**Improvement of Search Algorithm for Integral Distinguisher in Subblock-Based Block Cipher**

Haruhsia Kosuge, Hidema Tanaka  
National Defense Academy of Japan, Yokosuka, Japan

**Abstract**

Integral distinguisher is the main factor of integral attack. Conventionally, higher order integral distinguisher is obtained as an extension of first order integral (conventional algorithm). The algorithm was applied to many subblock-based block ciphers, however, the conventional algorithm has some problems. We find other integral distinguisher of two subblock-based block ciphers, TWINE and LBlock, which are different from the conventional evaluations. As a solution, we propose a new algorithm to search for higher order integral distinguisher. The point of a proposal algorithm is exploitation of bijective and injective components of cipher functions. Applying the proposal algorithm to TWINE and LBlock, we confirm the results of the proposal algorithm are consistent with the results which are calculated from computer experiment. The results are the optimal distinguisher and the most advantageous one for the attackers. Our proposal algorithm contributes to development of stronger block ciphers by obtaining such integral distinguisher.

**Keywords**  
Chosen plaintext attack, Light-weight block cipher, Integral attack, Saturation attack

**1. Introduction**

1.1. Background

Considering unit of operation in block ciphers, they are roughly divided into subblock-based block ciphers and bit-based ones. We define subblock-based block cipher by a block cipher in which all operations are executed in m-bit unit (m > 2), and we call the operation unit as subblock. Note that a block is divided into N subblocks in this paper. Also, we define bit-based block cipher by block ciphers in which not all operations are executed in the same bit unit. Subblock has a significant impact on the performances for software implementation [1].

Integral attack is one of the powerful chosen plaintext attacks, and the attack is effective against subblock-based block ciphers [2]. Recently, Todo shows the first full-round attack for MISTY1 [3] by integral attack [4]. Integral attack was originally proposed as SQUARE attack [5], and Knudsen et al. formalized it as integral attack [2]. The attacker can guess some round keys based on integral distinguisher. Integral distinguisher is the main factor of integral attack. It is obtained by $2^{nm}$ chosen plaintexts, where $n (n < N)$ is the order of integral distinguisher. A set of $2^{nm}$ plaintexts is encrypted for multiple rounds to make a set of outputs. An integrated value of the set of outputs is calculated. If there exist subblocks which are always 0 in such integrated value, we can define integral distinguisher. We call such subblock as balanced subblock. Knudsen et al.
studied integral distinguisher as first order initially, and considered higher order integral distinguisher as extension of first order one [2]. We call this algorithm as conventional algorithm.

Integral attack was applied to many subblock-based block ciphers, and these attacks are based on the conventional algorithm [6][7][8][9][10][11][12]. In such applications, we point out there is still a problem in Feistel structures whose \( N \) is large such as TWINE [13] and LBlock [8] (\( N = 16 \)). We found fifteenth order integral distinguisher of TWINE and LBlock which contradict results of the conventional algorithm by computer experiment in our previous work [10][13][14]. Also, we point out there is a restriction of input conditions in the conventional algorithm. Note that we define input condition by a manner to set chosen plaintexts. We can see that input conditions in the scope of the conventional algorithm are only ones of first order and their extensions.

1.2. Contribution

In this paper, we propose a new algorithm to search for higher order integral distinguisher in subblock-based block ciphers. We search for higher order integral properties from input to output (from plaintext side to ciphertext side). In other word, we do not use extension which is used in the conventional algorithm.

Bijective and injective components of cipher functions are exploited in the proposal algorithm. We divide the proposal algorithm into two algorithms. One is Algorithm A which exploits bijective component, and we use a new idea, bijective path. The other is Algorithm B which exploits injective component, and we use another new idea, independent.

We apply the proposal algorithm to TWINE and LBlock. As a result, we confirm that results of the proposal algorithm are consistent with ones calculated from computer experiment [14], and we conclude that they are the optimal integral distinguisher. Also, all of the input conditions are in the scope of the proposal algorithm. Therefore, we conclude that the proposal algorithm is more effective than the conventional one.

If integral distinguisher which holds in additional rounds is constructed, the number of rounds to be attacked can be extended. Even if only the number of balanced subblocks increases, it is also advantageous for the attacker. As the number of balanced subblocks increases, the number of round keys the attacker can guess from single integral distinguisher increases. In other word, the attacker needs less chosen plaintexts to guess all of the secret keys. From the viewpoint of designers, they need to select stronger cipher algorithm and key schedule by considering such vulnerabilities, and the proposal algorithm contributes to the purposes.

2. PRELIMINARIES

We use the notations shown in Table 1.

2.1. Integral Distinguisher

When the attacker chooses \( n \) subblocks as variable, he needs to prepare \( 2^{mn} \) plaintexts, and we call \( n \) as order. In \( 2^{mn} \) plaintexts, a concatenation of variable subblocks takes every possible
element of $F_{2^{mn}}$ and one of the constant subblocks takes a constant value. Let $S = \{S_i | S_i \in F_{2^{mn}}^N, 0 \leq i \leq 2^{mn} - 1\}$ be a set of chosen plaintexts which satisfies

$$e_0 || e_1 || ... || e_{n-1} = \{0, 1, ..., 2^{mn} - 1\},$$
$$c_0 || c_1 || ... || c_{N-m-1} = \text{const},$$

(1)

where $e_0 || e_1 || ... || e_{n-1}$ denotes a concatenation of variable subblocks, and $c_0 || c_1 || ... || c_{N-m-1}$ a concatenation of constant subblocks. We define input condition by such manner to define set of chosen plaintexts $S$. Let $\alpha$ be an index set of variable subblocks, and we denote input condition by $\alpha$.

Let $Enc^\gamma$ be $\gamma$ rounds cipher function and $RK^{1,2,3,...,\gamma}$ be a set of round keys used from first to $\gamma$-th round encryption. Let $S' = Enc^\gamma(S, RK^{1,2,3,...,\gamma})$ be a set of outputs, we define integral of $S'$ as follows.

$$\int S' = \bigoplus_{i=0}^{2^{mn}-1} Enc^\gamma(S_i, RK^{1,2,3,...,\gamma}),$$

(2)

