Evaluating the Security of Block Ciphers Against Zero-correlation Linear Attack in the Distinguishers Aspect*

Xichao Hu[1,3], Yongqiang Li[1,3], Lin Jiao[2], Zhengbin Liu[4] and Mingsheng Wang[1,3]

[1] State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing China
xchao_h@163.com,yongq.lee@gmail.com,wangmingsheng@iie.ac.cn
[2] State Key Laboratory of Cryptology, Beijing, China
jiaolin_jl@126.com
[3] School of Cyber Security, University of Chinese Academy of Sciences, Beijing, China
[4] Science and Technology on Communication Security Laboratory, Chengdu, China
zhengbinliu@126.com

Abstract. Zero-correlation linear attack is a powerful attack of block ciphers, the lower number of rounds (LNR) which no its distinguisher (named zero-correlation linear approximation, ZCLA) exists reflects the ability of a block cipher against the zero-correlation linear attack. However, due to the large search space, showing there are no ZCLAs exist for a given block cipher under a certain number of rounds is a very hard task. Thus, present works can only prove there no ZCLAs exist in a small search space, such as 1-bit/nibble/word input and output active ZCLAs, which still exist very large gaps to show no ZCLAs exist in the whole search space.

In this paper, we propose the meet-in-the-middle method and double-collision method to show there no ZCLAs exist in the whole search space. The basic ideas of those two methods are very simple, but they work very effectively. As a result, we apply those two methods to AES, Midori64, and ARIA, and show that there no ZCLAs exist for 5-round AES without the last Mix-Column layer, 7-round Midori64 without the last Mix-Column layer, and 5-round ARIA without the last linear layer.

As far as we know, our method is the first automatic method that can be used to show there no ZCLAs exist in the whole search space, which can provide sufficient evidence to show the security of a block cipher against the zero-correlation linear attack in the distinguishers aspect, this feature is very useful for designing block ciphers.

Keywords: ZCLAs · Search Tool · Proof Tool · SAT

1 Introduction

Zero-correlation linear cryptanalysis [BR14, BW12, BLNW12] is the variant and generalization of linear cryptanalysis. It has been applied to variour block ciphers, such as AES, CLEFIA, and so on. Those results show that zero-correlation linear cryptanalysis is one of the essential general techniques for analyzing a block cipher. Up to now, resisting the zero-correlation linear cryptanalysis is the criterion of design a block cipher.

ZCLAs are the distinguishers of zero-correlation linear cryptanalysis, it is a linear hull of a given cipher that presents no correlation. The most generic way to construct the ZCLAs

*This paper was submitted to Asiacrypt2022.

is meet-in-middle, that is, a pair of input and output masks is called ZCLA of a function if the input mask cannot propagate to the output mask through this function. Such ZCLAs are the ZCLAs that we studied through this paper. The number of rounds which the ZCLAs covered is closely related to the ability of zero-correlation linear cryptanalysis. On the one hand, in most cases, the best cryptanalysis is based on the distinguishers who has the longest rounds at present. On the other hand, the LNR which no ZCLAs exist indicate the ability of a block cipher against zero-correlation linear cryptanalysis in a sense. Thus, how to find ZCLAs that covers rounds as long as possible and get the LNR which no, ZCLAs exist are the most essential and critical problem in regard to zero-correlation linear cryptanalysis.

In recent years, with the improvement of computer power, cryptanalysts began to use automated methods to search ZCLAs. The original methods, such as U-method, UID-method and WW-method, which can be applied to search ZCLAs, but they cannot consider the detials of Sbox. To solve this problem, some cryptanalysts turn the problem that finding distinguishers into MILP problem, and solve such problem by MILP method. That is, they model the propagation of mask through each operation by inequalities. Then, the propagation of mask through a given number of rounds for a block cipher can be modeled by inequalities. Thus, by invoking the MILP solver such as Gurobi, one can determine whether a given input mask can propagate to a given output mask or not. If not, such input and output mask constructs an ZCLA. In this process, modeling the propagation of mask through **Xor**, **Copy**, and linear layer is relatively easy, and modeling the propagation of mask through Sbox isn't trivial and requires careful thinking. In a word, how to model the propagation of mask through Sbox is the critical problem for searching ZCLAs by MILP method.

In 2016, Cui et al. proposed a MILP-based tool [CJF+16] to search the impossible differentials and ZCLAs for lightweight block ciphers. They use Sun et al.'s method [SHW+14b, SHW+14a] to model the valid propagations of masks (differential) through small Sbox (whose size is no more than 7-bit) by inequalities, thus their tool can search the ZCLAs by considering the details of Sbox. Besides, there exist two works that don't focus on searching ZCLAs, but can be applied to search ZCLAs as well. In 2017, Sasaki and Todo (ST-method) also presented a MILP-based tool [ST17] to search the impossible differentials for block ciphers. For small Sbox, they method in accord with Cui et al.'s method (independently). For large Sbox (whose size is no less than 8-bit), they proposed the *arbitrary mode* creatively, although this mode cannot search ZCLAs by considering the details of Sbox, but it can detect contradictions that occur in linear layers. Later, Abdelkhalek et al. [AST+17] proposed a method to search the optimal differential characteristics for block ciphers with 8-bit S-boxes. In particular, they mdoel the propagation of differentials through Sbox whose size no more than 8-bit by inequalities with the help of third party tool Logic Friday [1].

The application scenarios of previous tools for searching ZCLAs are summarized as follows.

Cui et al.'s method. In their method, they modeled the propagation of masks through Sbox by inequalities with the help of third-party tool sagemath [2], which is computationally infeasible when the size of the S-box is more than 6 bits. Thus, for S-boxes based block ciphers, those two methods are suited for block ciphers whose S-box's size is less than or equal to 6-bit.

ST-method. For block ciphers whose S-box's size is less than or equal to 6 bits, the application scenario of this method is the same as Cui et al.'s method. For block

[1] http://sontrak.com
[2] https://www.sagemath.org

ciphers with large S-box, their method can detect any contradictions which occur in linear operations, but cannot consider the details of Sbox.

Abdelkhalek et al.'s method. In their method, they modeled the valid propagation of masks through S-box by inequalities with the help of third-party tool Logic Friday, Since Logic Friday can handle at most 16 variables only, their method is suited for block ciphers whose sum of the input and output's bits is less than or equal to 16 bits.

In the aspect of evaluating the security of a block cipher against zero-correlation linear cryptanalysis, Cui et al.'s method and ST-method can get the LNR which no ZCLAs exist by limiting the input and output mask are 1-bit/nibble/word active, there still exist a great gap to get the LNR which no ZCLAs exist without such limitation. Besides, some fancy works try to prove the security in theorem, for example, Wang et al. [WJ18] showed that there no exist impossible differential for 5-round AES without the last MC layer. But such works usually associated with the properties of the targe block cipehr, which is not universal. Thus, those works cannot get the LNR which no ZCLAs exist. All in all, although some advanced methods that are used to find new ZCLAs and get the LNR which no ZCLAs exist are proposed, there are still some limitations not resolved.

- Previous methods cannot search the ZCLAs for block ciphers which the sum of the input and output size of S-box is more than 16 bits in the case of considering all the details of S-boxes.

- Previous methods cannot get the LNR which no ZCLAs exist without the limitation of input mask and output mask.

- Previous methods cannot get all ZCLAs without the limitation of input mask and output mask.

Our contribution. In this paper, we propose some methods to search new ZCLAs for block ciphers which the sum of the input and output size of S-box is more than 16 bits in the case of considering all the details of S-boxes, and get the LNR which no ZCLAs exist for SPN structure block ciphers.

In the aspect of deriving new ZCLAs, we give a new sight on the representation of correlation, which allows us to propose a new modeling method for the propagation of masks through S-box directly. Specifically speaking, for an S-box whose input is n-bit, our modeling method allows us to model the valid propagation of masks through S-box by $2^n + 1$ expressions, which can be modeled by the statements in the CVC format of the SAT solver STP[3] easily. With the modeling method for other operations, we design an SAT-based automatic to search ZCLAs for the block ciphers which the sum of the input and output size of S-box is more than 16 bits.

In the aspect of getting LNR which no ZCLAs exist, it is directly to get the LNR which no 1-bit input and output active ZCLAs exist. Moreover, by studying the properties of SPN structure block cipher, we propose the outside-in strategy and the inside-out strategy, which allows us to show there no ZCLAs exist for a target number of rounds.

Noteworthy, based on the works above, we propose a method to get all ZCLAs without the limitation of input and output mask.

Compared with previous methods, our method has the following advantages, which shows our method is a stronger cryptanalytic and designed method.

- **Able to search the ZCLAs by considering all the details of S-box for block ciphers when the sum of input and output bits of S-box is more than**

[3] http://stp.github.io/

16**-bit.** As we introduced, when the sum of input and output bits of S-box is more than 16-bit, previous methods cannot search the ZCLAs by considering all the details of S-box. The main obstacle is it is hard to model the propagation of masks through such S-box. In our method, for an S-box whose input is n-bit, the valid propagation of masks through it can be modeled by $2^n + 1$ expressions, which allow us to search ZCLAs by considering all the details of S-box for block ciphers when the sum of input and output bits of S-box is more than 16-bit.

- **Able to get the LNR which no 1-bit input and output active ZCLAs exist for block ciphers when the sum of input and output bits of S-box is more than** 16**-bit.** Since previous methods cannot model the propagation of masks through the S-box whose the sum of input and output bits is more than 16-bit, they cannot get the LNR which no 1-bit input and output active ZCLAs exist for the block cipher with such S-box. With our new modeling method, we can get the LNR which no 1-bit input and output active ZCLAs exist for block ciphers when the sum of input and output bits of S-box is more than 16-bit.

- **Able to get the LNR which no ZCLAs exist for SPN structure block ciphers without the limitation of input mask and output mask** For SPN structure block ciphers, previous automatic search methods can only show there no 1-bit/nibble/word input and output active ZCLAs exist, and there no theoretical results to show that there no ZCLAs exist under a given number of round for SPN structure block ciphers. In our method, we propose the outside-in strategy and the inside-out strategy, which allows us to show there no ZCLAs exist for a target number of rounds without the limitation of input mask and output mask, which is a strong method of block cipher in design aspects. Besides, as a by-product, our method also can be used to show there no t-bit/nibble/word input and output active ZCLAs exist with small t.

- **Able to get all ZCLAs for SPN structure block ciphers without the limitation of input mask and output mask** Due to the large search space, detecting all ZCLAs directly is impossible. Inspiring by our method for getting the LNR which no ZCLAs exist, we propose a method to get all ZCLAs for SPN structure block ciphers without the limitation of input mask and output mask, which is a strong method of block cipher in cryptanalysis aspects.

To show the ability of our tool, we apply it to the block ciphers MISTY1, MISTY2 [Mat97], AES [DR02], Midori64 [BBI+15] and ARIA [KKP+03]. These results are shown as follows.

- In the aspect of searching new ZCLAs, we apply our tool to MISTY1 and MISTY2, which are two block ciphers that adopt 7-bit and 9-bit S-boxes. Since the structures of two successive rounds of MISTY1 and four successive rounds of MISTY2 are different, we applied our tool to those block ciphers where the input mask is placed at difference round. Finally, we get 660, 1016, 986, 495, 710, 486 4-round 1-bit input and output active ZCLAs for MISTY1 where the input mask is placed in $2t+1$-th and $2t$-th round, and MISTY2 where the input mask is placed in $4s+1$, $4s+2$, $4s+3$-th and $4s$-th round respectively (s and t are two integrals). For comparison, we apply previous method to search the ZCLAs, that is, we model the propagation of mask through 9-bit S-box by "arbitrary mode", and model the propagation of mask through other operations accurately. As a result, we get 460, 860, 452, 308, 404, 308 4-round 1-bit input and output active ZCLAs for MISTY1 and MISTY2 where the input mask is placed at difference rounds. Those results show that, our method can detect more ZCLAs than previous methods indeed.

- In the aspect getting LNR which no 1-bit input and output active ZCLAs exist for block ciphers when the sum of input and output bits of S-box is more than 16-bit. Our method directly show that, even consider all the details of S-boxes, there still no 5-round 1-bit input and output active ZCLAs for MISTY1 and MISTY2 where the input mask is placed at difference rounds. Those results reflect the ability of MISTY1 and MISTY2 that against the zero-correlation linear attack.
- In the aspect of getting the LNR which no ZCLAs exist for SPN structure block ciphers, we apply our method for AES and Midori64 without considering the last MC layer, and ARIA without considering the last linear layer. Finally, the LNR we got for AES, Midori64, and ARIA is 5, 7, and 5 respectively. That is, there no 5, 7, and 5-round ZCLAs exist for AES and Midori64 without considering the last MC layer, and ARIA without considering the last linear layer respectively. Those results show that our method is indeed a strong method for block cipher in design aspects.
- In the aspect of getting all ZCLAs for SPN structure block ciphers, we apply our method for ARIA, and get all 166460 truncated ZCLAs for 4-round ARIA without the last linear layer. Those results show that our method is indeed a strong method for block cipher in cryptanalysis aspects.

