsui, M.: On correlation between the order of S-boxes and the strength of DES. In: Advances in Cryptology - EUROCRYPT '94, Workshop on the Theory and Application of Cryptographic Techniques, Perugia, Italy, May 9-12, 1994, Proceedings. pp. 366–375 (1994). https://doi.org/10.1007/BFb0053451
26. Mouha, N., Preneel, B.: Towards finding optimal differential characteristics for ARX: Application to Salsa20. Tech. rep., Cryptology ePrint Archive, Report 2013/328 (2013)
27. Mouha, N., Wang, Q., Gu, D., Preneel, B.: Differential and linear cryptanalysis using mixed-integer linear programming. In: Information Security and Cryptology - 7th International Conference, Inscrypt 2011, Beijing, China, November 30 - December 3, 2011. Revised Selected Papers. pp. 57–76 (2011). https://doi.org/10.1007/978-3-642-34704-7_5

28. Sasaki, Y., Todo, Y.: New impossible differential search tool from design and cryptanalysis aspects - revealing structural properties of several ciphers. In: Advances in Cryptology - EUROCRYPT 2017 - 36th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Paris, France, April 30 - May 4, 2017, Proceedings, Part III. pp. 185–215 (2017). https://doi.org/10.1007/978-3-319-56617-7_7
29. Sinz, C.: Towards an optimal CNF encoding of Boolean cardinality constraints. In: Principles and Practice of Constraint Programming - CP 2005, 11th International Conference, CP 2005, Sitges, Spain, October 1-5, 2005, Proceedings. pp. 827–831 (2005). https://doi.org/10.1007/11564751_73
30. Sun, L., Wang, W., Wang, M.: Accelerating the search of differential and linear characteristics with the SAT method. IACR Trans. Symmetric Cryptol. **2021**(1), 269–315 (2021). https://doi.org/10.46586/tosc.v2021.i1.269-315
31. Sun, L., Wang, W., Wang, M.: Improved attacks on GIFT-64. Cryptology ePrint Archive, Report 2021/1179 (2021), https://ia.cr/2021/1179
32. 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 - 20th International Conference on the Theory and Application of Cryptology and Information Security, Kaoshiung, Taiwan, R.O.C., December 7-11, 2014. Proceedings, Part I. pp. 158–178 (2014). https://doi.org/10.1007/978-3-662-45611-8_9
33. Xiang, Z., Zhang, W., Bao, Z., Lin, D.: Applying MILP method to searching integral distinguishers based on division property for 6 lightweight block ciphers. In: Advances in Cryptology - ASIACRYPT 2016 - 22nd International Conference on the Theory and Application of Cryptology and Information Security, Hanoi, Vietnam, December 4-8, 2016, Proceedings, Part I. pp. 648–678 (2016). https://doi.org/10.1007/978-3-662-53887-6_24
34. Zhang, W., Bao, Z., Lin, D., Rijmen, V., Yang, B., Verbauwhede, I.: RECTANGLE: A bit-slice lightweight block cipher suitable for multiple platforms. Sci. China Inf. Sci. **58**(12), 1–15 (2015). https://doi.org/10.1007/s11432-015-5459-7
35. Zhu, B., Dong, X., Yu, H.: MILP-based differential attack on round-reduced GIFT. In: Matsui, M. (ed.) Topics in Cryptology - CT-RSA 2019 - The Cryptographers' Track at the RSA Conference 2019, San Francisco, CA, USA, March 4-8, 2019, Proceedings. Lecture Notes in Computer Science, vol. 11405, pp. 372–390. Springer (2019). https://doi.org/10.1007/978-3-030-12612-4_19