<table>
<thead>
<tr>
<th>Notation</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R$</td>
<td>number of rounds which a cipher function iterates.</td>
</tr>
<tr>
<td>$N$</td>
<td>number of subblocks.</td>
</tr>
<tr>
<td>$m$</td>
<td>bit-length of a subblock ($m \geq 2$).</td>
</tr>
<tr>
<td>$n$</td>
<td>the order of integral distinguisher ($1 \leq n \leq N - 1$).</td>
</tr>
<tr>
<td>$RK^\gamma$</td>
<td>round key used in $r$-th round ($RK^r \in F_{2^{mn}}^N, 1 \leq r \leq R$).</td>
</tr>
<tr>
<td>$Enc^\gamma$</td>
<td>$\gamma$ rounds cipher function, s.t., $F_{2^{mn}}^N \times F_{2^{mn}}^R \rightarrow F_{2^{mn}}^N (1 \leq \gamma \leq R)$.</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>input condition; index set of variable subblocks. When an index of variable subblocks is ${a_0, a_1, ..., a_{n-1}}$, it is written as $\alpha_{a_0a_1...a_{n-1}} (1 \leq n \leq N - 1)$.</td>
</tr>
<tr>
<td>$\beta$</td>
<td>output property; index set of balanced subblocks. When an index of variable subblocks is ${b_0, b_1, ..., b_{n-1}}$, it is written as $\beta_{b_0b_1...b_{n-1}} (1 \leq k \leq N)$.</td>
</tr>
<tr>
<td>$S$</td>
<td>set of chosen plaintexts which is defined by $\alpha$. ($S = {S_i</td>
</tr>
<tr>
<td>$f \cdot S'$</td>
<td>integrated value of a set of outputs $S'$ which is obtained by $\gamma$ rounds encryption of $S$, and balanced subblocks are denoted by $\beta$ ($f \cdot S' = \bigoplus_{i=0}^{2^{mn}-1} Enc^\gamma(S_i, RK^{1,2,3,...,\gamma})$).</td>
</tr>
<tr>
<td>$\alpha \rightarrow^\gamma \beta$</td>
<td>integral distinguisher; integrated value has output property $\beta$ when a set of chosen plaintexts defined by input condition $\alpha$ is encrypted for $\gamma$ rounds.</td>
</tr>
<tr>
<td>$S_{\alpha}$</td>
<td>$b$-th subblock of $\gamma$-th round ($0 \leq b \leq N - 1, 0 \leq r \leq R$).</td>
</tr>
<tr>
<td>$x_{\alpha}$</td>
<td>sequence of values of $S_{\alpha} (x_{\alpha} = (x_{\alpha})<em>{l=0}^{2^{mn}-1}, x</em>{\alpha} \in F_{2^{mn}}^N)$.</td>
</tr>
<tr>
<td>$X_{\alpha}$</td>
<td>first order integral property of $x_{\alpha}$.</td>
</tr>
<tr>
<td>$X_{\alpha}$</td>
<td>higher order integral property of $x_{\alpha}$.</td>
</tr>
<tr>
<td>$G$</td>
<td>partial function of subblock-based block cipher, s.t., $F_{2^{mn}}^N \rightarrow F_{2^{mn}}^N$.</td>
</tr>
<tr>
<td>$H$</td>
<td>partial function of subblock-based block cipher, s.t., $F_{2^{mn}}^N \rightarrow F_{2^{mn}}^N$.</td>
</tr>
<tr>
<td>$V$</td>
<td>set of input variable subblocks.</td>
</tr>
<tr>
<td>$V'$</td>
<td>set of intermediate variable subblocks.</td>
</tr>
<tr>
<td>$P$</td>
<td>set of variable subblocks in bijective path.</td>
</tr>
<tr>
<td>$E$</td>
<td>set of variable subblocks in endpoint of bijective path.</td>
</tr>
</tbody>
</table>

Where $\oplus$ denotes XOR (XClusive OR) summation. If there exist at least one subblock which is always 0 in $S'$ for any values of constant subblocks and round keys, we can define integral distinguisher. We call such subblocks as balanced subblocks, and we define output property using them. Let _ be output property which denotes index set of balanced subblocks.
We denote integral distinguisher by input condition \( \alpha \) and output property \( \beta \). We define integral distinguisher by \( \alpha \rightarrow \gamma \beta \). It denotes that an integrated value has output property \( \beta \) when a set of chosen plaintexts defined by input condition \( \alpha \) is encrypted for \( r \) rounds. Due to the limited space, we omit the integral attack scenario using integral distinguisher, and typical case is shown in [11].

### 2.2. Conventional Algorithm

To obtain higher order integral distinguisher, the conventional algorithm execute following two steps. In the first step, first order integral distinguisher is obtained. In first order integral, only one subblock is chosen as variable and input condition is written as \( \alpha_{[a]} \) \((0 \leq a \leq N - 1)\). In the second step, first order integral distinguisher is extended to higher order integral distinguisher. In higher order integral, multiple subblocks are variables in input condition \( \alpha \) and it is denoted as \( \alpha_{[a_0,a_1,...,a_n]} \) \((n \leq n \leq N - 1)\). We show the first step in Sec.4.7 and show the outline of second step in the following.

A typical example is the evaluation of CLEFIA [6]. The designers of CLEFIA extend first order integral distinguisher to second one. Fig.1 shows the outline. Suppose that following first order integral distinguisher is obtained in the first step.

\[
\alpha_{[1]} \rightarrow \beta_{[1]} \tag{3}
\]

Then we regard \( \alpha_{[1]} \) as input condition of \( r \)-th round. The input condition \( \alpha_{[1]} \) denotes a subblock \( x_i \) is variable, and it takes every possible element of \( \mathbb{P}_2^{32} \). Considering an input condition which can be extended to one of \((r - 1)\)-th round, we assign \( x_i \) as a variable subblock to make \( x_i \) take every possible element of \( \mathbb{P}_2^{64} \). A mapping from \([x_i^1, x_i^2]\) to \([x_i^{2^{-1}}, x_i^{3^{-1}}]\) is bijective when all secret keys are constant. Therefore, \( x_i^{2^{-1}} \parallel x_i^{3^{-1}} \) takes every possible element of \( \mathbb{P}_2^{64} \), and input condition to denote \((r - 1)\)-th round is \( \alpha_{[2,3]} \). Inverting input-output relation, \( \alpha_{[2,3]} \) of \((r - 1)\)-th round results in \( \alpha_{[1]} \) of \( r \)-th round. Then, the input condition \( \alpha_{[1]} \) is considered as \( 2^{32} \) sets of plaintexts which all accord to \( \alpha_{[1]} \) from Eq.(1). Therefore, output property \( \beta_{[1]} \) also holds in one caused by \( \alpha_{[1]} \), since integration of outputs from \( \alpha_{[1]} \) is calculated as integration of \( 2^{32} \) integrated values denoted by \( \beta_{[1]} \). For these reasons, integral distinguisher of \( \alpha_{[1]} \) can be extended to integral distinguisher of \( \alpha_{[2,3]} \) as follows.

\[
\begin{align*}
\alpha_{[2,3]} \rightarrow \beta_{[1]} & \Rightarrow \alpha_{[2,3]} \rightarrow \beta_{[1]} \\
\alpha_{[1]} \rightarrow \beta_{[1]} & \Rightarrow \alpha_{[2,3]} \rightarrow \beta_{[1]} \tag{4}
\end{align*}
\]

In the same way, integral distinguisher of \( \alpha_{[2,3]} \) can be extended to integral distinguisher of \( \alpha_{[0,1,3]} \). Then we have following.

\[
\begin{align*}
\alpha_{[0,1,3]} \rightarrow \beta_{[1]} & \Rightarrow \alpha_{[0,1,3]} \rightarrow \beta_{[1]} \\
\alpha_{[2,3]} \rightarrow \beta_{[1]} & \Rightarrow \alpha_{[0,1,3]} \rightarrow \beta_{[1]} \tag{5}
\end{align*}
\]
the second step of the conventional algorithm, we extend first order integral such as Eq.(3) round by round, exploiting bijective components. Conventionally, the order is increased to \( N - 1 \), since \((N - 1)\)-th order integral distinguisher is the upper bound of integral distinguisher.

3. PROBLEMS OF CONVENTIONAL ALGORITHM

In Feistel structures whose number of subblocks \( N \) is small such as CLEFIA \((N = 4)\), the conventional algorithm is effective. However, we find it is problematic to apply the algorithm for Feistel structures whose \( N \) is large such as TWINE and LBlock \((N = 16)\).

We found new integral distinguisher of TWINE and LBlock which can not be searched by the conventional algorithm in our previous work [14]. We executed computer experiment to search for such integral distinguisher. In particular, the first step of the conventional algorithm was replaced with the search by computer experiment. We used the second step to extend the result of the computer experiment. Note that computer experiment is to compute an integration of outputs of multiple rounds encryption such as Eq.(2). We checked positions of balanced subblocks in integrated values. However, unbalanced subblocks in integrated values become 0 with probability \( 2^{-m} \). To distinguish such unbalanced subblocks with balanced ones, we executed computer experiment for 10 times by choosing random values for constant subblocks and round keys.

We compare the results with the ones of the conventional algorithm in the following. Due to the limitations of space, see the details of each algorithms in [13] and [8].