Outline. We introduce the notations and related work in Section 2. Our method for searching ZCLAs is discussed in Section 3. The methods to get the LNR which no ZCLAs exist is proposed in Section 4. In Section 5, we propose our method for get all ZCLAs for SPN structure block ciphers. In Section 6, we conclude this paper.

2 Preliminaries

2.1 Zero-correlation Linear Approximation

Let us briefly recall the notations and concepts of zero-correlation linear approximation [Mat93, Nyb94, DGV94, CCH10]. The definition of zero-correlation linear approximation is close to the concept of correlation. We recall the definitions of inner product and correlation function first.

Definition 1 (Inner Product)**.** For $\lambda \in \mathbb{F}_2^n$, denote $(\lambda_{n-1}, \cdots, \lambda_1, \lambda_0)$ as the bit representation of λ. For any two elements $\alpha, x \in \mathbb{F}_2^n$, the inner product of them is defined as

$$\alpha \cdot x = \oplus_{i=0}^{n-1} \alpha_i x_i.$$

Definition 2 (Correlation)**.** For a function $f : \mathbb{F}_2^n \to \mathbb{F}_2^m$, the correlation of the linear approximation for an input mask $\alpha \in \mathbb{F}_2^n$ and an output mask $\beta \in \mathbb{F}_2^m$ is defined as

$$C_f(\alpha, \beta) = 2^{-n} \sum_{x \in \mathbb{F}_2^n} (-1)^{\alpha \cdot x \oplus \beta \cdot f(x)}.$$

Now, let us recall the correlation of linear trail for iterated function. Let $f : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be a function that is the iterated composition of round functions $f_i : \mathbb{F}_2^n \to \mathbb{F}_2^n$:

$$f := f_r \circ \cdots \circ f_2 \circ f_1.$$

Definition 3 (Linear Trail)**.** An r-round **linear trail** is an $(r+1)$-tuple $U = (u_0, u_1, \ldots, u_r)$, and the correlation of this linear trail is

$$C_f(U) = \prod_{i=1}^{r} C_{f_i}(u_{i-1}, u_i).$$

Based on the definition of linear trail, for an iterated function $f = f_r \circ \cdots \circ f_2 \circ f_1$, the correlation of the linear approximation for an input mask $\alpha \in \mathbb{F}_2^n$ and an output mask $\beta \in \mathbb{F}_2^m$ is calculated as

$$C_f(\alpha,\beta) = \sum_{u_0=\alpha, u_r=\beta, U=(u_0,\ldots,u_r)} C_f(U).$$

Then, the concept of zero-correlation linear approximation can be restated as follows.

Definition 4 (Zero-correlation Linear Approximation)**.** For an iterated function $f = f_r \circ \cdots \circ f_2 \circ f_1$, a pair of input and output masks (α, β) is an r-round **zero-correlation linear approximation** if and only if $C_f(\alpha,\beta) = 0$.

After the concept of zero-correlation linear attack is proposed, various types of ZCLAs are proposed. Actually, according to the values of input and output mask, the ZCLAs are divided into various types. To introduce those types of ZCLA, we summarize three types of masks as follows.

-k Bits Active Mask A n-bit mask α is called $k (k \le n)$ *bits active mask* if and only if the number of 1s in α is equal k.

-k Nibbles Active Mask For a n-bit mask $\alpha = (\alpha_{n-1}, \ldots, \alpha_1, \alpha_0)$, a $\frac{n}{4}$-bit vector $v = (v_{\frac{n}{4}-1}, \ldots, v_1, v_0)$ is called *the nibble active pattern* of it if and only if

$$\alpha_{4j+3}|\alpha_{4j+2}|\alpha_{4j+1}|\alpha_{4j} = \begin{cases} 1 & \text{if } v_j = 1, \\ 0 & \text{otherwise}, \end{cases}$$

where $0 \le j \le \frac{n}{4} - 1$. Such relation is donated as $\mathcal{N}(\alpha) = v$. Moreover, α is called k *nibbles active mask* if and only if the number of $1s$ in v is equal k.

-k Words Active Mask For a n-bit mask $\alpha = (\alpha_{n-1}, \ldots, \alpha_1, \alpha_0)$, a $\frac{n}{n}$-bit vector $v = (v_{\frac{n}{n}-1}, \ldots, v_1, v_0)$ is called *the nibble active pattern* of it if and only if

$$\alpha_{8j+3}|\alpha_{8j+2}|\cdots|\alpha_{8j+1}|\alpha_{8j} = \begin{cases} 1 & \text{if } v_j = 1, \\ 0 & \text{otherwise}, \end{cases}$$

where $0 \le j \le \frac{n}{8} - 1$. Such relation is donated as $\mathcal{W}(\alpha) = v$. Moreover, α is called k *nibbles active mask* if and only if the number of $1s$ in v is equal k.

With those definitions of masks, the more specific definitions of ZCLA are summarized as follows.

Definition 5 ((k_i, k_0)-bit (nibble, word) Active Input and Output ZCLA)**.** A pair of input and output mask (α, β) is called (k_i, k_0)*-bit (nibble, word) active input and output ZCLA (AIAO-ZCLA)* if and only if α is k_i-bit (nibble, word) active mask, β is k_o-bit (nibble, word) active mask, and α cannot propagate to β.

Definition 6 ((u, v)-nibble (word) Pattern ZCLA)**.** A pair of input and output mask (α, β) is called (u, v)*-nibble (word) pattern ZCLA* if and only if $\mathcal{N}(\alpha) = u(\mathcal{W}(\alpha) = u)$, $\mathcal{N}(\beta) = v(\mathcal{W}(\beta) = v)$, and α cannot propagate to β.

Definition 7 ((u, v)-nibble (word) Pattern Truncated ZCLA)**.** For a block cipher, we call it has (u, v)*-nibble (word) pattern truncated ZCLA* if and only if for $\forall \mathcal{N}(\alpha) = u(\forall \mathcal{W}(\alpha) = u)$ and $\forall \mathcal{N}(\beta) = v(\forall \mathcal{W}(\beta) = v)$, α cannot propagate to β.

Definition 8 ((k_i, k_0)-nibble (word) Active Input and Output Truncated ZCLA)**.** For a block cipher, we call it has (k_i, k_0)*-nibble (word) active input and output truncated ZCLA (AIAO-T-ZCLA)* if and only if for all k_i nibbles (words) active mask α and k_o nibbles (words) active mask β, α cannot propagate to β.

2.2 ZCLAs Search with SAT Method

The SAT problem is a classic scientific computation problem aiming to determine whether there exists a solution for a set of boolean formulas, where each boolean formula is called a clause and the variable in each clause is binary. Although the SAT problem is NP-completed, modern solvers can solve such problem with tens of thousands of variables and clauses.

The problem of searching ZCLAs for a given function F can be solved by SAT method. For example, for a given search space and any input and output mask (α, β) in this space, the SAT method can be used to determine whether (α, β) is a ZCLA for F or not as follows:

1. Modeling the propagation of mask through F, limiting the input mask be α and output mask be β by boolean formulas;

2. Using the SAT solver to determine whether there exists a solution for the boolean formulas. If not, then (α, β) is a ZCLA for a given function F; Otherwise, it isn't.

Thus, by testing all input and output mask in a given search space, one can get all ZCLAs in this space or ensure there no ZCLAs exist in such search space. Similar approach can be applied to search truncated ZCLAs. In the next, we recall the method for modeling the propagation of mask and the limitation of input and output mask by boolean formulas, which are the most necessary parts for searching ZCLAs by SAT method.

2.2.1 Modeling the Propagation of Mask

In this part, we recall the method for modeling the propagation of mask through the operations **XOR**, **COPY**, **Linear Map**, and **S-box** by boolean formulas. For the sake of simplicity, for $\alpha \in \mathbb{F}_2^n$, we donate $(\alpha_0, \alpha_1, \ldots, \alpha_{n-1})$ as the binary representation of α, where $\alpha_i \in \mathbb{F}_2 (0 \leq i \leq n-1)$.

XOR. Let f be a XOR function, assume $\alpha_0, \alpha_1 \in \mathbb{F}_2^n$ are the input masks, and $\beta \in \mathbb{F}_2^n$ is the output mask. Then, the correlation $C_f((\alpha_0, \alpha_1), \beta) \neq 0$ if and only if $\alpha_0 = \alpha_1 = \beta$. Thus, let $x_0 = (x_{0,0}, x_{0,1}, \cdots, x_{0,n-1}), x_1 = (x_{1,0}, x_{1,1}, \cdots, x_{1,n-1})$ be the variables for the input masks, and $y = (y_0, y_1, \cdots, y_{n-1})$ be the variable for output mask, the following boolean formulas can model the propagation of mask through XOR:

$$\begin{cases} y_i = x_{0,i}, \\ y_i = x_{1,i}, \end{cases} (0 \leq i \leq n-1). \tag{1}$$

In particular, the operation XORed with the key or constant only changes the sign of the correlation.

COPY. Let f be a COPY function, assume $\alpha \in \mathbb{F}_2^n$ is the input mask, and the $\beta_0, \beta_1 \in \mathbb{F}_2^n$ are the output masks. Then, the correlation $C_f(\alpha, (\beta_0, \beta_1)) \neq 0$ if and only if $\alpha = \beta_0 \oplus \beta_1$ [Bih94]. Thus, let $x = (x_0, x_1, \cdots, x_{n-1})$ be the variable for the input mask, and $y_0 = (y_{0,0}, y_{0,1}, \cdots, y_{0,n-1}), y_1 = (y_{1,0}, y_{1,1}, \cdots, y_{1,n-1})$ be the variables for output mask, the following boolean formulas can model the propagation of mask through COPY:

$$x_i = y_{0,i} \oplus y_{1,i}, (0 \leq i \leq n-1). \tag{2}$$

Linear Map. Let $M = (m_{i,j})_{0 \leq i \leq n-1, 0 \leq j \leq n-1}$ be the binary matrix representation of Linear Map, assume $\alpha \in \mathbb{F}_2^n$ is the input mask and $\beta \in \mathbb{F}_2^n$ is the output mask. Then, the correlation $C_f(\alpha, \beta) \neq 0$ if and only if $\beta = M^t \alpha$, where M^t is the transpose of M [DGV94]. Thus, let $x = (x_0, x_1, \cdots, x_{n-1})$ be the variable for the input mask, and $y = (y_0, y_1, \cdots, y_{n-1})$ be the variable for output mask, the following boolean formulas can model the propagation of mask through Linear Map:

$$y_t = \oplus_{p \in \{q | m_{t,q}=1\}} x_p, (0 \leq t \leq n-1). \tag{3}$$

S-box. Let S be a S-box which substitute n-bit value to m-bit value, assume $\alpha \in \mathbb{F}_2^n$ is the input mask and the $\beta \in \mathbb{F}_2^m$ is the output mask. Then, $\alpha \xrightarrow{S} \beta$ with correlation $C_S(\alpha, \beta)$ [DGV94]. That is, the input mask α can propagate to the output mask β if and only if $C_S(\alpha, \beta) \neq 0$. The Abdelkhalek et al's method can be used to model the propagation of mask through S-box.

The origin of Abdelkhalek et al's method is applied to derived the linear inequalities for modeling the propagation of difference. With the same idea, we can use this method to get the boolean formulas for modeling the propagation of mask. First, we recall the definition of product-of-sum representation of boolean function.

Definition 9 (Product-of-Sum Representation of Boolean Function)**.** Let $\overline{f}(x, y)$ be a boolean function of $(n + m)$-bit inputs, where the input is $x = (x_{n-1}, \ldots, x_1, x_0)$ and $y = (y_{m-1}, \ldots, y_1, y_0)$ and the output of $\overline{f}(x, y)$ is 0 or 1. The product-of-sum representation of the boolean function $\overline{f}(x, y)$ is defined as

$$\overline{f}(x, y) = \wedge_{\boldsymbol{c}=(c_0,\ldots,c_{n-1},c_{n+0},\ldots,c_{n+m-1})\in T}(((\vee_{i=0}^{i=n-1}(x_i \oplus c_i)) \vee ((\vee_{i=0}^{i=m-1}(y_i \oplus c_{n+i})))$$

where T is a set that determined by $\overline{f}(x, y)$, and $\wedge$ and $\vee$ denote logical AND and OR operations, respectively.