TWINE [13]: We found following first order integral distinguisher of TWINE by computer experiment.

\[
\alpha_{[1]} \rightarrow^0 \beta_{[1,3,13,15]} \tag{6}
\]

This integral distinguisher can be extended to fifteenth order integral distinguisher as follows.

\[
\alpha_{[14]} \rightarrow^{15} \beta_{[1,3,13,15]} \tag{7}
\]

where \( \alpha_{[14]} \) denotes all subblocks other than fourteenth subblock are variable. On the other hand, following eleventh order integral distinguisher of TWINE was obtained by our computer experiment [14].

\[
\alpha_{[1,4,11,13,14]} \rightarrow^{11} \beta_{[1,3,5,7,9,11,13,15]} \tag{8}
\]

This integral distinguisher can be extended to fifteenth order integral distinguisher as follows.

\[
\alpha_{[14]} \rightarrow^{15} \beta_{[1,3,5,7,9,11,13,15]} \tag{9}
\]

The number of balanced subblocks is increased in Eq.(9) compared with Eq.(7), though they have the same input conditions.

LBlock [8]: In LBlock, following fifteenth order integral distinguisher obtained by the conventional algorithm is known [8][10].

\[
\left\{ \begin{array}{l}
\alpha_{[5]} \rightarrow^{9} \beta_{[9,1,1,13,15]} \\
\alpha_{[8]} \rightarrow^{15} \beta_{[9,1,13,15]}
\end{array} \right. \tag{10}
\]
We found following fifteenth order integral distinguisher by computer experiment and its extension [14].

\[
\left\{ \begin{array}{c}
\alpha_{(1,7,9,11,14)}^{(5)} \rightarrow^{11} \beta_{(8-15)}^{} \Rightarrow \\
\alpha_{(1,7,9,11,14)}^{(5)} \rightarrow^{15} \beta_{(8-15)}^{}
\end{array} \right. \tag{11}
\]

As same as TWINE, the number of balanced subblocks is increased in Eq.(11) compared with Eq.(10).

In fifteenth order integral distinguisher of TWINE and LBlock such as Eq.(9) and (11), the number of balanced subblocks are increased from 4 to 8 (32-bit). These results suggest a practical problem of the conventional algorithm.

Also, there is a theoretical problem in the conventional algorithm. Input conditions in the scope of the algorithm are limited to first order integral distinguisher and their extensions. The number of input conditions in the scope is calculated as \(2^{N-2} \) Subtraction denotes a full code book and single chosen plaintext which are out of scope of integral attack.

In CLEFIA (\(N = 4\)), the number of possible input conditions is 14, however, one of the conventional algorithm is 8. As for TWINE (\(N = 16\)), the number of possible input conditions is \(2^{16} - 2\), however, one of the conventional algorithm is 56. In this way, difference between the number of possible input conditions and one of the conventional algorithm becomes huge when \(N\) is large.

4. FUNDAMENTALS OF PROPOSAL SEARCH ALGORITHM

4.1. Subblock-based Block Cipher

In this paper, we limit the scope of target to subblock-based block cipher. To define such cipher, we suppose a block cipher which composes of bijective functions and addition on \(F_{2^m}^m\). With respect to bijective functions, we denote \(L\) and \(F\) be linear and nonlinear unctions, s.t., \(F_{2^m}^m \rightarrow F_{2^m}^m\)

For example, we consider a nonlinear function \(F\) which contains key addition as follows.

\[
F(x; RK) = S(x \oplus RK),
\]

where \(S\) is a S-box of \(F_2^{2^m} \rightarrow F_2^{2^m}\). Suppose \(RK\) is a random constant, \(F\) is regarded as a random function chosen by \(RK\), s.t., \(F_2^{2^m} \rightarrow F_2^{2^m}\). Although there are \(2^m\) possible random functions, we ignore the detail of functions and denote them by \(F\), since our interest is only in their bijection. We denote \(\oplus\) and \(\oplus\) as XOR and addition mod \(2^m\), respectively.

We suppose \(G\) and \(H\) are composite functions of \(L, F\) and above two types of additions.

4.2. Sequence of SubBlocks

We analyze higher order integral distinguisher in subblock unit. For analyzing subblocks, their values were analyzed as multiset in the previous work [2]. However, sequence allows us to
analyze more detailed properties, since it represents detailed states. We denote a sequence of a subblock \( \lambda \) as follows.

\[
\lambda = (\lambda_i)_{i=0}^{2m-1}, \quad \lambda_i \in \mathbb{F}_2^m
\]

Also, a sequence of a concatenation of \( l \) subblocks \( \_ \) is written as

\[
\widetilde{\lambda} = (\lambda_i)_{i=0}^{2m-1} = (\lambda_0||\lambda_1||...||\lambda_{l-1,1})_{i=0}^{2m-1},
\]

where \( \lambda = \lambda_0||\lambda_1||...||\lambda_{l-1} \) (\( 1 \leq l \leq N \)).

The sequential order to input chosen plaintexts does not affect output property from Eq.(2). Therefore, we can define a mapping to sort a sequence of \( \lambda \), s.t.,

\[
\Xi : (\lambda_i)_{i=0}^{2m-1} \rightarrow (\lambda_{\sigma(i)})_{i=0}^{2m-1},
\]

where \( \sigma \) is any given permutation function.

\[\text{Figure 2: Trail of endpoint of bijective path \( E \) in CLEFIA.}\]

### 4.3. First order Integral Property

First order integral distinguisher is easily obtained by symbolization of sequences of subblock values. Symbols to denote the property of each sequence of a subblock is called as integral property. Especially, integral property of first order integral is called as first order integral property. We introduce a definition used in [6].

**Definition 1 (First order Integral Property)** Let \( x_{b,i} = (x_{b,i})_{i=0}^{2m-1} (x_{b,i} \in \mathbb{F}_2^m) \) be a sequence of subblock \( x_{b,i} \) and \( X_{b,i} \) be a first order integral property of \( x_{b,i} \). First order integral property \( X_{b,i} \) is categorized as follows.

- **Constant(C):** \( \forall i, i', x_{b,i} = x_{b,i'} \)
- **All(A):** \( \forall i, i' (i \neq i'), x_{b,i} \neq x_{b,i'} \)
- **Balance(B):** \( \sum_{i=0}^{2m-1} x_{b,i} = 0 \)
- **Random(R):** Others

(17) \( \text{(17)} \)

(18) \( \text{(18)} \)

(19) \( \text{(19)} \)

(20) \( \text{(20)} \)
4.4. Higher order Integral Property

Integral property whose order \( n \) equals to or greater than 2 is called as higher order integral property. To analyze higher order integral property, we define new symbolization as follows.

**Definition 2 (Higher order Integral Property)** Let \( \vec{x}_b^e = (x_{b,j})_{j=0}^{2^m-1} \) be a sequence of a subblock \( x_{b,j} \) and \( y_b^{(k)}(k) \) multiplicity of \( x_{b,j} = k \) in the sequence. We denote \( N_b^e \) as higher order integral property of \( \vec{x}_b^e \).

Higher order integral property \( N_b^e \) is categorized as follows:

- **Constant** \( C \): \( y_i, y'_i, x_{b,i} = x_{b,i} \)
- **Uniform** \( U \): \( y_k, x_{b,j} = 2^{2m(n-1)} \)
- **Event** \( E \): \( y_k, x_{b,j} \mod 2 = 0 \)
- **Balance** \( B \): \( \sum_{j=0}^{2^m-1} x_{b,j} = 0 \)
- **Random** \( R \): Others

A sequence \( \vec{x}_b^e \) may have multiplicity, since the length of \( \vec{x}_b^e \) is \( 2^m \) and \( x_{b,j} \in \mathbb{F}_2^m \). Therefore, we define \( U \) instead of \( A \) of first order integral, and \( U \) is different property from \( B \). Also, we define special property \( E \) which is the intermediate property between \( U \) and \( B \). If an input sequence is \( U \) or \( E \), an output sequence of bijective function \( F \) has the same property as input.

From definition of higher order integral properties, we have following inclusion relation among these properties.

\[ \mathcal{B} \supset \mathcal{E} \supset \mathcal{U} \]  

There is a case that analysis of one subblock results in multiple properties. If it is analyzed as \( \mathcal{B} \) and \( \mathcal{E} \), higher order integral property of the subblock is \( \mathcal{E} \), since it does not contradict the analysis of \( \mathcal{B} \). By using such inclusion relation, higher order integral property of the subblock is \( \mathcal{U} \) when it is analyzed as \( \mathcal{B}, \mathcal{E} \) and \( \mathcal{U} \).
4.5. Search Algorithm Using Bijection

Integral distinguisher mainly depends on bijection between input variable subblocks and output variable subblocks of intermediate rounds. Let $V$ be a set of input variable subblocks. In $n$-th order integral, the number of subblocks in $V$ is $n$. Considering a set of intermediate variable subblocks $V'$, there are some sets to satisfy that $V \rightarrow V'$ is bijective. To analyze bijective characteristics, we define bijective path $P$ which is a set of such subblocks as follows.

**Definition 3 (Bijection Path)** Let $V$ be a set of input variable subblocks ($|V| = n$). We denote $G$ as a function of $\mathbb{F}_2^m \rightarrow \mathbb{F}_2^m$ supposed in Eq.(13). We define bijective path $P$ and endpoint of bijective path $E$ by following recursive conditions. Note that $P$ and $E$ are sets of subblocks.

(i) A set of input variable subblocks $V$ is substituted for $P$ and $E$.

(ii) Consider a set of intermediate subblocks $E'$, s.t., $E' = G(E)$. If $E'$ satisfies that $E \rightarrow E'$ is bijective, substitute $P = P \cup E'$ and $E = E'$. Note that any subblocks which are not in $E'$ must not be codomain of any mapping from $E$.

(iii) If there is not any $E'$ which satisfies $E \rightarrow E'$ is bijective in condition (ii), $E$ is the set of subblocks in endpoint of bijective path.

We use recursive conditions of Definition 3 as procedure in Algorithm A (see Sec.5.11; search for bijective path). We update subblocks in endpoint of bijective path $E$, and bijective path $P$ is a set of all subblocks which are temporarily or finally in $E$. We demonstrate the search in CLEFIA as a toy example. We show the trail of endpoint of bijective path $E$ when we set input condition $\_f0;1;3g$ in Fig.2. Filled squares indicate subblocks in $E$. From left to right, subblocks in $E$ are updated, and the rightmost one is finally obtained as the endpoint of bijective path $E$. We use a function $\phi$ as follows.

$$
\{x'_0, x'_1\} = \phi(x_0, x_1) = \{x_0, F(x_0; RK) \oplus x_1\},
$$

$$
x_0, x_1, RK \in \mathbb{F}_2^3, \tag{27}
$$

where $RK$ is a constant key value. The function $\phi$ has following characteristics.

(i) If $x_0, x_1 \in E$ holds, $\phi$ is bijective of $\mathbb{F}_2^m \rightarrow \mathbb{F}_2^m$.

(ii) If $x_1 \in E$ and $x_0 \notin E$ hold, $\phi$ is bijective of $\mathbb{F}_2^m \rightarrow \mathbb{F}_2^m$.

(iii) If $x_0 \in E$ and $x_1 \notin E$ hold, $\phi$ is not bijective.

We summarize the above characteristics in Fig.3.

Using bijective path and endpoint of bijective path, we can estimate higher order integral properties of subblocks are $\mathcal{U}$ or $E$ from following propositions.

**Proposition 1** If a subblock $x'_b$ is in bijective path $P$ ($x'_b \in P$), higher order integral property of $x'_b$ is $\mathcal{U}$ ($X'_b = \mathcal{U}$).

**Proof** Suppose that $V \rightarrow V'$ is bijective. From Definition 3, at least one set of $n$ variable subblocks $V'$ exists in $P$. From condition of chosen plaintexts (Eq.(1)), $\bar{V}$ contains every possible element of $\mathbb{F}_2^m$ without multiplicity. Therefore, all sequences in $\bar{V}$ satisfy the definition of $\mathcal{U}$ (Eq.(22)). Since $V \rightarrow V'$ is bijective, $\bar{V}'$ also contains every possible element of $\mathbb{F}_2^m$ without multiplicity. Therefore, all sequences in $\bar{V}'$ satisfy the definition of $\mathcal{U}$. Hence, higher order integral properties of all subblocks in $P$ are $\mathcal{U}$. □
Proposition 2 Let \( V \) be a set of input variable subblocks, and \( E = \{e_0, e_1, \ldots, e_{n-1}\} \) endpoint of bijective path. Let \( \{X'_i(e_0), X'_i(e_1), \ldots, X'_i(e_{n-1})\} \) be a set of first order integral properties of subsequences, s.t., \( E \backslash e_t \) \((0 \leq t \leq n-1)\) are constant. If following condition holds, higher order integral property \( X'_b \) is \( \mathcal{U} \).

\[ \exists e_t \in E, X'_b(e_t) = A \Rightarrow X'_b = \mathcal{U} \]  

(28)

Proof Let \( H \) be a function of \( \mathbb{F}_2^{m} \rightarrow \mathbb{F}_2^{m} \) supposed in Eq.(13). Suppose a subblock value is calculated as

\[ x'_b = H(e_0, e_1, \ldots, e_{n-1}). \]  

(29)

A sequence \( \tilde{x}'_b \) is written as

\[ \tilde{x}'_b = (H(e_0, e_1, \ldots, e_{n-1}))_{0}^{2^{m-1}}, \]

(30)

where \( e_t \) is \( t \)-th element of \( \tilde{e}_t \) and \( E_t \) is a set of \( t \)-th elements of sequences of subblocks in \( \tilde{E} = \{\tilde{e}_0, \tilde{e}_1, \ldots, \tilde{e}_{n-1}\} \). By substituting \( t = 2^m k + l (0 \leq k \leq 2^{m(n-1)} - 1, 0 \leq l \leq 2^m - 1) \), we can rewrite Eq.(30) as

\[ x'_b = (H(e_{2^m k+l}, E_{2^m k+l} \backslash e_{2^m k+l}))_{0}^{2^{m(n-1)}}. \]  

(31)

From an assumption that \( V \rightarrow E \) is bijective, the number of distinct combinations of \( \tilde{E} \) is \( 2^{mn} \) and one of \( \tilde{E} \) is \( 2^{mn} \). The multiplicity of each combination in \( \tilde{E} \) is \( 2^{m} \), and \( \tilde{e}_t \) takes \( 2^{m} \) distinct values for every possible combination of \( \tilde{E} \) in \( \tilde{e}_t \). Therefore, it is possible to sort \( x'_b \) which satisfies following two conditions in Eq.(31).

(i) \( \forall k, l, l' (l \neq l'), e_{l,2^m k+l} \neq e_{l',2^m k+l} \)

(ii) \( \forall k, l, l' (l \neq l'), E_{2^m k+l} \backslash e_{l,2^m k+l} = E_{2^m k+l} \backslash e_{l',2^m k+l} \)

As an example, we show a sorted sequence, s.t., each element is in \( \mathbb{F}_2^m \) and \( n = 15 \), in Fig.4.