Thus, for a given boolean function $\overline{f}(x, y)$, once we get the product-of-sum representation of the boolean function of it, the values that satisfies the following boolean formulas are just the values such that $\overline{f}(x, y) = 1$.

$$(\vee_{i=0}^{i=n-1}(x_i \oplus c_i)) \vee (\vee_{i=0}^{i=m-1}(y_i \oplus c_{n+i})) = 1, \boldsymbol{c} \in T. \tag{4}$$

In Abdelkhalek et al's method, they proposed the software Logic Friday[4] to get the product-of-sum representation of the boolean function. Moreover, the Logic Friday can get the product-of-sum representation of a given boolean function with minimum number of terms or as small as possible number of terms[5].

Let $f : \mathbb{F}_2^n \times \mathbb{F}_2^m \longrightarrow \mathbb{F}_2$, and

$$f(x, y) = \begin{cases} 1, & \text{if } C_s(x, y) \neq 0, \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

We use the Logic Friday to get the product-of-sum representation of f with minimum number of terms or as small as possible number of terms. Then, we can get the boolean formulas such that the values that satisfies those boolean formulas are just the values such that $f(x, y) = 1$, that is, the valid propagation of mask through S-box.

2.2.2 Limiting the Input and Output Mask

In this part, we recall the limitation of input and output mask for basic type of ZCLAs. Then, we discuss the spaces for searching ZCLAs. Actually, according to the limitation of input and output mask, the ZCLAs can be divided into two classes:

Fixed values ZCLA: For a given n-bit function F and a pair of input and output mask $(\alpha, \beta)(\alpha, \beta \in \mathbb{F}_2^n)$, assume $x = (x_0, x_1, \cdots, x_{n-1})$ and $y = (y_0, y_1, \cdots, y_{n-1})$ be the variable for the input and output mask, respectively. To determine whether (α, β) is an ZCLA of F or not, we limit the input and output mask as

$$\begin{cases} x_i = \alpha_i, \\ y_i = \beta_i, \end{cases} (0 \leq i \leq n - 1).$$

[4] http://sontrak.com/.

[5] The relation of the number of boolean formulas and the solve efficiency of SAT solver is unclear. However, as a consensus, less number of boolean formulas often leads to high solve efficiency of SAT solver. Thus, the number of boolean formulas to model the propagation of mask through S-box should as small as possible.

(u, v)-nibble (word) pattern truncated ZCLA: For a given n-bit function F and a pair of nibble (word) active pattern (u, v), assume $x = (x_0, x_1, \cdots, x_{n-1})$ and $y = (y_0, y_1, \cdots, y_{n-1})$ be the variable for the input and output mask, respectively. To determine whether there exist an (u, v)-nibble (word) pattern truncated ZCLA of F or not, we limit the input and output mask as

$$\begin{cases} x_{mp+m-1}|x_{mp+m-2}|\cdots|x_{mp} = u_p(0 \leq p \leq \frac{n}{m}), \\ y_{mq+m-1}|y_{mq+m-2}|\cdots|x_{mq} = v_q(0 \leq q \leq \frac{n}{m}). \end{cases}$$

With those two basic types of ZCLAs, we recall the search space to get the ZCLAs which are much attracted by the cryptanalysts.

(k_i, k_o)-bit AIAO-ZCLAs. A pair of input and output mask (α, β) is called (k_i, k_o)-bit AIAO-ZCLA of a function F, if $(\alpha, \beta) \in \mathcal{S} = \{(\alpha, \beta)|\alpha$ is the k_i-bitactive mask, and β is the k_o-bit active mask$\}$ and (α, β) is a fixed value ZCLA. The size of overall search space is $\binom{n}{k_i} \times \binom{n}{k_o}$, it is infeasible to get all ZCLAs when k_i and k_0 are large. Thus, cryptanalysts usually focus on small value of k_i and k_0, such as $(1, 1)$-bit AIAO-ZCLAs.

(u, v)-nibble (word) pattern ZCLAs. A pair of input and output mask (α, β) is called (u, v)-nibble (word) pattern ZCLA of a function F, if $(\alpha, \beta) \in \mathcal{S}_{u,v} = \{(\alpha, \beta)|\mathcal{N}(\alpha) = u, \mathcal{N}(\beta) = v\}(\mathcal{S}_{u,v} = \{(\alpha, \beta)|\mathcal{W}(\alpha) = u, \mathcal{W}(\beta) = v\})$ and (α, β) is a fixed value ZCLA. The size of overall search space is $(2^m - 1)^{wt(u)+wt(v)}$, where $m = 4(m = 8)$.

(k_i, k_o)-nibble (word) AIAO-ZCLAs. A pair of input and output mask (α, β) is called (u, v)-nibble (word) pattern ZCLA of a function F, if $(\alpha, \beta) \in \mathcal{S} = \bigcup_{\{(u,v)|wt(u)=k_i, wt(v)=k_o\}} \mathcal{S}_{u,v}$ and (α, β) is a fixed value ZCLA. The size of overall search space is $(\binom{\frac{n}{m}}{k_i}(2^m - 1)^{k_i}) \times (\binom{\frac{n}{m}}{k_o}(2^m - 1)^{k_o})$, it is infeasible to get all ZCLAs when k_i and k_0 are large. Thus, cryptanalysts usually focus on small value of k_i and k_0, such as $(1, 1)$-nibble (word) AIAO-ZCLAs.

(k_i, k_o)-nibble (word) AIAO-T-ZCLAs. To search all such ZCLAs of a function F, let $\mathcal{S} = \{(u, v)|wt(u) = k_i, wt(v) = k_o\}$, and for $(u, v) \in \mathcal{S}$, we determine whether there exist an (u, v)-nibble (word) pattern truncated ZCLA of F or not. The size of overall search space is $\binom{\frac{n}{m}}{k_i} \times \binom{\frac{n}{m}}{k_o}$.

Since ZCLAs play an important role in the zero-correlation linear attack, the lower number of rounds which no ZCLAs exist indicates the security of a given block cipher against the zero-correlation linear attack, which is very helpful in the design of block ciphers.

3 Evaluating the Security by Meet-in-the-Middle Approach

For a block cipher E, let S be the whole search space, if for all input and output mask in S are not ZCLA of E, we can determine there no ZCLAs exist for E in S. However, as the discussion above, the search space S is too large, which is infeasible by present searching method. To solve this problem, we propose the meet-in-the-middle (MITM) approach.

3.1 The Overview of Meet-in-the-Middle Approach

In this part, we give our MITM approach to show there no ZCLAs exist for a block cipher E. First, we give the following theorem.

Theorem 1. *Let E is a n bits SPN structure block cipher whose Sbox's size is m bits, α and β be the input mask and output mask of E. Let $E = F_1 \circ F \circ F_0$, $\Phi_I(\beta) = \{\gamma_I \in \mathbb{F}_2^{n*} | \gamma_I \xrightarrow{F_0} \beta\}$ and $\Phi_O(\alpha) = \{\gamma_O \in \mathbb{F}_2^{n*} | \alpha \xrightarrow{F_1} \gamma_O\}$, if the following conditions are hold:*

1. *There exits h sets H_i and $\widehat{H}_i (0 \leq i \leq h-1)$, such that*
   - *$H_0 \bigcup \cdots \bigcup H_{h-1} = \mathbb{F}_2^{n*}$;*
   - *For $\forall \eta \in \widehat{H}_i (0 \leq i \leq h-1)$, it holds $H_i \subseteq \Phi_I(\eta)$.*

   Meanwhile, there exists t sets T_j and $\widehat{T}_j (0 \leq j \leq t-1)$ such that
   - *$T_0 \bigcup \cdots \bigcup T_{t-1} = \mathbb{F}_2^{n*}$;*
   - *For $\forall \theta \in \widehat{T}_j (0 \leq j \leq t-1)$, it hold that $T_j \subseteq \Phi_O(\theta)$.*
2. *For each $\widehat{H}_i (0 \leq i \leq h-1)$ and $\widehat{T}_j (0 \leq j \leq t-1)$, there exits at least an $\eta \in \widehat{H}_i$ and an $\theta \in \widehat{T}_j$, such that $\eta \xrightarrow{F} \theta$.*

Then, there no ZCLA exists for E.

Proof. For any input mask γ and output mask δ, since $\gamma \in \mathbb{F}_2^{n*}$ and $\delta \in \mathbb{F}_2^{n*}$, according to condition 1, there exist $i, j (0 \leq i \leq h-1, 0 \leq j \leq t-1)$, such that $\gamma \in H_i$ and $\delta \in T_j$. According to condition 3, there exist $\eta \in \widehat{H}_i$ and $\theta \in \widehat{T}_j$, such that $\eta \xrightarrow{F} \theta$. Moreover, according to condition 1 again, we have $\gamma \xrightarrow{F_0} \eta$ and $\theta \xrightarrow{F_1} \delta$. Thus, the input mask γ can propagate to the output mask δ through E. The whole process is shown in Figure 1. □

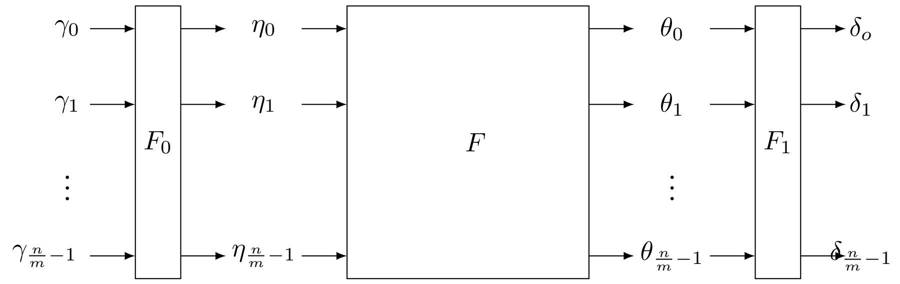

Figure 1: Connection of Any Input and Output Mask

For a block cipher E, we can adopt the following approach to show there no ZCLAs exist.

Function divide. Choosing appropriate F_1, F, and F_0, such that $E = F_1 \circ F \circ F_0$.

State divide. Choosing appropriate h and t, H_i and $\widehat{H}_i (0 \leq i \leq h-1)$, and T_j and $\widehat{T}_j (0 \leq j \leq t-1)$, such that condition 2 holds.

Masks connection. For each $\widehat{H}_i (0 \leq i \leq h-1)$ and $\widehat{T}_j (0 \leq j \leq t-1)$, we show that there exits at least an $\eta \in \widehat{H}_i$ and an $\theta \in \widehat{T}_j$, such that $\eta \xrightarrow{F} \theta$.

The step of masks connection can be implemented by SAT method. For the other steps, form the view of function, the functions F_1, F, and F_0, the values h and t, and the set H_i and $\widehat{H}_i (0 \leq i \leq h-1)$, and T_j and $\widehat{T}_j (0 \leq j \leq t-1)$ can be treated as parameters, which decide the efficiency of our proof for showing there no ZCLAs exist. For example, in the case of F_0 and F_1 both are identity function, the MITM approach to show there no ZCLAs exist for E is indeed to show all input and output mask are not ZCLA, which is infeasible as we discussed above. Thus, we should deal those parameters carefully.

3.2 The Selection of Parameters of Function Divide

As the first step of MITM approach, the selection of F_0 and F_1 has a great influence on the efficiency of the proof. When F_0 or F_1 covers more rounds, due to the clusters of masks, the values of h or t may be small, thus we need less tests in masks connection. However, this also causes that F covers less rounds, which makes finding an $\eta \in \widehat{H}_i$ and an $\theta \in \widehat{T}_j$, such that $\eta \xrightarrow{F} \theta$ difficult. Thus, we need to weigh the pros and cons when we divides the block cipher E.

As we focus on the SPN structure block ciphers, we choose $F_0 = F_1 = \widehat{S} \circ \widehat{P} \circ \widehat{S}$, where $\widehat{P}$ donates the linear layer of E and $\widehat{S}$ donates the Sboxes layer.