Let \( x''_b \) be a subsequence of sorted \( x'_b \) for given \( k \), and \( x''_b \) is written as

\[ x''_b = (H(e_{2^m k+l}, V''_{2^m k+l} \backslash e_{2^m k+l}))_{0}^{2^{m-1}}. \]  

(32)

We can regard integral property of \( x''_b \) as first order, and \( X''_b(e_t) \) holds for any values of constant sequences from the definition of integral distinguisher shown in Sec.2.3. Suppose a first order integral property of \( X''_b \) is \( A (X''_b(e_t) = A) \), \( x''_b \) contains every possible element of \( \mathbb{F}_2^m \) for any \( k \) in Eq.(32). In addition, \( x''_b \) is a

![Figure 4: Example of sorted sequence.](image)
concatenation of all subsequences $x''_B$ which are respectively chosen by $k$ ($0 \leq k \leq 2^{m(n-1)} - 1$). Therefore, $x''_B$ contains $2^{m(n-1)}$ subsequence which contain every possible element of $\mathbb{F}^m_2$, and this satisfies the definition of $\mathcal{U}$ (Eq.(22)). □

**Proposition 3** Let $V$ be a set of input variable subblocks and $E = \{e_0, e_1, ..., e_{n-1}\}$ endpoint of bijective path. Let $\{X''_B(e_0), X''_B(e_1), ..., X''_B(e_{n-1})\}$ be a set of first order integral properties of subblocks, s.t., $E \setminus e_t \ (0 \leq t \leq n - 1)$ are constant. If following condition holds, higher order integral property $X''_B$ is at least $E$.

\[ \exists e_t \in E, \ X''_B(e_t) = C \Rightarrow X''_B = E \]  \hspace{1cm} (33)

**Proof** Since $X''_B(e_t) = C$, there is not any mapping of $e_t \rightarrow x''_B$. By substituting $i = 2^{m(n-1)}k+l \ (0 \leq k \leq 2^{m-1}, 0 \leq l \leq 2^{m(n-1)}-1)$ in Eq.(30), we can rewrite Eq.(30) as follows.

\[ x''_B = \left( (H(E_{2^{m(n-1)}l}\setminus e_t, 2^{m(n-1)}l, k, l)) \right)_{k=0}^{2^{m-1}-1} \]  \hspace{1cm} (34)

where $E_t$ is a set of $i$-th elements of sequences of subblocks in $E = \{\tilde{e}_0, \tilde{e}_1, ..., \tilde{e}_{n-1}\}$, and $H$ be a function of $\mathbb{F}^m_2 \rightarrow \mathbb{F}^m_2$ supposed in Eq.(13). Since $V \rightarrow E$ is bijective, $E \setminus e_t$ contains $2^{m(n-1)}$ distinct combinations, and multiplicity of $E \setminus e_t$ is $2^m$. Therefore, we can sort $x''_B$ as $E_{2^{m(n-1)}l}\setminus e_t, 2^{m(n-1)}l, k, l$ whose multiplicity of each element of $\mathbb{F}^m_2$ is the same for any $k$ in Eq.(34). Therefore, $x''_B$ is regarded as a concatenation of $2^m$ subsequences which has the same multiplicity for any element of $\mathbb{F}^m_2$. Hence, multiplicity of each element of $\mathbb{F}^m_2$ in the sequence $x''_B$ is the product of $2^m$ and this multiplicity satisfies the definition of $E$ (Eq.(23)). □

Using Proposition 1, 2 and 3, we search for subblocks of $\mathcal{U}$ and $E$. Also, subblocks of $C$ are easily searched by considering the relation with input variable subblocks in $V$ as follows.

\[ \forall v_u \in V, \ X''_B(v_u) = C \Rightarrow X''_B = C \]  \hspace{1cm} (35)

Then, we search for subblocks of $B$ in the same manner of first order integral property. Higher order integral properties of all subblocks which are not analyzed as above properties are $R$. Note that subblocks of $B$ and $E$ may be analyzed as $U$ by using injection (see Proposition 4), and $R$ can be updated to $B$ from such analysis.

4.6. Search Algorithm Using Injection

We exploit similar property of Proposition 2 from injection among subblocks to update higher order integral properties obtained by bijective path. We can obtain any integral properties only by an algorithm using injection, since property obtained by bijection is also obtained by injection.

However, the algorithm using bijection is more efficient than one of injection. Therefore, we execute the former at first, and execute the latter as finalization.

To analyze injection, we use new idea, independent. This idea resembles to linear independent. We define independent as follows.

**Definition 4 (Independent)** Let $H_i \ (0 \leq i \leq n'-1, 1 \leq n')$ be nonlinear functions of $\mathbb{F}^m_2 \rightarrow \mathbb{F}^m_2$ which are supposed in Eq.(13). We denote $E = \{e_0, e_1, ..., e_{n-1}\}$ as endpoint of bijective path. Suppose a set of $n'$ intermediate variable subblocks $V' = \{v_0, v_1, ..., v_{n'-1}\}$ are calculated as
Using the definition of independent, we have a following proposition.

**Proposition 4** Let $H$ be a function of $\mathbb{F}_2^{2n'} \to \mathbb{F}_2^n$ supposed in Eq.(13). Let $V' = \{v'_0, v'_1, ..., v'_{n'-1}\}$ be a set of intermediate variable subblocks, and subblock $x'_k$ and $V'$ satisfy

$$x'_k = H(v'_0, v'_1, ..., v'_{n'-1}).$$  \hfill (37)

Suppose $v'_t \in V'$ is independent of $V'$ (see Definition 4). If following condition holds, higher order integral property $X'_k$ is $\mathcal{U}$.

$$X'_k(v'_t) = A \Rightarrow X'_k = \mathcal{U}$$  \hfill (38)

**Proof** A sequence $x'_k$ can be written as

$$x'_k = \left( H(v'_i, V'_i \setminus v'_t) \right)_{i=0}^{2^{n'-1}},$$  \hfill (39)

where $v'_i$ is $i$-th element of $v'_t$ and $V'_i$ is a set of $i$-th elements of sequences of subblocks in $\hat{V}' = \{v'_0, v'_1, ..., v'_{n'-1}\}$. By substituting $i = 2^mk + l$ ($0 \leq k \leq 2^{m(n-1)} - 1, 0 \leq l \leq 2^{m} - 1$), we can rewrite Eq.(39) as

$$x'_k = \left( H(v'_{2^mk+l}, V'_{2^mk+l} \setminus v'_{2^mk+l}) \right)_{l=0}^{2^{m(n-1)}-1}.$$  \hfill (40)

From condition (ii) of Definition 4, $V' \setminus v'_t$ contains $2^{m(n-1)}$ distinct combinations, and multiplicity of each distinct combinations is $2^m$. From condition (i) of Definition 4, $V'$ contains $2^n$ distinct combinations, and there is no multiplicity. Therefore, $v'_t$ takes $2^m$ distinct values for every possible combination of $V' \setminus v'_t$. We sort $x'_k$ as following two conditions in Eq.(40).

(i) $\forall k, l, l' \in \mathbb{Z}$, $v'_{2^mk+l} \neq v'_{2^mk+l'}$

(ii) $\forall k, l, l' \in \mathbb{Z}$, $V'_{2^mk+l} \setminus v'_{2^mk+l} = V'_{2^mk+l'} \setminus v'_{2^mk+l'}$

Let $\tilde{x}'_k$ be a subsequence of sorted sequence $x'_k$ for given $k$, and $\tilde{x}'_k$ is written as

$$\tilde{x}'_k = \left( H(v'_{2^mk+k}, V'_{2^mk+k} \setminus v'_{2^mk+k}) \right)_{k=0}^{2^{m-1}}.$$  \hfill (41)

Therefore, we can regard integral property of subsequence $\tilde{x}'_k$ as first order, and $X'_k(v'_t)$ holds for any values of constant sequences from the definition of integral distinguisher. Suppose a first order integral property of $\tilde{x}'_k$ is $A$ ($X'_k(v'_t) = A$), $\tilde{x}'_k$ contains every possible element of $\mathbb{F}_2^n$ for any $k$ in Eq.(41). In addition, $\tilde{x}'_k$ is a concatenation of subsequences $\tilde{x}'_k$ which are respectively chosen by $k$. Therefore, $\tilde{x}'_k$ contains $2^{m(n-1)}$ subsequences which contains every possible element of $\mathbb{F}_2^n$, and this satisfies the definition of $\mathcal{U}$ (Eq.(22)).

Using Proposition 4, we can update higher order integral properties of subblocks which are determined as $\mathcal{B}$ or $\mathcal{E}$ in the algorithm shown in Sec.4.9. Then, we search for subblocks of $\mathcal{R}$ updated to $\mathcal{B}$ in the same manner of first order integral property.
5. **PROCEDURE OF PROPOSAL ALGORITHM**

We divide search algorithm into two algorithms, Algorithm A and B. They are based on bijective path (Sec.4.9) and independent (Sec.4.10), respectively. In addition, we propose Algorithm B′ to check injection in Algorithm B.