3.3 The Selection of Parameters of State Divide

Since we choose $F_0 = F_1$, the method for choose the values of h and the set H_i and $\widehat{H}_i (0 \leq i \leq h-1)$ are same as choosing the values of t and and T_j and $\widehat{T}_j (0 \leq j \leq t-1)$. Thus, we only focus on choosing the values of h and the set H_i and $\widehat{H}_i (0 \leq i \leq h-1)$ of $F_0 = \widehat{S} \circ \widehat{P} \circ \widehat{S}$. The method for choosing such parameters are closely related to the property of block ciphers, we mainly study two types of block ciphers: nibble-level linear layer based block ciphers and AES-like block ciphers. For the sake of presentation, for an m-bit Sbox, let β be the output mask of S, donate $\phi_i(\beta) = \{\gamma_i \in \mathbb{F}_2^m | \gamma_i \xrightarrow{S} \beta\}$, $\widetilde{O}$ is a set such that $\bigcup\limits_{\beta \in \widetilde{O}} \phi_i(\beta) = \mathbb{F}_2^{m*}$, $\widetilde{O}_A = \{\widetilde{O} | \bigcup\limits_{\beta \in \widetilde{O}} \phi_i(\beta) = \mathbb{F}_2^{m*}\}$, $M_O(S) = min_{\widetilde{O} \in \widetilde{O}_A} \#\widetilde{O}$, $A_O(S) = \{\widetilde{O} | \widetilde{O} \in \widetilde{O}_A, \#\widetilde{O} = M_O(S)\}$.

3.3.1 Nibble-level Linear Layer Based Block Ciphers

The nibble-level linear layer based block ciphers is a class of common block ciphers, such as ARIA, Robin, Modori and so on. For a n-bit such block cipher, the state s can be arranged as an $1 \times q$ matrix as follows:

$$s = \begin{pmatrix} s_0 & s_1 & \cdots & s_{q-1} \end{pmatrix},$$

where $s_i \in \mathbb{F}_2^{\mu} (0 \leq i \leq q-1)$, this representation also hold for the mask of a state.

Meanwhile, one round of the encryption process (omit the key add layer) can be expressed as follows.

SubBytes(SB): Applying the μ-bit S-box S to each nibble or byte in parallel of the cipher internal state.

MixColumns(MC): The whole states are multiplied by the $q \times q$ word (nibble) level matrix $W = (w_{i,j})_{0 \leq i \leq q-1, 0 \leq j \leq q-1}$.

Let $\odot$ be an operate such that $(y_0, \ldots, y_{q-1})^T = M \odot (x_0, \ldots, x_{q-1})^T$ if and only if $y_i = m_{0,i}x_0 || \cdots || m_{q-1,i}x_{q-1}$ and $P_{\odot}(M) = \{(y_0, \ldots, y_{q-1})^T = M \odot (x_0, \ldots, x_{q-1})^T | (x_0, \ldots, x_{q-1}) \in \mathbb{F}_2^{q*}\}$, where $M = (m_{i,j})_{0 \leq i \leq q-1, 0 \leq j \leq q-1} = (W^{-1})^T$ and T is the symbol of transposition. Assume $P_{\odot}(M) = \{\boldsymbol{y_i} | 0 \leq i \leq t-1\}$, let $H_i = \{\alpha \in \mathbb{F}_2^{n*} |$ the active pattern of α is $\boldsymbol{x}$, where $M^t \odot \boldsymbol{x} = \boldsymbol{y}\}$. Then, it is easily to verify $H_0 \bigcup \cdots \bigcup H_t = \mathbb{F}_2^{n*}$. Moreover, if the block cipher has the following property, we can get $\widehat{H}_i$ for each $H_i (0 \leq i \leq t)$ easily.

Property 1. For a given block cipher E, for two consecutive rounds SB layer $(S_0^0, \ldots, S_{q-1}^0)$ and $(S_0^1, \ldots, S_{q-1}^1)$, if there at least exist two disjoint sets G_0 and G_1, such that

- $G_0, G_1 \subseteq A_0(S_0^0) \bigcap \cdots \bigcap A_0(S_{q-1}^0)$.

- $Z_i = \{z \in \mathbb{F}_2^{\mu} |$ for $\forall x \in G_0 \bigcup G_1 \bigcup \{z_0 \oplus z_1 | z_0 \in G_0, z_1 \in G_1\}, x \overset{S_i^1}{\to} z\} \neq \emptyset (0 \leq i \leq q-1)$.

Then, we call E has $P_{H,arl}^t$ property.

Theorem 2. *Let E a block cipher that fulfills property $P_{H,arl}^t$, for any $0 \leq i \leq t$, constructing $R = R_0 \times \cdots \times R_{q-1}$, where $R_j = Z_j$ when $(\boldsymbol{y_i})_j = 1$ and $R_j = \{0\}$ otherwise $(0 \leq j \leq q-1)$. Then, $\forall \beta \in R$, it holds that $H_i \subseteq \Phi_I(\beta)$.*

Proof. For $\forall \alpha \in H_i$, there exists a $\boldsymbol{x}$ such that the active pattern of α is $\boldsymbol{x}$ and $M \odot \boldsymbol{x} = \boldsymbol{y_i}$. Let $V = \{v | x_v = 1\} = \{v_p | 0 \leq p \leq l\}$ and $R' = R'_0 \times \cdots \times R'_{q-1}(R'_{p'} = G_0$ if $p' = v_0, R'_q = G_1$ if $p' \neq v_0$ and $p' \in V$, and $R_{p'} = \{0\}$ otherwise$, 0 \leq p' \leq q-1)$. Since $G_0, G_1 \subseteq A_0(S_0^0) \bigcap \cdots \bigcap A_0(S_{q-1}^0)$, there exists $\gamma \in R'$, such that $\alpha \overset{\widehat{S}}{\to} \gamma$. Meanwhile, since M is the word (nibble) level matrix, for $\forall \gamma' \in R'$ and $\mu = M\gamma'$, we have $\mu_k = 0((\boldsymbol{y_i})_k = 0, 0 \leq k \leq q-1)$ and $\mu_k((\boldsymbol{y_i})_k = 1, 0 \leq k \leq q-1)$ is the xor of some values of $\{\gamma'_{k'} | \boldsymbol{x}_{k'} = 1\}$. Thus, for $0 \leq k \leq q-1((\boldsymbol{y_i})_k = 1)$, $\mu_k \in G_0 \bigcup G_1 \bigcup \{z_0 \oplus z_1 | z_0 \in G_0, z_1 \in G_1\}$. According to the property 1, for $\forall \varepsilon_k \in Z_k$, it holds $\mu_k \overset{S_k^1}{\to} \varepsilon_k$. Therefore, for $\forall \beta \in R_0 \times \cdots \times R_{q-1}$, it holds that $H_j \subseteq \Phi_i(\beta)$. □

Form Theorem 2, for each $H_i (0 \leq i \leq t)$, we can choose $\widehat{H}_i = R_0 \times \cdots \times R_{q-1}$, where $R_j = Z_j$ when $(\boldsymbol{y_i})_j = 1$ and $R_j = \{0\}$ otherwise $(0 \leq j \leq q-1)$.

3.3.2 AES-like block ciphers

For the sake of representation, we depict the AES-like block ciphers as follows. Actually, for a n-bit AES-like block cipher, the state s can be arranged as an $p \times q$ matrix as follows:

$$s = \begin{pmatrix} s_{0,0} & s_{0,1} & \cdots & s_{0,q-1} \\ s_{1,0} & s_{1,1} & & \\ \vdots & & \ddots & \\ s_{r-1,0} & & & s_{r-1,c-1} \end{pmatrix},$$

where $s_{i,j} \in \mathbb{F}_2^{\mu} (0 \leq i \leq p-1, 0 \leq j \leq q-1)$, this representation also hold for the mask of a state.

Meanwhile, one round of the encryption process (omit the key add layer) can be summarized as follows, where

SubBytes(SB): Applying the μ-bit S-box S to each nibble or byte in parallel of the cipher internal state.

ShuffleCell(SC): The permutation of the nibbles or bytes. For the sake of representation, assume the nibble or byte in i_0-th row and j_0-th column is transform to the i_1-th row and j_1-th column, we define the function λ and τ to describe the transform of positions, this is, $\lambda(i_0, j_0) = (i_1, j_1)$ and $\tau(i_1, j_1) = (i_0, j_0)$.

MixColumns(MC): Each column of the internal state is multiplied by the mix-column matrix M.

Before we start, we introduce some notations firstly. For $J \subseteq \{0, 1, \ldots, c-1\}$ and a mask $\alpha = (\alpha_{i,j})_{0 \leq i \leq r-1, 0 \leq j \leq c-1}$, let $R_J = R_0 \times \cdots \times R_{c-1}(R_j \subseteq \mathbb{F}_2^{\mu *}$ for $j \in J, R_j = \{0\}$ for $j \notin J, 0 \leq j \leq c-1)$, $\alpha_j = (\alpha_{0,j}, \ldots, \alpha_{r-1,j})$, and $D_{R_J} = \{\alpha | \alpha_j \in R_j, 0 \leq j \leq c-1\}$. Analogously, we define $\alpha_j^{SC^{-1}} = (\alpha_{x_0^j, y_0^j}, \ldots, \alpha_{x_{r-1}^j, y_{r-1}^j})$, where $\tau(l, j) = (x_l^j, y_l^j)$, and $D_{R_J}^{SC^{-1}} = \{\alpha | \alpha_j^{SC^{-1}} \in R_j, 0 \leq j \leq c-1\}$. In particular, let $R_J = R_0 \times \cdots \times R_{c-1}(R_j =$

$\mathbb{F}_2^{\mu *}$ for $j \in J, R_j = \{0\}$ for $j \notin J, 0 \leq j \leq c-1$). Then, for $\forall \beta \in D_{R_J}^{SC^{-1}}$, we have $\{\gamma | \beta \xrightarrow{F_0} \gamma\} \subseteq D_{R_J}$. Let $\overline{S} = \underbrace{(S, S, \ldots, S)}_{r}$, $\overline{F}_0 = \overline{S} \circ W \circ \overline{S}$. For $\forall \delta \in (\mathbb{F}_2^{\mu})^r$, we define $\overline{\Phi}_I(\delta) = \{\epsilon | \epsilon \xrightarrow{\overline{F}_0} \delta\}$. Then, if the block cipher fulfills the following property, we can get H_i and $\widehat{H}_i (0 \leq i \leq t)$.

Property 2. For a AES-like block cipher E, if there exist $2l$ sets $L_i, \widehat{L}_i (0 \leq i \leq l-1)$, such that $L_0 \bigcup \cdots \bigcup L_{l-1} = (\mathbb{F}_2^{\mu})^{r*}$, and for $\forall \beta \in \widehat{L}_i$, $L_i \subseteq \overline{\Phi}_I(\beta)$. Then, we call E has $P_{H,ael}^{l}$ property.

For a block cipher, it always has property $P_{H,ael}^{r \times \mu}$. However, from an application point of view, the value of l should as small as possible. Here, we propose how to detect the value of l as small as possible in two cases.

-$u = 4$: In this case, we build a $(r \times \mu) \times (r \times \mu)$ table U, where for $\forall \alpha, \beta \in (\mathbb{F}_2^{\mu})^{r*}$, the α row and the β column equal 1 if and only if $\alpha \in \sigma(\beta)$, and 0 otherwise.

When $Branch(M) = r + 1$: Since for $\forall \alpha \in (\mathbb{F}_2^{\mu})^{r*}$, there may exists β such that $\boldsymbol{wt}(\beta) = r$ and $\alpha \xrightarrow{\overline{F}_0} \beta$, we choose $l = 1$, $L_0 = (\mathbb{F}_2^{\mu})^{r*}$, and detect where there exist some such values of β. That is, we search all $\beta \in (\mathbb{F}_2^{\mu *})^r$ and detect where there exists β such that, for $\forall \alpha \in L_0$, it holds that $U(\alpha, \beta) = 1$. We add the β which meets the criteria into $\widehat{L}_0$.

When $Branch(M) = r$: In this case, we choose $l = r+1$. Specifically speaking, we divide $(\mathbb{F}_2^{\mu})^{r*}$ into $r+1$ sets $L_i (0 \leq i \leq r)$, where $L_i = \{\alpha |$ the active pattern of α is $(b_0, \ldots, b_{r-1})$, where $b_i = 1$, and $b_j = 0 (j \neq i)\} (0 \leq i \leq r-1)$ and $L_r = \{\alpha | \boldsymbol{wt}(\alpha) \geq 2\}$. Then, for each i, we search β in the set $\{\gamma | \exists \alpha \in L_i, \alpha \xrightarrow{\overline{F}_0} \gamma\}$ and detect where there exists β such that, for $\forall \alpha \in L_i$, it holds that $U(\alpha, \beta) = 1$. We add the β which meets the criteria into $\widehat{L}_i$.

When $Branch(M) < r$: The handling method is similar to the above two cases.

If $\widehat{L}_i$ is empty, we add 1 to l and divide the L_i into 2 parts by observation and repeat the process above until $\widehat{L}_0, \ldots, \widehat{L}_{l-1}$ are not empty.

-$u = 8$: In this case, building a $(r \times \mu) \times (r \times \mu)$ table is infeasible. Thus, we make use of the set A_O. For the sake of clarity, for $\widetilde{O}_0, \ldots, \widetilde{O}_{r-1} \in A_O$ and a pattern $b = (b_0, \ldots, b_{r-1}) \in \mathbb{F}_2^r$, we donate $Q^b = Q_0 \times \cdots \times Q_{r-1} (Q_i = \widetilde{O}_i$ if $b_i = 1$, and $Q_i = \{0\}$ otherwise).

When $Branch(M) = r + 1$: We choose first $l = 1$, $L_0 = (\mathbb{F}_2^{\mu})^{r*}$, and pick $\widetilde{O}_0, \ldots, \widetilde{O}_{r-1} \in A_O$ randomly. Then, we detect the β, such that for all $\alpha \in \{(M^{-1})^T x | x \in Q^b, b \in \mathbb{F}_2^{r*}\}$, it hold that $\alpha \xrightarrow{\overline{S}} \beta$. We add the β which meets the criteria into $\widehat{L}_0$. Moreover, if computing power permits, we can repeat above process many times.

When $Branch(M) = r$: In this case, we choose $l = r + 1$, divide $(\mathbb{F}_2^{\mu})^{r*}$ into $r + 1$ sets $L_i (0 \leq i \leq r)$ as the situation $u = 4$, and pick $\widetilde{O}_0, \ldots, \widetilde{O}_{r-1} \in A_O$ randomly. Then, for each i, we detect the β, such that for all $\alpha \in \{M^t x | x \in Q^b,$ the set of active pattern of element in L_i contain $b\}$, We add the β which meets the criteria into $\widehat{L}_i$. Moreover, if computing power permits, we can repeat above process many times.

When $Branch(M) < r$: The handling method is similar to the above two cases.

If $\widehat{L}_i$ is empty, we add 1 to l and divide the L_i into 2 parts by observation and repeat the process above until $\widehat{L}_0, \ldots, \widehat{L}_{l-1}$ are not empty.

Under the discussion above, we give the following Theorem, which allows us to construct H_i and $\widehat{H}_i(0 \leq i \leq h-1)$.

Theorem 3. *Let E a block cipher that fulfills property $P^l_{H,ael}$, for $\forall J \subseteq \{0, 1, \ldots, c-1\}$, and $w \in W_J = \{(i_0, \ldots, i_{c-1}) | 0 \leq i_j \leq l-1 \text{ for } j \in J, \text{ and } i_j = 0 \text{ for } j \notin J\}$, let $R_{J,w} = R_0 \times \cdots \times R_{c-1}(R_j = L_{w_j} \text{ for } j \in J, R_j = \{0\} \text{ for } j \notin J, 0 \leq j \leq c-1)$ and $R'_{J,w} = R'_0 \times \cdots \times R'_{c-1}(R'_j = \widehat{L}_{w_j} \text{ for } j \in J, R_j = \{0\} \text{ for } j \notin J, 0 \leq j \leq c-1)$. Then, for $\forall \beta \in D^{SC^{-1}}_{R_{J,w}}$, it holds that $\{\gamma | \beta \xrightarrow{F_0} \gamma\} \subseteq D_{R'_{J,w}}$.*

From Theorem 3, if a block cipher E fulfills property $P^l_{H,ael}$, let $h = \sum_{i=1}^{c} \binom{c}{i} l^i = (l+1)^c - 1$, we can choose $H_i = D^{SC^{-1}}_{R_{J,w}}$ and $\widehat{H}_i = D_{R'_{J,w}}(0 \leq i \leq h-1, J \subseteq \{0, 1, \ldots, c-1\}, w \in W_J)$.

4 Application MITM Approach to Block Ciphers

4.1 Modori64

Midori64 is a 64-bit block cipher, it is a SPN structure block cipher with almost MDS matrix. The internal state is viewed as a 4×4 square array of bytes as follows, where $s_i \in \mathbb{F}_2^4 (0 \leq i \leq 16)$.

$$s = \begin{pmatrix} s_0 & s_4 & s_8 & s_{12} \\ s_1 & s_5 & s_9 & s_{13} \\ s_2 & s_6 & s_{10} & s_{14} \\ s_3 & s_7 & s_{11} & s_{15} \end{pmatrix}.$$

One encryption round of Midori64 is illustrated in Figure 2, it is composed of the following four operations:

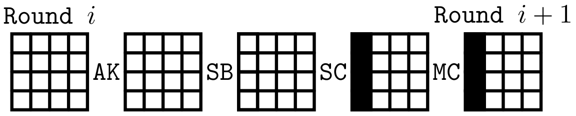

Figure 2: One Round of Midori64

AddRoundKey(AK): The 64-bit round key XORed with the 64-bit state.

SubBytes(SB): Applying the 4-bit involutive S-box S to each nibble in parallel of the cipher internal state.

ShiftRows(SC): The i-th nibble of the internal state is permuted to the perm(i)-th nibble, where perm $= [0, 10, 5, 15, 14, 4, 11, 1, 9, 3, 12, 6, 7, 13, 2, 8]$.

MixColumns(MC): Each column of the internal state is multiplied by the involutive almost-MDS matrix M.

Previous methods have not been evaluated the ability of Midori64 against the zero-correlation linear attack. But, according to the dual between the impossible differentials and ZCLAs, ST-method can be generalized to get the LNR which no ZCLAs exist. However, this method limits the input and output mask only 1-nibble active, which implicit there is still a big gap to get the LNR which no ZCLAs exist. In the next, we adopt the outside-in strategy to fill this gap.

Let $\overline{F}_0 = \overline{S} \circ M \circ \overline{S}$, where $\overline{S} = (S, S, S, S)$. Since S is an 4-bit Sbox, we can build a $(r \times \mu) \times (r \times \mu)$ table U as discussed in Section **??**. Then, we first divide $(\mathbb{F}_2^4)^{4*}$ into 5 sets $L_i (0 \leq i \leq 4)$, where $L_i = \{\alpha|$ the active pattern of α is $(b_0, \ldots, b_3)$, where $b_i = 1$, and $b_j = 0 (j \neq i)\} (0 \leq i \leq 3)$ and $L_r = \{\alpha | \boldsymbol{wt}(\alpha) \geq 2\}$. Unfortunately, there doesn't exist an β such that, for $\forall \alpha \in L_4$, it holds that $U(\alpha, \beta) = 1$. Thus, by observation and verification, we divide $(\mathbb{F}_2^4)^{4*}$ into 6 sets $L_i (0 \leq i \leq 4)$, where $L_i = \{\alpha|$ the active pattern of α is $(b_0, \ldots, b_3)$, where $b_i = 1$, and $b_j = 0 (j \neq i)\} (0 \leq i \leq 3)$, $L_4 = \{\alpha | \boldsymbol{wt}(\alpha) = 3\}$, and $L_5 = \{\alpha | \boldsymbol{wt}(\alpha) = 2 \text{ or } 4\}$. And, for each L_i, we can detect a set $\widehat{L}_i \neq \emptyset$, such that for $\forall \alpha \in L_i$ and $\forall \beta \in \widehat{L}_i$, it holds that $\alpha \xrightarrow{\overline{F}_0} \beta$. Moreover, our experiment show that if $\#\widehat{L}_i$ is too large, the process of proof may be less efficient. Thus, we choose $\#\widehat{L}_i = 16$ by experimental measurement, which allows us to finish our proof. The choice of $\#\widehat{L}_i$ is shown in Table 1. Note that, both S and M are involutive, thus $\overline{F}_0 = (\overline{F}_1)^{-1}$, which implicit that $\forall \alpha \in L_i$ and $\forall \beta \in \widehat{L}_i$, it holds that $\beta \xrightarrow{\overline{F}_1} \alpha$. With those discussion, we apply the outside-in strategy to show there no ZCLAs exist for 7-round Midori64 without the last MC layer.

Table 1: The Values of Set $\widehat{L}_i$

i	$\widehat{L}_i$
0	$\{273, 275, 276, 278, 281, 282, 284, 287, 305, 307, 308, 310, 313, 314, 316, 319\}$
1	$\{4113, 4115, 4116, 4118, 4121, 4122, 4124, 4127, 4145, 4147, 4148, 4150, 4153, 4154, 4156, 4159\}$
2	$\{4353, 4355, 4356, 4358, 4361, 4362, 4364, 4367, 4865, 4867, 4868, 4870, 4873, 4874, 4876, 4879\}$
3	$\{4368, 4400, 4416, 4448, 4496, 4512, 4544, 4592, 4880, 4912, 4928, 4960, 5008, 5024, 5056, 5104\}$
4	$\{6297, 6348, 6537, 6552, 6553, 6557, 6617, 7308, 7368, 7372, 7373, 7388, 7577, 7628, 14489, 14540\}$
5	$\{5035, 5038, 5050, 5055, 5098, 5103, 5115, 5118, 5291, 5294, 5306, 5311, 5354, 5359, 5371, 5374\}$

Application of the MITM approach. The specific steps we take are as follows.

1. We choose $F_0 = \widehat{S} \circ \widehat{P} \circ \widehat{S}$, $F = (\widehat{P} \circ \widehat{S})^3 \circ \widehat{P}$, and $F_1 = \widehat{S} \circ \widehat{P} \circ \widehat{S}$, where $\widehat{P}$ represents the composition of the operation of SR and MC.

2. For $\forall J, J' \subset \{0, 1, 2, 3\}$, $w \in W_J = \{(i_0, \ldots, i_3) \neq 0 | 0 \leq i_j \leq 5 \text{ for } j \in J, \text{ and } i_j = 0 \text{ for } j \notin J\}$, and $w' \in W_{J'} = \{(i_0, \ldots, i_3) \neq 0 | 0 \leq i_j \leq 5 \text{ for } j \in J', \text{ and } i_j = 0 \text{ for } j \notin J'\}$, let $R_{J,w} = R_0 \times \cdots \times R_{c-1} (R_j = \widehat{L}_{w_j} \text{ for } j \in J, R_j = \{0\} \text{ for } j \notin J, 0 \leq j \leq 3)$, $R'_{J',w'} = R'_0 \times \cdots \times R'_{c-1} (R'_j = \widehat{L}_{w'_j} \text{ for } j \in J', R_j = \{0\} \text{ for } j \notin J', 0 \leq j \leq 3)$. Then, we show there exits at least a $\eta \in D_{R_{J,w}}$ and a $\theta \in D^{SC^{-1}}_{R_{J,w}}$, such that $\eta \xrightarrow{F} \theta$ by SAT method.

Under the outside-in strategy, after $((6+1)^4 - 1)^2 \approx 2^{22.45}$ invocations of SAT solver, we get the following Theorem. The whole process costs around 28 days totally in a single core. Actually, we run 8 tasks parallelly, which allows us to finish our proof no more than 4 days.

Theorem 4. *For 7-round Midori64 without the last MC layer, there no ZCLAs exist even though the details of the S-box are considered, under the assumption that round keys are independent and uniformly random.*

4.2 ARIA

ARIA is a 128-bit block cipher, it is a SPN structure block cipher. The internal state is viewed as a 1×16 array of bytes as follows, where $s_i \in \mathbb{F}_2^8 (0 \leq i \leq 16)$.

$$s = (s_0, s_0, \ldots, s_{15})$$

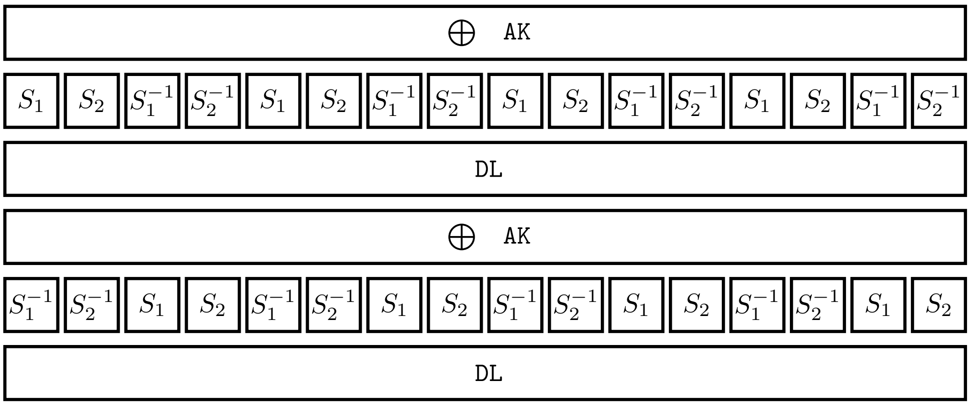
⊕ AK
S_1 S_2 S_1^{-1} S_2^{-1} S_1 S_2 S_1^{-1} S_2^{-1} S_1 S_2 S_1^{-1} S_2^{-1} S_1 S_2 S_1^{-1} S_2^{-1}
DL
⊕ AK
S_1^{-1} S_2^{-1} S_1 S_2 S_1^{-1} S_2^{-1} S_1 S_2 S_1^{-1} S_2^{-1} S_1 S_2 S_1^{-1} S_2^{-1} S_1 S_2
DL

Figure 3: Two Rounds of ARIA

Two encryption round of ARIA is illustrated in Figure 3, it is composed of the following three operations:

AddRoundKey(AK): The 128-bit round key XORed with the 128-bit state.