5.1. **Algorithm A**

In Algorithm A, higher order integral properties $X'_{b_r}$ for all subblocks $x_b^r$ ($0 \leq r \leq R$, $0 \leq b \leq N - 1$) are estimated by setting input condition $\alpha_{[\alpha_0, \alpha_1, \ldots, \alpha_{b-1}]}$. Note that $X'_{b_r}$ is an element of $\{C, U, E, B, R\}$ from Definition 2. Using bijective path, we can estimate higher order integral distinguisher. Algorithm A has following five steps and the flowchart is shown in Fig.5.

- **Step-1 (Initialization)** Input $\alpha_{[\alpha_0, \alpha_1, \ldots, \alpha_{b-1}]}$, and set $V = \{x_0^0, x_1^0, \ldots, x_{a_1-1}^0\}$. A value of $r$ is initialized to 0, and substitute $E^r = V$. Partition a round function $G$ into $M$ partial functions $\{G_0, G_1, \ldots, G_{M-1}\}$. Go to Step-2.

- **Step-2 (Bijective Path)** Substitute $P = P \cup E^r$. Partition $E^r$ into $\{E'_0, E'_1, \ldots, E'_{M-1}\}$ based on input relation for $\{G_0, G_1, \ldots, G_{M-1}\}$. Considering $E'_i$, s.t., $E'_i = G_i(E'_i)$ ($E'_i \subseteq E'$) for $0 \leq i \leq M - 1$, if $E'_i \rightarrow E'_j$ is bijective, then substitute $E^{r+1} = E^{r+1} \cup E'_j$, else substitute $E = E \cup E'_j$. If $|E^{r+1}| = 0$ or $r > R$ hold, then go to Step-3, else substitute $r = r + 1$ and repeat Step-2.

- **Step-3 (Inside Bijective Path)** For $x'_b^r \in P$, substitute $X'_{b_r} = U$ (Proposition 1). Go to Step-4.

- **Step-4 (Outside Bijective Path)** For $x'_b^r \notin P$, calculate first order integral properties $X'_{b_r}(v_o)$ and $X'_{b_r}(e_t)$ for $v_o \in V$ and $e_t \in E$. From the results, following four cases are applied for $x'_b^r \notin P$.
  - (i) If $v_o$, s.t., $X'_b(v_o) = C$ holds, then substitute $X'_{b_r} = C$ (Eq.(35)).
  - (ii) If $e_t$, s.t., $X'_b(e_t) = A$ holds, then substitute $X'_{b_r} = U$ (Proposition 2).
  - (iii) If $e_t$, s.t., $X'_b(e_t) = C$ and $X'_b \notin \{C, U\}$ hold, then substitute $X'_{b_r} = E$ (Proposition 3).
  - (iv) Other than the above, substitute $T = T \cup x'_b$. Go to Step-5.

- **Step-5 (Other Properties)** For $x'_b^r \in T$, if $x'_b^r = x \oplus x'$ or $x \oplus x'$, s.t., $X, X' \in \{U, E, S\}$, then substitute $X'_{b_r} = B$ and subtract $X'_{b_r}$ from $T (T = T \setminus x'_b)$. For all $x'_b \in T$, substitute $X'_{b_r} = R$. Output $E$ and $X'_{b_r}$ of $x'_b$. End algorithm.
In Step-1, we initialize variables. Based on input condition, we construct a set of input variable subblocks $V$. Let $E'$ be a set of subblocks in $r$-th round which can be ones in endpoint of bijective path $E$, and $r$ is initialized to 0. We substitute $E' = V$. To analyze a round function in partial function manner such as function $\phi$ (see Fig.3), $G$ is partitioned into $M$ partial functions $\{G_0, G_1, ..., G_{M-1}\}$.

In Step-2, we search for subblocks in bijective path $P$ and endpoint of bijective path $E$. Step-2 is repeated with increasing the value of $r$. We substitute $P = P \cup E'$ and partition $E'$ into $\{E'_0, E'_1, ..., E'_{M-1}\}$ following the manner of input for partial functions $\{G_0, G_1, ..., G_{M-1}\}$. We consider $E'_i = G_i(E'_i)$ for $0 \leq i \leq M - 1$ ($E'_i \subset E$). If $E'_i \rightarrow E'_i$ is bijective, we substitute $E^{r+1} = E^{r+1} \cup E'_i$. Otherwise we substitute $E = E \cup E'_i$ following condition (ii) of Definition 3. Following condition (iii) of Definition 3, if $E^{r+1}$ is empty set or $r$ is more than the number of full rounds $R$, $E$ and $P$ are determined and we go to Step-3. Otherwise, we repeat Step-2 with increasing the value of $r$.

We divide search scope into two domains, inside and outside of bijective path $P$. In Step-3 for inside $P$, higher order integral properties of subblocks are determined as $U$ from Proposition 1. In Step-4 for outside $P$, we calculate first order integral properties of each subblocks from subblocks in $V$ and $E$. Using Eq.(35), Proposition 2 and 3, higher order integral properties are estimated as $C$, $U$ or $E$. There are subblocks whose higher order integral properties are not estimated. Let $T$ be a set of such subblocks.

In Step-5, we estimate higher order integral properties of subblocks in $T$. They can be $B$ or $R$ in the same manner of first order integral.

In Algorithm A, subblocks of $B$, $E$ and $R$ are estimated. Their higher order integral properties are finalized in Algorithm B.

### 5.2. Algorithm B

In Algorithm B, higher order integral properties obtained in Algorithm A are inputted, and they are finalized as output property. Using independent subblocks, some subblocks of $E$ and $B$ can be updated to $U$ from Proposition 4. Algorithm B has following six steps and the flowchart is shown in Fig.6.
• **Step-1 (Initialization)** Input $E$ and $X'_b$ of $Yx'_b$ obtained in Algorithm A. Let $T$ be a set of $Yx'_b$, s.t., $X'_b \in \{B,E\}$. Go to Step-2.

• **Step-2 (Extraction from Temporary Subset Block)**
  If $|T| = 0$ holds, then go to Step-6, else extract a subblock $x'_b$ from $T$ and initialize $r_0 = r - 1$. Go to Step-3.

• **Step-3 (Independent Condition (i))** If $r_0 < 0$ holds, then subtract $x'_b$ from $T$ ($T = T \setminus x'_b$) and return to Step-2. Let $V'$ be a set of subblocks in $r_0$-th round, s.t., $x'_b = H(V')$. If $E \rightarrow V'$ is injective, then go to Step-4, else substitute $r_0 = r_0 - 1$ and repeat Step-3.

• **Step-4 (Independent Condition (ii))** For all combinations $\{e', e\}$ ($e' \in V', e \in E$), test injection of $E \setminus e_0 \rightarrow V' \setminus e_0'$ and calculate $X'_b(e_0')$. If $E \setminus e_0 \rightarrow V' \setminus e_0'$ is injective and $X'_b(e_0') = A$ holds, then substitute $X'_b = \mathcal{U}$ and go to Step-5, else return to Step-3 (Proposition 4).

• **Step-5 (Update Properties)** Consider $Yx'_b$, s.t., $X'_b = \mathcal{R}$. If $x'_b = x \oplus x'$ or $x \boxplus x'$, s.t., $X, X' \in \{\mathcal{U}, E, B\}$, then substitute $X'_b = B$ and $T = T \cup x'_b$. Return to Step-2.

• **Step-6 (Output Property)** Search for $\gamma (0 \leq \gamma \leq R)$ which is the number of rounds that $\exists x'_b$, s.t., $X'_b \neq \mathcal{R}$. For $0 \leq b \leq N - 1$, if $X'_b \neq \mathcal{R}$ holds, then $b$ becomes an element of $\beta_{\{b_0, b_1, \ldots, b_{N-1}\}}$. Output $\beta_{\{b_0, b_1, \ldots, b_{N-1}\}}$ and $\gamma$. End Algorithm.