Substitution Layer(SL): In the odd or even number of round, applying the 8-bit S-box S_1, S_2, S_1^{-1}, and S_2^{-1}) with difference order as shown in Figure 3 to each word in parallel of the cipher internal state.

Diffusion Layer(DL): The internal state is multiplied by the involutive word-level matrix M.

Previous methods can only show that there no ZCLAs exist by limiting the input and output mask 1-word active at most, which implicit there is still a big gap to get the LNR which no ZCLAs exist. Thus, we apply the outside-in strategy to show there no ZCLAs exist for 5-round ARIA without the last linear layer.

Let $F_0 = \widehat{S}^2 \circ M \circ \widehat{S}^1$, $F = M \circ \widehat{S}^1 \circ M$, and $F_1 = \widehat{S}^1 \circ M \circ \widehat{S}^2$, where $\widehat{S}^1$ represents the substitution layer in the odd number of round and $\widehat{S}^2$ represents the substitution layer in the even number. By calculation, we have $P_{\odot}(M) = \{\boldsymbol{y_i} | 0 \le i \le 196\}$ and there exist some sets $G_0, G_1 \subseteq A_O(S_1) \bigcap A_0(S_2) \bigcap A_0(S_1^{-1}) \bigcap A_0(S_2^{-1})$, such that $G_0 \bigcap G_1 = \emptyset$. We pick $G_0 = \{1, 8\}$ and $G_1 = \{2, 4\}$. Then, $G_0 \bigcup G_1 \bigcup \{z_0 \oplus z_1 | z_0 \in G_0, z_1 \in G_1 = \{1, 2, 3, 4, 5, 8, 10, 12\}$. Let $Z_0 = \{z \in \mathbb{F}_2^{8*} | \text{ for } \forall x \in \{1, 2, 3, 4, 5, 8, 10, 12\}, x \xrightarrow{S_1} z\}$, $Z_1 = \{z \in \mathbb{F}_2^{8*} | \text{ for } \forall x \in \{1, 2, 3, 4, 5, 8, 10, 12\}, x \xrightarrow{S_2} z\}$, $Z_2 = \{z \in \mathbb{F}_2^{8*} | \text{ for } \forall x \in \{1, 2, 3, 4, 5, 8, 10, 12\}, x \xrightarrow{S_1^{-1}} z\}$, and $Z_3 = \{z \in \mathbb{F}_2^{8*} | \text{ for } \forall x \in \{1, 2, 3, 4, 5, 8, 10, 12\}, x \xrightarrow{S_2^{-1}} z\}$. It is easily to verify $Z_i \neq \emptyset (0 \le i \le 3)$. Thus, ARIA has property $P_{H,arl}^{197}$. Let $Z = (Z_2, Z_3, Z_0, Z_1, \ldots, Z_2, Z_3, Z_0, Z_1)$, then for any input mask $\alpha \in \mathbb{F}_2^{n*}$, there exist $0 \le i \le 196$, such that for $\forall \beta \in R = R_0 \times \cdots \times R_{q-1}$, where $R_j = Z_j$ when $(\boldsymbol{y_i})_j = 1$ and $R_j = \{0\}$ otherwise $(0 \le j \le q-1)$, it holds that $\alpha \xrightarrow{F_0} \beta$.

In turn, since $\widehat{S}^1 = (\widehat{S}^2)^{-1}$ and $M^{-1} = M$, we have $F_1^{-1} = (\widehat{S}^1 \circ M \circ \widehat{S}^2)^{-1} = \widehat{S}^1 \circ M \circ \widehat{S}^2$. Then, ARIA also has property $P_{T,arl}^{196}$. Let $Z' = (Z_0, Z_1, Z_2, Z_3, \ldots, Z_0, Z_1, Z_2, Z_3)$, then for any input mask $\alpha \in \mathbb{F}_2^{n*}$, there exist $0 \le i' \le 197$, such that for $\forall \beta \in R' = R_0 \times \cdots \times R_{q-1}$, where $R_j = Z_j$ when $(\boldsymbol{y_{i'}})_j = 1$ and $R_j = \{0\}$ otherwise $(0 \le j' \le q-1)$, it holds that $\alpha \xrightarrow{(F_1)^{-1}} \beta$.

Application of MITM approach. The specific steps we take are as follows.

1. We choose $F_0 = \widehat{S}^2 \circ M \circ \widehat{S}^1$, $F = M \circ \widehat{S}^1 \circ M$, and $F_1 = \widehat{S}^1 \circ M \circ \widehat{S}^2$.
2. For each $0 \leq i, i' \leq 196$, we construct the set R and R' and show there exits at least a $\eta \in R$ and a $\theta \in R'$, such that $\eta \xrightarrow{F} \theta$ by SAT method.

Under the outside-in strategy, after $(197) \approx 2^{15.2}$ invocations of SAT solver, we get the following Theorem. The whole process costs around 8 days totally in a single core. Actually, we run 8 tasks parallelly, which allows us to finish our proof in around 1 day.

Theorem 5. *For 5-round ARIA without the last MC linear layer, there no ZCLAs exist even though the details of the S-box are considered, under the assumption that round keys are independent and uniformly random.*

5 Evaluating the Security by Double-Collision Approach

In this part, we give our double-collision approach to show there no ZCLAs exist for a block cipher E. With different statements of Theorem 1, we have the following Corollary.

Corollary 1. *Let α and β be the input mask and output mask of a block cipher $E = F_1 \circ F \circ F_0$, donate $\Phi_I(\beta) = \{\gamma_i \in \mathbb{F}_2^{n*} | \gamma_i \xrightarrow{F_0} \beta\}$ and $\Phi_O(\alpha) = \{\gamma_o \in \mathbb{F}_2^{n*} | \alpha \xrightarrow{F_1} \gamma_o\}$, if the following conditions are hold:*

1. *There exits h sets $H_i (0 \leq i \leq h-1)$ and t sets $T_j (0 \leq j \leq t-1)$, such that $H_0 \bigcup \cdots \bigcup H_{h-1} = \mathbb{F}_2^{n*}$ and $T_0 \bigcup \cdots \bigcup T_{t-1} = \mathbb{F}_2^{n*}$.*
2. *For each $i, j (0 \leq i \leq h-1, 0 \leq j \leq t-1)$, there exists $\eta, \theta \in \mathbb{F}_2^{n*}$, such that $H_i \subseteq \Phi_I(\eta)$, $T_j \in \Phi_O(\theta)$, and $\eta \xrightarrow{F_I} \theta$.*

Then, there no ZCLA exists for E.

Let $\widehat{H}_i' (0 \leq i \leq h-1)$ be a set such that, for $\forall \alpha \in \widehat{H}_i'$, all mask in H_i may propagate to α through F_0 forward. Analogously, let $\widehat{T}_j' (0 \leq j \leq t-1)$ be a set such that, for $\forall \alpha \in \widehat{T}_j'$, all mask in T_j may propagate to α through F_1 backward. Then, the double-collision approach can be summarized as follows.

Function divide. Choosing appropriate F_1, F, and F_0, such that $E = F_1 \circ F \circ F_0$.

State divide. Choosing appropriate h and t, H_i and $\widehat{H}_i (0 \leq i \leq h-1)$, and T_j and $\widehat{T}_j (0 \leq j \leq t-1)$, such that condition 2 holds, the technique are .

Masks connection. For each $\widehat{H}_i (0 \leq i \leq h-1)$ and $\widehat{T}_j (0 \leq j \leq t-1)$, we get an $\eta_i \in \widehat{H}_i$ and an $\theta \in \widehat{T}_j$, such that $\theta_j \xrightarrow{F} \theta$.

Double collision. For each $\eta_i (0 \leq i \leq h-1)$ and $\theta_j (0 \leq j \leq t-1)$, we show $H_i \subseteq \Phi_I(\eta)$ and $T_j \in \Phi_O(\theta)$.

6 Application Double-Collision Approach to AES

AES is a 128-bit block cipher, it supports the key size of 128, 192, and 256-bit. There is no doubt that AES is one of the most famous block ciphers all around the world. Its design philosophy has had a profound impact on block ciphers. The internal state is viewed as a 4×4 square array of bytes as follows, where $s_i \in \mathbb{F}_2^8 (0 \leq i \leq 16)$.

$$s = \begin{pmatrix} s_0 & s_4 & s_8 & s_{12} \\ s_1 & s_5 & s_9 & s_{13} \\ s_2 & s_6 & s_{10} & s_{14} \\ s_3 & s_7 & s_{11} & s_{15} \end{pmatrix}.$$

One encryption round of AES is illustrated in Figure 4, it is composed of the following four operations:

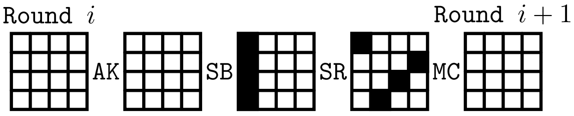

Figure 4: One Round of AES-128

AddRoundKey(AK): The 128-bit round key which is derived from the key schedule XORed with the state.

SubBytes(SB): Applying the 8-bit S-box S to each byte in parallel of the cipher internal state.

ShiftRows(SR): The i-th rows $(0 \leq i \leq 3)$ of the internal state is rotated by i bytes form right to left.

MixColumns(MC): Each column of the internal state is multiplied by the MDS matrix M.

As AES plays an important role in the development of block ciphers, the security of it has received extensive research. Significant work has been achieved in the research of impossible differentials. In 2018, Wang et al. [WJ18] show that there no 1-word active input and output impossible differentials exist for 5-round AES without the last MC layer in the case of considering the key schedule. Later, they [WJ19] show that there no impossible differentials exist for 5-round AES without the last MC layer in the case of considering the key schedule. However, their approach is specific and cannot applied to analysis the existence of ZCLAs. Besides, In EUROCRYPT2016, Sun et al. [SLG+16] showed that there exist no zero-correlation linear approximations covering more than four rounds for AES, but they work cannot consider the details of Sbox. Thus, the LNR which no ZCLAs exist for AES in the case of considering the details of the Sbox is still unknown. To fill this gap, with the technique in Section **??**, we study the property of the S-box of AES first.

Property 3. Let S be the S-box of AES, it holds that $M_I(S) = M_O(S) = 2$.

Under the Property 3, we can get A_I and A_O by exhaustive search as discussion in Section **??**. Let $\overline{F}_0 = \overline{S} \circ M \circ \overline{S}$, $\overline{S} = (S, S, S, S)$. Since $branch(M) = 5$, for $\forall \alpha \in (\mathbb{F}_2^8)^{4*}$, there may exists a β, such that $\alpha \xrightarrow{\overline{F}_0} \beta$. In turn, for $\forall \beta \in (\mathbb{F}_2^8)^{4*}$, there may exists a α, such that $\alpha \xrightarrow{\overline{F}_0} \beta$. Thus, we adopt the inside-out strategy to show there no ZCLAs exist for 5-round AES without the last MC layer in the case of considering the details of S-box.

Application of the inside-out strategy. The specific steps we take are as follows.

1. We choose $F_0 = \widehat{S} \circ \widehat{P} \circ \widehat{S}$, $F = \widehat{P} \circ \widehat{S} \circ \widehat{P}$, and $F_1 = \widehat{S} \circ \widehat{P} \circ \widehat{S}$, where $\widehat{P}$ represents the composition of the operation of SR and MC.
2. For $\forall J, J' \subset \{0,1,2,3\}$, let $R_J = R_0 \times \cdots \times R_{c-1}(R_j = \mathbb{F}_2^{8*}$ for $j \in J, R_j = \{0\}$ for $j \notin J, 0 \leq j \leq 3)$, $R_{J'} = R_0 \times \cdots \times R_{c-1}(R_j = \mathbb{F}_2^{8*}$ for $j \in J', R_j = \{0\}$ for $j \notin J', 0 \leq j \leq 3)$. Then,

(a) Detecting an $\alpha \in D_{R_J}$ and an $\beta \in D_{R_{J'}}^{SR^{-1}}$, such that $\alpha \xrightarrow{F} \beta$ by SAT method.