In Algorithm B, we reanalyse higher order integral properties of subblocks which are estimated as $B$ or $E$ in Algorithm A. Let $T$ be a set of such subblocks. We construct $T$ in Step-1, and one of the elements is extracted as $x'_b$ in Step-2. Higher order integral property of $x'_b$ can be updated to $\mathcal{U}$ from Proposition 4.

![Figure 7: $\mu - \epsilon$ table of function $\phi$.](image)

Let $V' = \{x'^0_0, x'^0_1, \ldots, x'^0_{\mu' - 1}\}$ be subblocks in $r_0$-th round which satisfies

$$x'^0_b = H(x'^0_0, x'^0_1, \ldots, x'^0_{\mu' - 1}), 0 \leq r_0 \leq r - 1, 1 \leq \mu' \leq N.$$ (42)

We test injection of $E \rightarrow V'$ with increasing the value of $r_0$ in Step-3 (see condition (i) of Definition 4). Then, we search for a combination $\{e', e\}$ which satisfies $E \setminus e_0 \rightarrow V' \setminus e_0'$ is injective in Step-4 (see condition (ii) of Definition 4). If we find at least one combination, $e_0'$ is independent of $V'$. Also if $X'_b(e_0') = A$ holds, we obtain $X'_b = \mathcal{U}$ from Proposition 4. In Step-5, we search for subblocks of $\mathcal{R}$ which can be updated to $B$ in the same manner of first order integral, and such subblocks must be tested whether they can be updated to $\mathcal{U}$ or not.

We apply the above procedure for all subblocks in $T$, and update its elements. If $T$ is empty set, we go to Step-6. In Step-6, output property $\beta_{\{b_0, b_1, \ldots, b_{N-1}\}}$ and its round $\gamma$ are determined. Note that only subblock of $\mathcal{R}$ is not balanced.
In this way, we can finalize higher order integral properties and obtain integral distinguisher. In addition, we show a detailed algorithm of Step-3 and -4 in Algorithm B’.

5.3. Algorithm B

In Algorithm B’, we test injection such as Step-3 and -4 of Algorithm B. Let \( W (|W| = n) \) be a set of subblocks of intermediate rounds, and \( G \) a function of \( G : \mathbb{F}_2^{mn} \rightarrow \mathbb{F}_2^{m'} \) such as Eq.(36). We test whether \( W \rightarrow W' \) is injective or not. Note that we substitute \( W = E \) or \( W = E \backslash e_{i} \) in Algorithm B.

To test injection, we use contraposition of definition of injection. Suppose we input two plaintexts into cipher function. Let \( w_0 \) and \( w_1 \) be values of concatenation of every subblock in \( W \) \((w_0, w_1 \in \mathbb{F}_2^{mn})\). Let \( w_0' \) and \( w_1' \) be values of concatenation of every subblock in \( W' = G(W) \) \((w_0', w_1' \in \mathbb{F}_2^{m'})\). Supposing that \( w_0' = w_1' \) holds, we test whether \( w_0 = w_1 \) holds or not. Then, if \( w_0 = w_1 \) holds, we can determine \( W \rightarrow W' \) is injective.

We define two status of subblocks by equal (\( e \)) and unequal (\( \mu \)). Suppose a subblock is derived from \( x_b' = H(W) \), where \( H \) is a function of \( \mathbb{F}_2^{m} \rightarrow \mathbb{F}_2^{m'} \) (see Eq.(13)). As same as \( w_0 \) and \( w_1 \), we consider two values of \( x_0', x_1 \) \((x_0', x_1 \in \mathbb{F}_2^{m'})\). If \( x_0 = x_1 \) holds, \( x_b' = e \) and we write the status as \( x_b' = e \). Otherwise, \( x_b' = \mu \) and we write \( x_b' = \mu \).

The status \( e \) is substituted for subblocks in \( W' \) and constant subblocks and \( \mu \) is for the others.

\[
(x_b' \in W') \lor (x_b' = C) \Rightarrow x_b' = e \\
(x_b' \notin W') \land (x_b' \neq C) \Rightarrow x_b' = \mu
\]  \hspace{1cm} (43)

Considering input-output relation of each function such as function \( \phi \) (see Eq.(27)), we can determine whether each subblock is \( \mu \) or \( e \). For example, function \( \phi \) has following three characteristics.

(i) \((x_0, x_1) = (e, e) = (x_0', x_1') = (e, e)\)
(ii) \((x_0, x_1) = (e, \mu) = (x_0', x_1') = (e, \mu)\)
(iii) \((x_0, x_1) = (\mu, e) = (x_0', x_1') = (\mu, e)\)

These characteristics holds if we replace \((x_0, x_1)\) as \((x_0', x_1')\). We show six characteristics in Fig.7. We call such table as \( \mu - e \) table.

Using \( \mu - e \) table, injection of \( W \rightarrow W' \) \((|W'| = n') \) is tested in the following three steps.

- **Step-1 (Initialization)** Set \( W \) and \( W' \). If \(|W| > |W'| \) holds, then output “\( W \rightarrow W' \) is not injective” and end algorithm, else substitute \( e \) for \( \forall x_b' \in W' \) or \( \forall X_b'^{-1} \), s.t., \( X_b'^{-1} = C \). Substitute \( \mu \) for \( \forall x_b' \), s.t., \( x_b' \neq e \). Go to Step-2.

- **Step-2 (Update Status)** First, update status using \( \mu - e \) table from output to input. Next, update status using \( \mu - e \) table from input to output. As a result, if status of some subblocks are updated, then repeat Step-2, else goto Step-3.

- **Step-3 (Determination of Injection)** If \( \forall x_b' \in W, x_b' = e \) holds, then output “\( W \rightarrow W' \) is injective”, else output “\( W \rightarrow W' \) is not injective.” End algorithm.

In Step-1, we initialize status of all subblocks using Eq.(43). In Step-2, we update status from \( \mu \) to \( e \) using \( \mu - e \) table. There are two direction, from \( W' \) to \( W \) (output to input) and from \( W \) to \( W' \) (input to output). Until there is no update by applying \( \mu - e \) table, we repeat Step-2. In Step-3, if status of all subblocks in \( W \) are \( e \), \( W \rightarrow W' \) is injective from contraposition of definition of injection.
6. APPLICATION OF PROPOSAL ALGORITHM

As an application of the proposal algorithm shown in Sec.5, we search for fifteenth order integral distinguisher of TWINE and LBlock. As an example, we show fifteenth order integral distinguisher of TWINE in Fig.8 and demonstrate the proposal algorithm.

In Algorithm A, higher order integral properties are estimated by bijective path. In Step-1, \(f_{14}\) is inputted. In Step-2, we search for bijective path using bijective characteristics of function \(\phi\) such as Fig.3. Underlined subblocks are in bijective path \(P\) and ones squared with solid lines are subblocks in endpoint of bijective path \(E\). In Step-3, higher order integral properties of these subblocks are \(U\). In Step-4, higher order integral properties of subblocks outside \(P\) are estimated as \(C\), \(U\) or \(E\). Note that there are not any subblocks of \(E\). In Step-5, we search for subblocks of \(B\) and \(R\).

In Algorithm B, higher order integral properties are finalized. From Step-2 to Step-5, we search for subblocks updated from \(B\) to \(U\) or \(R\) to \(B\). Subblocks squared with broken lines are subblocks updated from \(B\) to \(U\) or \(R\) to \(B\).

We show the way to update higher order integral property of \(x_{13}^{13}\) as an example. Let \(V'_{7}\) be a set of subblocks of 7-th round. The subblock is calculated as \(x_{0}^{13} = H(V')\), where \(V' = V'_{7}\). In Step-3, we confirm \(E \rightarrow V'\) is injective using Algorithm B'. In Step-4, we find \(\{x_{11}^{13}, x_{10}^{13}\} (x_{11}^{13} \in V', x_{10}^{13} \in E)\), s.t., \(E \setminus x_{10}^{13} \rightarrow V' \setminus x_{11}^{13}\) is injective. Therefore, \(x_{11}^{13}\) is independent of \(V'\). Since \(x_{0}^{13}(x_{11}^{13}) = A\), \(x_{13}^{13}\) is updated from \(B\) to \(U\). In the same way, \(x_{13}^{13}, x_{13}^{13}\) and \(x_{13}^{13}\) are updated from \(B\) to \(U\). In Step-5, we search for subblocks of \(R\) updated to \(B\). Higher order integral properties of subblocks calculated as \(U \oplus U\) are \(B\). Therefore, \(x_{0}^{14}, x_{1}^{14}, x_{10}^{14}\) and \(x_{14}^{14}\) are updated from \(R\) to \(B\).
As a result, we obtain \( \alpha_{[2i]} \rightarrow^{15} \beta_{11,12,13,14} \), and this is the same as Eq.(9). Also, we find the same output properties are obtained if we choose a constant subblock whose index is even in TWINE. The results are summarized as follows.

\[
\alpha_{[2i]} \rightarrow^{15} \beta_{11,12,13,14}, \quad 0 \leq i \leq 7, \tag{44}
\]

where \( Q_{[2i]} \) denotes all subblocks other than \( 2i \)-th subblock are variable.

In the same way, we search for fifteenth order integral distinguisher of LBlock, and we obtain following results.

\[
\alpha_{[0]} \rightarrow^{15} \beta_{8-15}, \quad 0 \leq i \leq 7 \tag{45}
\]
These results in TWINE and LBlock are consistent with the results which are calculated from computer experiment [14], and we confirm that they are the optimal integral distinguisher.

7. DISCUSSIONS

7.1. Comparison with Conventional Search Algorithm

There are two major differences between the proposal algorithm and the conventional one. One is an approach to search for higher order integral properties from input to output. From this approach, the proposal algorithm is feasible independent of the order of integral distinguisher and the selection of input variable subblocks. Therefore, all input conditions are in the scope of the proposal algorithm. Also, integral properties of all subblocks are obtained. In the conventional algorithm, unsearched domain inevitably exists, since they extend first order integral.

The other is an approach to analyze higher order integral property. Only first order integral property is analyzed in the conventional algorithm. Since the number of plaintexts is $2^m n$, the definition of first order integral property does not represent detailed property. On the other hand, our definition of higher order integral property represents detailed property.

7.2. Computational Complexity of Proposal Algorithm

Supposing that we obtain $(N -1)th$ order integral distinguisher, we calculate computational complexity of the proposal algorithm. We estimate it by the number of times to execute LUT (lookup table), since the other operations such as variable assignment are negligible. Computational complexity of each algorithm is summarized in Table 2.

Table 2: Computational complexity of each algorithm.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>step</th>
<th>computational complexity</th>
<th>table size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algorithm A</td>
<td>Step-2</td>
<td>$N \times R$</td>
<td>$t_{A1}$</td>
</tr>
<tr>
<td></td>
<td>Step-4</td>
<td>$2 \times N^2 \times R$</td>
<td>$t_{A2}$</td>
</tr>
<tr>
<td></td>
<td>Step-5</td>
<td>$N \times R$</td>
<td>$t_{A3}$</td>
</tr>
<tr>
<td>Algorithm B</td>
<td>Step-2, -3, and -4</td>
<td>$N^5 \times R^4$</td>
<td>$t_{B'}$</td>
</tr>
<tr>
<td></td>
<td>Step-5</td>
<td>$N \times R$</td>
<td>$t_{A3}$</td>
</tr>
</tbody>
</table>

Algorithm A: In Step-2, we lookup small tables such as Fig.3 to search for subblocks in endpoint of bijective path $E$. Let $t_{A1}$ be a table size of bijective characteristics of functions. The maximum computational complexity is $N \times R$ times of LUT and the table size is $t_{A1}$.

In Step-4, we calculate first order integral property of all subblocks from each subblock in $V$ and $E$. Since $|V| = |E| < N$, we regard the number of subblocks in $V$ and $E$ as $N$. We prepare a table to
calculate first order integral property such as $A \oplus A = B$. Let $t_{A2}$ be a table size of such calculations. There are $2 \times N$ subblocks in $V$ and $E$, and first order integral properties of $N \times R$ subblocks are calculated at most. Therefore, the maximum computational complexity becomes $(2 \times N) \times (N \times R) = 2 \times N^2 \times R$ times of LUT and the table size is $t_{A2}$.

In Step-5, we calculate higher order integral properties in the same manner of first order integral. Let $t_{A3}$ be a table size of such calculations. The maximum computational complexity is $N \times R$ times of LUT and the table size is $t_{A3}$.

As the result, the maximum computational complexity of Algorithm A is $2 \times N^2 \times R$ times of LUT. And the maximum block ciphers. Let $t_A = \max[t_{A1}, t_{A2}, t_{A3}, t_{A4}, t_{A5}]$ be such table size.

**Algorithm B:** From Step-2 to Step-4, we test whether subblocks of $\mathcal{B}$ and $\mathcal{E}$ are updated to $\mathcal{U}$. The number of subblocks to test is $N \times R$ at most (Step-2). For every $N \times R$ subblocks, Step-3 is executed. We repeat Step-3 for $R$ times, so that, we execute it for $(N \times R) \times R = N^2 \times R^2$ times at most. We suppose that Step-4 is executed for each time to execute Step-3. In Step-4, we test injection of $E \cup E'$ for $N^2$ combinations of $\{e_i, e_i'\} (e_i, e_i' \in V', e_i, e_i' \in E, |V'| \leq N)$. Therefore, we execute Algorithm B' for $(N \times R)^2 \times N^2 = N^5 \times R^3$ times from Step-2 to Step-4. Then, we consider computational complexity of Algorithm B'. In Step-2 of Algorithm B', we lookup $\mu \to e$ table to determine whether subblocks are updated from $\mu$ to $e$. Let $t_B$ be size of $\mu - e$ table. The maximum computational complexity is $N^5 \times R^3$ times of LUT and the table size is $t_B$. As the result, the maximum computational complexity of Step-2, -3 and -4 of Algorithm B is $(N^3 \times R^3) \times (N^2 \times R^2) = N^5 \times R^4$ times LUT and the table size is $t_B$.

In Step-5, we calculate higher order integral properties in the same manner of first order integral. Since the calculation is the same as Step-5 of Algorithm A, computational complexity is $N \times R$ times of LUT and the table size is $t_{A3}$.

The maximum computational complexity of Algorithm B is $N^5 \times R^4$ times of LUT. The maximum table size of all tables used in Algorithm B is $t_B = \max[t_{B1}, t_{B2}, t_{B3}, t_{B4}]$.

From the above, the maximum computational complexity of the proposal algorithm is $N^5 \times R^4$ times LUT. And the maximum table size is $\max[t_A, t_B]$. On the contrary, computational complexity of the conventional algorithm is estimated as $N \times R$ times of LUT and table size is $t_{A5}$. Therefore, the conventional algorithm is faster than the proposal one.

However, we can execute the proposal algorithm even in general-purpose computers, since $N \leq 32$ and $R \leq 100$ in all of existing subblock-based block ciphers.

### 8. CONCLUSION

In this paper, we propose a new algorithm to search for higher order integral distinguisher in subblock-based block ciphers. Unlike the conventional algorithm, we search for higher order integral properties from input to output. Exploiting bijective and injective components of cipher functions, we can search for higher order integral distinguisher effectively. We apply the proposal
algorithm to TWINE and LBlock. As the result, we confirm that our results of the proposal algorithm are the optimal integral distinguisher. Designers of block ciphers must consider integral distinguisher obtained by the proposal algorithm. There is a possibility that the number of rounds to be attacked increases by the algorithm. Even if it is not, it can be less difficult to guess all of the secret keys from increased balanced subblocks. Therefore, the designers need to consider such vulnerabilities and select stronger cipher algorithm and key schedule.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number 24560491.

REFERENCES


61