(b) For each $\alpha_j \in (\mathbb{F}_2^{8*})^4 (j \in J)$, we determine whether for $\forall \gamma_j \in (\mathbb{F}_2^8)^{4*}$, such that $\gamma_j \xrightarrow{\overline{F_0}} \alpha_j$; And for each $\beta_{j'} \in (\mathbb{F}_2^{8*})^4 (j' \in J')$, we determine whether for $\forall \delta_{j'} \in (\mathbb{F}_2^8)^{4*}$, such that $\beta_{j'} \xrightarrow{\overline{F_1}} \delta_{j'}$.

If conditions in 2(b) hold, then for $\forall \gamma \in D_{R_J}^{SR^{-1}}$ and $\forall \delta \in D_{R_{J'}}$, it holds that $\gamma \xrightarrow{F_1 \circ F \circ F_0} \delta$. Otherwise, we repeat the steps of 2(a) and 2(b).

Note that, the verification of 2(b) can be achieved by random verification. Picking $j \in J$, we use the determination that for $\forall \gamma_j \in (\mathbb{F}_2^8)^{4*}$, such that $\gamma_j \xrightarrow{\overline{F_0}} \alpha_j$ as an example. Here, for the sake of narrative, we donate $Q^b = Q_0 \times \cdots \times Q_3 (Q_i = \widetilde{O}_i$ if $b_i = 1$, and $Q_i = \{0\}$ otherwise), where $\widetilde{O}_0, \ldots, \widetilde{O}_3 \in A_O$ and $b = (b_0, \ldots, b_3) \in \mathbb{F}_2^4$. For each $b \in \mathbb{F}_2^{4*}$, we pick $\widetilde{O}_0, \ldots, \widetilde{O}_3 \in A_O$ randomly, and verify for $\forall \epsilon_j \in \{M^t x | x \in Q^b\}$, it hold that $\varepsilon_j \xrightarrow{\overline{S}} \alpha_j$. If for all $b \in \mathbb{F}_2^{4*}$, there always exist $\widetilde{O}_0, \ldots, \widetilde{O}_3 \in A_O$, such that $\forall \epsilon_j \in \{M^t x | x \in Q^b\}$, it hold that $\varepsilon_j \xrightarrow{\overline{S}} \alpha_j$. Then, $\forall \gamma_j \in (\mathbb{F}_2^8)^{4*}$, it holds that $\gamma_j \xrightarrow{\overline{F_0}} \alpha_j$.

In our experiments, the repetition of steps of 2(a) and 2(b) is never occur. That is, by $15 \times 15 = 225$ invocations of SAT solver and some time for random verification, we get the following Theorem. The whole process costs 3356 seconds totally, which is very efficient.

Theorem 6. *For 5-round AES without the last MC layer, there no ZCLAs exist even though the details of the S-box are considered, under the assumption that round keys are independent and uniformly random.*

To show the correctness of our method, we give specific propagation of masks as an example.

Example 1. For $J = \{0\}$ and $J' = \{0, 1\}$, let

$$\alpha = \begin{pmatrix} 94 & 0 & 0 & 0 \\ 247 & 0 & 0 & 0 \\ 96 & 0 & 0 & 0 \\ 195 & 0 & 0 & 0 \end{pmatrix} \in D_{R_J}, \beta = \begin{pmatrix} 12 & 120 & 0 & 0 \\ 0 & 166 & 11 & 0 \\ 0 & 0 & 196 & 209 \\ 78 & 0 & 0 & 184 \end{pmatrix} \in D_{R_{J'}}^{SR^{-1}}.$$

Then, it can be verified that $\alpha \xrightarrow{F} \beta$. Besides, as shown is Table 2, Table 3 and Table 4, $\forall \gamma_0^{SR^{-1}} \in (\mathbb{F}_2^8)^{4*}$, there exists $1 \leq i \leq 15$, such that $\gamma_0^{SR^{-1}} \in \Lambda_i^1$. Then, there exists $\overline{\gamma}_0^{SR^{-1}} \in \overline{\Lambda^1}_i$, such that $\gamma_0^{SR^{-1}} \xrightarrow{\overline{S}} \overline{\gamma}_0^{SR^{-1}}$. Meanwhile, we can verify that $\forall \epsilon_0 \in \{M^t x | x \in \overline{\Lambda^1}_i\}$, it hold that $\varepsilon_0 \xrightarrow{\overline{S}} [94, 247, 96, 195]$. Thus, $\forall \gamma_0^{SR^{-1}} \in (\mathbb{F}_2^8)^{4*}$, we have $\gamma_0^{SR^{-1}} \xrightarrow{\overline{F_0}} [94, 247, 96, 195]$. With the same reason, for $\forall \delta_0, \delta_1 \in (\mathbb{F}_2^8)^{4*}$, we have $[12, 166, 196, 184] \xrightarrow{\overline{F_1}} \delta_0$ and $[120, 11, 209, 78] \xrightarrow{\overline{F_1}} \delta_1$. Therefore, for $\forall \gamma \in D_{R_J}^{SR^{-1}}$ and $\forall \delta \in D_{R_{J'}}$, it holds that $\gamma \xrightarrow{F_1 \circ F \circ F_0} \delta$.

7 Conclusion and Future work

In this paper, we propose some method for ZCLAs. Those methods are able to search the ZCLAs and get the LNR which no 1-bit AIAO-ZCLAs exist by considering all the details of S-box for block ciphers when the sum of input and output bits of S-box is more than 16-bit. Notable, those method can also get the LNR which no ZCLAs exist and all ZCLAs SPN structure block ciphers without the limitation of input mask and output mask.

Table 2: Result of Random Verification of $[94, 247, 96, 195]$

1	Λ_i^1	$\overline{\Lambda^1}_i$
2	$\mathbb{F}_2^{8*} \times \{0\} \times \{0\} \times \{0\}$	$\{1,2\} \times \{0\} \times \{0\} \times \{0\}$
3	$\{0\} \times \mathbb{F}_2^{8*} \times \{0\} \times \{0\}$	$\{0\} \times \{2,4\} \times \{0\} \times \{0\}$
4	$\mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*} \times \{0\} \times \{0\}$	$\{1,2\} \times \{1,5\} \times \{0\} \times \{0\}$
5	$\{0\} \times \{0\} \times \mathbb{F}_2^{8*} \times \{0\}$	$\{0\} \times \{0\} \times \{1,5\} \times \{0\}$
6	$\mathbb{F}_2^{8*} \times \{0\} \times \mathbb{F}_2^{8*} \times \{0\}$	$\{1,2\} \times \{0\} \times \{4,6\} \times \{0\}$
7	$\{0\} \times \mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*} \times \{0\}$	$\{0\} \times \{1,2\} \times \{1,2\} \times \{0\}$
8	$\mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*} \times \{0\}$	$\{1,2\} \times \{3,5\} \times \{1,2\} \times \{0\}$
9	$\{0\} \times \{0\} \times \{0\} \times \mathbb{F}_2^{8*}$	$\{0\} \times \{0\} \times \{0\} \times \{1,2\}$
10	$\mathbb{F}_2^{8*} \times \{0\} \times \{0\} \times \mathbb{F}_2^{8*}$	$\{1,2\} \times \{0\} \times \{0\} \times \{1,5\}$
11	$\{0\} \times \mathbb{F}_2^{8*} \times \{0\} \times \mathbb{F}_2^{8*}$	$\{0\} \times \{1,2\} \times \{0\} \times \{2,4\}$
12	$\mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*} \times \{0\} \times \mathbb{F}_2^{8*}$	$\{1,2\} \times \{2,4\} \times \{0\} \times \{1,8\}$
13	$\{0\} \times \{0\} \times \mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*}$	$\{0\} \times \{0\} \times \{1,2\} \times \{2,4\}$
14	$\mathbb{F}_2^{8*} \times \{0\} \times \mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*}$	$\{1,2\} \times \{0\} \times \{1,2\} \times \{1,8\}$
15	$\{0\} \times \mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*}$	$\{0\} \times \{1,2\} \times \{2,4\} \times \{1,9\}$

Table 3: Result of Random Verification of $[12, 166, 196, 184]$

Λ_i^2	$\overline{\Lambda^2}_i$
$\mathbb{F}_2^{8*} \times \{0\} \times \{0\} \times \{0\}$	$\{1,2\} \times \{0\} \times \{0\} \times \{0\}$
$\{0\} \times \mathbb{F}_2^{8*} \times \{0\} \times \{0\}$	$\{0\} \times \{1,8\} \times \{0\} \times \{0\}$
$\mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*} \times \{0\} \times \{0\}$	$\{1,2\} \times \{5,17\} \times \{0\} \times \{0\}$
$\{0\} \times \{0\} \times \mathbb{F}_2^{8*} \times \{0\}$	$\{0\} \times \{0\} \times \{1,2\} \times \{0\}$
$\mathbb{F}_2^{8*} \times \{0\} \times \mathbb{F}_2^{8*} \times \{0\}$	$\{1,2\} \times \{0\} \times \{9,10\} \times \{0\}$
$\{0\} \times \mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*} \times \{0\}$	$\{0\} \times \{1,2\} \times \{5,6\} \times \{0\}$
$\mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*} \times \{0\}$	$\{1,2\} \times \{1,2\} \times \{9,10\} \times \{0\}$
$\{0\} \times \{0\} \times \{0\} \times \mathbb{F}_2^{8*}$	$\{0\} \times \{0\} \times \{0\} \times \{1,2\}$
$\mathbb{F}_2^{8*} \times \{0\} \times \{0\} \times \mathbb{F}_2^{8*}$	$\{1,2\} \times \{0\} \times \{0\} \times \{8,13\}$
$\{0\} \times \mathbb{F}_2^{8*} \times \{0\} \times \mathbb{F}_2^{8*}$	$\{0\} \times \{1,2\} \times \{0\} \times \{7,9\}$
$\mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*} \times \{0\} \times \mathbb{F}_2^{8*}$	$\{1,2\} \times \{1,2\} \times \{0\} \times \{17,18\}$
$\{0\} \times \{0\} \times \mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*}$	$\{0\} \times \{0\} \times \{1,2\} \times \{6,11\}$
$\mathbb{F}_2^{8*} \times \{0\} \times \mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*}$	$\{1,2\} \times \{0\} \times \{1,2\} \times \{36,47\}$
$\{0\} \times \mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*}$	$\{0\} \times \{1,2\} \times \{1,2\} \times \{8,13\}$

Table 4: Result of Random Verification of $[120, 11, 209, 78]$

Λ_i^3	$\overline{\Lambda^3}_i$
$\mathbb{F}_2^{8*} \times \{0\} \times \{0\} \times \{0\}$	$\{1,8\} \times \{0\} \times \{0\} \times \{0\}$
$\{0\} \times \mathbb{F}_2^{8*} \times \{0\} \times \{0\}$	$\{0\} \times \{1,2\} \times \{0\} \times \{0\}$
$\mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*} \times \{0\} \times \{0\}$	$\{1,2\} \times \{7,9\} \times \{0\} \times \{0\}$
$\{0\} \times \{0\} \times \mathbb{F}_2^{8*} \times \{0\}$	$\{0\} \times \{0\} \times \{3,5\} \times \{0\}$
$\mathbb{F}_2^{8*} \times \{0\} \times \mathbb{F}_2^{8*} \times \{0\}$	$\{1,2\} \times \{0\} \times \{5,11\} \times \{0\}$
$\{0\} \times \mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*} \times \{0\}$	$\{0\} \times \{1,2\} \times \{7,9\} \times \{0\}$
$\mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*} \times \{0\}$	$\{1,2\} \times \{1,2\} \times \{8,9\} \times \{0\}$
$\{0\} \times \{0\} \times \{0\} \times \mathbb{F}_2^{8*}$	$\{0\} \times \{0\} \times \{0\} \times \{2,10\}$
$\mathbb{F}_2^{8*} \times \{0\} \times \{0\} \times \mathbb{F}_2^{8*}$	$\{1,2\} \times \{0\} \times \{0\} \times \{9,12\}$
$\{0\} \times \mathbb{F}_2^{8*} \times \{0\} \times \mathbb{F}_2^{8*}$	$\{0\} \times \{1,2\} \times \{0\} \times \{12,15\}$
$\mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*} \times \{0\} \times \mathbb{F}_2^{8*}$	$\{1,2\} \times \{1,2\} \times \{0\} \times \{8,9\}$
$\{0\} \times \{0\} \times \mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*}$	$\{0\} \times \{0\} \times \{1,2\} \times \{11,12\}$
$\mathbb{F}_2^{8*} \times \{0\} \times \mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*}$	$\{1,2\} \times \{0\} \times \{1,2\} \times \{12,15\}$
$\{0\} \times \mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*} \times \mathbb{F}_2^{8*}$	$\{0\} \times \{1,2\} \times \{1,2\} \times \{12,15\}$

As a result, we apply our method to derive new ZCLAs for the block ciphers MISTY1 and MISTY2, and show there no 5, 7, 5-round ZCLAs exist for AES, Midori64 and ARIA. What's more, we get all 4-round ZCLAs for ARIA. Those results show that our methods are indeed the advanced search method of ZCLAs from design and cryptanalysis aspects.

Proving the security of block ciphers against known attacks is very import, one of our works shows the security of block ciphers against zero-correlation linear attack for the view of the LNR which no ZCLAs exist for SPN structure. Unfortunately, our method is not suited for other structures. How to extend this work to more structures is still unknown.

References

[AST$^+$17] Ahmed Abdelkhalek, Yu Sasaki, Yosuke Todo, Mohamed Tolba, and Amr M. Youssef. MILP modeling for (large) s-boxes to optimize probability of differential characteristics. *IACR Trans. Symmetric Cryptol.*, 2017(4):99–129, 2017.

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

[Bih94] Eli Biham. On matsui's linear cryptanalysis. In Alfredo De Santis, editor, *Advances in Cryptology - EUROCRYPT '94, Workshop on the Theory and Application of Cryptographic Techniques, Perugia, Italy, May 9-12, 1994, Proceedings*, volume 950 of *Lecture Notes in Computer Science*, pages 341–355. Springer, 1994.

[BLNW12] Andrey Bogdanov, Gregor Leander, Kaisa Nyberg, and Meiqin Wang. Integral and multidimensional linear distinguishers with correlation zero. In Xiaoyun Wang and Kazue Sako, editors, *Advances in Cryptology - ASIACRYPT 2012 - 18th International Conference on the Theory and Application of Cryptology and Information Security, Beijing, China, December 2-6, 2012. Proceedings*, volume 7658 of *Lecture Notes in Computer Science*, pages 244–261. Springer, 2012.

[BR14] Andrey Bogdanov and Vincent Rijmen. Linear hulls with correlation zero and linear cryptanalysis of block ciphers. *Des. Codes Cryptogr.*, 70(3):369–383, 2014.

[BW12] Andrey Bogdanov and Meiqin Wang. Zero correlation linear cryptanalysis with reduced data complexity. In Anne Canteaut, editor, *Fast Software Encryption - 19th International Workshop, FSE 2012, Washington, DC, USA, March 19-21, 2012. Revised Selected Papers*, volume 7549 of *Lecture Notes in Computer Science*, pages 29–48. Springer, 2012.

[CCH10] Claude Carlet, Yves Crama, and Peter L. Hammer. Boolean functions for cryptography and error-correcting codes. In Yves Crama and Peter L. Hammer, editors, *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*, pages 257–397. Cambridge University Press, 2010.

[CJF$^+$16] Tingting Cui, Keting Jia, Kai Fu, Shiyao Chen, and Meiqin Wang. New automatic search tool for impossible differentials and zero-correlation linear approximations. *IACR Cryptol. ePrint Arch.*, 2016:689, 2016.

[DGV94] Joan Daemen, René Govaerts, and Joos Vandewalle. Correlation matrices. In Bart Preneel, editor, *Fast Software Encryption: Second International Workshop. Leuven, Belgium, 14-16 December 1994, Proceedings*, volume 1008 of *Lecture Notes in Computer Science*, pages 275–285. Springer, 1994.

[DR02] Joan Daemen and Vincent Rijmen. *The Design of Rijndael: AES - The Advanced Encryption Standard.* Information Security and Cryptography. Springer, 2002.

[KKP$^+$03] Daesung Kwon, Jaesung Kim, Sangwoo Park, Soo Hak Sung, Yaekwon Sohn, Jung Hwan Song, Yongjin Yeom, E-Joong Yoon, Sangjin Lee, Jaewon Lee, Seongtaek Chee, Daewan Han, and Jin Hong. New block cipher: ARIA. In Jong In Lim and Dong Hoon Lee, editors, *Information Security and Cryptology - ICISC 2003, 6th International Conference, Seoul, Korea, November 27-28, 2003, Revised Papers*, volume 2971 of *Lecture Notes in Computer Science*, pages 432–445. Springer, 2003.

[Mat93] Mitsuru Matsui. Linear cryptanalysis method for DES cipher. In Tor Helleseth, editor, *Advances in Cryptology - EUROCRYPT '93, Workshop on the Theory and Application of of Cryptographic Techniques, Lofthus, Norway, May 23-27, 1993, Proceedings*, volume 765 of *Lecture Notes in Computer Science*, pages 386–397. Springer, 1993.

[Mat97] Mitsuru Matsui. New block encryption algorithm MISTY. In Eli Biham, editor, *Fast Software Encryption, 4th International Workshop, FSE '97, Haifa, Israel, January 20-22, 1997, Proceedings*, volume 1267 of *Lecture Notes in Computer Science*, pages 54–68. Springer, 1997.

[Nyb94] Kaisa Nyberg. Linear approximation of block ciphers. In Alfredo De Santis, editor, *Advances in Cryptology - EUROCRYPT '94, Workshop on the Theory and Application of Cryptographic Techniques, Perugia, Italy, May 9-12, 1994, Proceedings*, volume 950 of *Lecture Notes in Computer Science*, pages 439–444. Springer, 1994.

[SHW$^+$14a] Siwei Sun, Lei Hu, Meiqin Wang, Peng Wang, Kexin Qiao, Xiaoshuang Ma, Danping Shi, Ling Song, and Kai Fu. 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. https://eprint.iacr.org/2014/747.

[SHW$^+$14b] Siwei Sun, Lei Hu, Peng Wang, Kexin Qiao, Xiaoshuang Ma, and Ling Song. Automatic security evaluation and (related-key) differential characteristic search: Application to simon, present, lblock, DES(L) and other bit-oriented block ciphers. In Palash Sarkar and Tetsu Iwata, editors, *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*, volume 8873 of *Lecture Notes in Computer Science*, pages 158–178. Springer, 2014.

[SLG$^+$16] Bing Sun, Meicheng Liu, Jian Guo, Vincent Rijmen, and Ruilin Li. Provable security evaluation of structures against impossible differential and zero correlation linear cryptanalysis. In Marc Fischlin and Jean-Sébastien Coron, editors, *Advances in Cryptology - EUROCRYPT 2016 - 35th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Vienna, Austria, May 8-12, 2016, Proceedings, Part I*, volume 9665 of *Lecture Notes in Computer Science*, pages 196–213. Springer, 2016.

[ST17] Yu Sasaki and Yosuke Todo. New impossible differential search tool from design and cryptanalysis aspects - revealing structural properties of several ciphers. In Jean-Sébastien Coron and Jesper Buus Nielsen, editors, *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*, volume 10212 of *Lecture Notes in Computer Science*, pages 185–215, 2017.

[WJ18] Qian Wang and Chenhui Jin. Upper bound of the length of truncated impossible differentials for AES. *Des. Codes Cryptogr.*, 86(7):1541–1552, 2018.

[WJ19] Qian Wang and Chenhui Jin. A method to bound the number of active s-boxes for a kind of aes-like structure. *Comput. J.*, 62(8):1121–1131, 2019.

A Finding the set $\widetilde{I}$ with minimum number of elements

Let us focus the following problem.

Problem 1. *For a give integer N and a m-bit S-box S, whether there exists a set $\widetilde{I} \subseteq \mathbb{F}_2^{m*}$ with $|\widetilde{I}| = N$ and $\bigcup\limits_{\alpha\in\widetilde{I}} \phi_o(\alpha) = \mathbb{F}_2^{m*}$?*

If for $\forall N(1 \leq N \leq 2^m - 1)$, we can solve the Problem 1. Then, we can test N for 1 to $2^m - 1$ until the Problem 1 has a solution. In this time, the solution of Problem 1 is the set $\widetilde{I}$ we desired. Thus, solving the problem that finding a $\widetilde{I}$ with the minimum number of elements turns into solving Problem 1. Again, to solve Problem 1, we study the following Problem.

Problem 2. *Let S be an m-bit S-box, $P_i = \{j \in \mathbb{F}_2^{m*} | i \xrightarrow{S} j\}$, and $Q_j = \{i \in \mathbb{F}_2^{m*} | i \xrightarrow{S} j\}$, whether there exists a solution $\{x_i\}$ for*

$$\begin{cases} \sum\limits_{i=1}^{2^m-1} x_i = N, \\ t_{i,j} = x_i (j \in P_i, 1 \leq i \leq 2^m - 1), \\ t_{i,j} = 0 (j \notin P_i, 1 \leq i \leq 2^m - 1), \\ \bigvee\limits_{i\in Q_j} t_{i,j} = 1 (1 \leq j \leq 2^m - 1), \\ x_i \in \{0,1\}, t_{i,j} \in \{0,1\} (1 \leq i,j \leq 2^m - 1)? \end{cases}$$

Note that Problem 2 is a SAT problem and can be solved by the SAT solvers. More importantly, Problem 2 is equivalent to Problem 1. That is, the solutions of Problem 1 and Problem 2 can be derived from each other, which leads to the following lemma.

Theorem 7. *For an integer N, if $\widetilde{I}$ is a solution of Problem 1, let $x_i = 1(i \in \widetilde{I})$ and $x_i = 0(i \notin \widetilde{I})$, then $\{x_i\}$ is a solution of Problem 2. In turn, if $\{x_i\}$ is a solution of Problem 2, then $\widetilde{I} = \{i | x_i = 1\}$ is a solution of Problem 1.*

Proof. On the one hand, since $\widetilde{I}$ is a solution of Problem 1, then $|\widetilde{I}| = N$ and $\bigcup\limits_{\alpha\in\widetilde{I}} \phi_o(\alpha) = \mathbb{F}_2^{m*}$. Let $x_i = 1(i \in \widetilde{I})$ and $x_i = 0(i \notin \widetilde{I})$, then $\sum\limits_{i=1}^{2^m-1} x_i = N$. Since $\bigcup\limits_{\alpha\in\widetilde{I}} \phi_o(\alpha) = \mathbb{F}_2^{m*}$, for $\forall j \in \mathbb{F}_2^{m*}$, there exists at least one $i'(i' \in \widetilde{I})$, such that $i' \overset{E}{\to} j$, thus $i' \in Q_j$ and $x_{i'} = 1$. Let $t_{i,j} = x_i(j \in P_i, 1 \leq i \leq 2^m - 1)$ and $t_{i,j} = 0(j \notin P_i, 1 \leq i \leq 2^m - 1)$, it holds that $\bigvee\limits_{i\in Q_j} t_{i,j} = 1(1 \leq j \leq 2^m - 1)$. Thus, we have constructed a solution x_i of Problem 2.

On the other hand, if $x_i(1 \leq i \leq 2^m - 1)$ is a solution of Problem 2, let $\widetilde{I} = \{i|x_i = 1\}$. Then, we have $|\widetilde{I}| = N$. Moreover, for $\forall j \in \mathbb{F}_2^{m*}$, since $\bigvee\limits_{i\in Q_j} t_{i,j} = 1$, then there exists at least one $i' \in Q_j$, such that $t_{i',j} = 1$, thus $x_{i'} = t_{i',j} = 1$, which implies $i' \in \widehat{I}$. Thus, we have $\bigcup\limits_{i\in\widetilde{I}} \phi_o(j) = \mathbb{F}_2^{m*}$. □

From the discussion above, for a given Sbox S, we can get the set $\widetilde{I}$ with minimum number of elements. For the sake of simplicity, we donate such minimum number as $M_I(S)$. Moreover, we can get the set $A_I = \{\widetilde{I}|\#\widetilde{I} = M_I(S)\}$ who contains all set $\widetilde{I}$ with minimum number of elements in the following two ways:

SAT solver Once we get a solution, we remove this solution and ask SAT solver for another solution until SAT solver return no solution. Then, we can get the set A_I.

Exhaustive search If permits, we can exhaustive search the set $\underbrace{F_2^{m*} \times \cdots \times F_2^{m*}}_{M_I(S)}$ to get all $\widetilde{I}$ with $\#\widetilde{I} = M_i(S)$.