# Links among Impossible Differential, Integral and Zero Correlation Linear Cryptanalysis [⋆]

Bing Sun[1,3], Zhiqiang Liu[2,3,], Vincent Rijmen[3], Ruilin Li[4], Lei Cheng[1], Qingju Wang[2,3], Hoda Alkhzaimi[5], Chao Li[1]

[1] College of Science, National University of Defense Technology, Changsha, Hunan, P. R. China, 410073
[2] Dept. Computer Science and Engineering, Shanghai Jiao Tong University, China
[3] Dept. Electrical Engineering (ESAT), KU Leuven and iMinds, Belgium
[4] College of Electronic Science and Engineering, National University of Defense Technology,
Changsha, Hunan, P. R. China, 410073
[5] Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark
`happy_come@163.com,ilu_zq@sjtu.edu.cn`

**Abstract.** As two important cryptanalytic methods, impossible differential cryptanalysis and integral cryptanalysis have attracted much attention in recent years. Although relations among other important cryptanalytic approaches have been investigated, the link between these two methods has been missing. The motivation in this paper is to fix this gap and establish links between impossible differential cryptanalysis and integral cryptanalysis.

Firstly, by introducing the concept of structure and dual structure, we prove that $a \to b$ is an impossible differential of a structure $\mathcal{E}$ if and only if it is a zero correlation linear hull of the dual structure $\mathcal{E}^{\perp}$. More specifically, constructing a zero correlation linear hull of a Feistel structure with $SP$-type round function where $P$ is invertible, is equivalent to constructing an impossible differential of the same structure with $P^T$ instead of $P$. Constructing a zero correlation linear hull of an SPN structure is equivalent to constructing an impossible differential of the same structure with $(P^{-1})^T$ instead of $P$. Meanwhile, our proof shows that the automatic search tool presented by Wu and Wang could find all impossible differentials of both Feistel structures with $SP$-type round functions and SPN structures, which is useful in provable security of block ciphers against impossible differential cryptanalysis.

Secondly, by establishing some boolean equations, we show that a zero correlation linear hull always indicates the existence of an integral distinguisher while a special integral implies the existence of a zero correlation linear hull. With this observation we improve the integral distinguishers of Feistel structures by 1 round, build a 24-round integral distinguisher of CAST-256 based on which we propose the best known key recovery attack on reduced round CAST-256 in the non-weak key model, present a 12-round integral distinguisher of SMS4 and an 8-round integral distinguisher of Camellia without $FL/FL^{-1}$. Moreover, this result provides a novel way for establishing integral distinguishers and converting known plaintext attacks to chosen plaintext attacks.

Finally, we conclude that an $r$-round impossible differential of $\mathcal{E}$ always leads to an $r$-round integral distinguisher of the dual structure $\mathcal{E}^{\perp}$. In the case that $\mathcal{E}$ and $\mathcal{E}^{\perp}$ are linearly equivalent, we derive a direct link between impossible differentials and integral distinguishers of $\mathcal{E}$. Specifically, we obtain that an $r$-round impossible differential of an SPN structure, which adopts a bit permutation as its linear layer, always indicates the existence of an $r$-round integral distinguisher. Based on this newly established link, we deduce that impossible differentials of SNAKE(2), PRESENT, PRINCE and ARIA, which are independent of the choices of the $S$-boxes, always imply the existence of integral distinguishers.

Our results could help to classify different cryptanalytic tools. Furthermore, when designing a block cipher, the designers need to demonstrate that the cipher has sufficient security margins against important cryptanalytic approaches, which is a very tough task since there have been so many cryptanalytic tools up to now. Our results certainly facilitate this security evaluation process.

**Keywords:** Impossible Differential, Integral, Zero Correlation Linear, Feistel, SPN, Camellia, CAST-256, SMS4, SNAKE(2), PRESENT, PRINCE, ARIA

[⋆] The work in this paper is supported by the Natural Science Foundation of China(No: 61103192, 61070215, 61202371, 61402515).

## 1 Introduction

Block ciphers are considered vital elements in constructing many symmetric cryptographic schemes such as encryption algorithms, hash functions, authentication schemes and pseudo-random number generators. The core security of these schemes depends on the resistance of the underlying block ciphers to known cryptanalytic techniques. So far a variety of cryptanalytic techniques have been proposed such as impossible differential cryptanalysis [1, 2], integral cryptanalysis [3], zero correlation linear cryptanalysis [4], etc.

Impossible differential cryptanalysis was independently proposed by Knudsen [1] and Biham [2]. One of the most popular impossible differentials is called a truncated impossible differential. It is independent of the choices of the $S$-boxes. Several approaches have been proposed to derive truncated impossible differentials of a block cipher/structure effectively such as the $\mathcal{U}$-method [5], *UID*-method [6] and the extended tool of the former two methods generalized by Wu and Wang in Indocrypt 2012 [7]. Integral cryptanalysis [3], also known as square attack[8], saturation attack [9], multi-set attack [10], higher-order differential attack [11, 12], was first proposed by Knudsen and Wagner. With some special inputs, we check whether the sum of the corresponding ciphertexts is zero or not. Usually, we do not need to investigate the details of the $S$-boxes and only view the $S$-boxes as some bijective transformations over finite fields. Zero correlation linear cryptanalysis, proposed by Bogdanov and Rijmen in [4], tries to construct some linear hulls with correlation exactly zero. In most cases, as in impossible differential and integral cryptanalysis, we do not need to investigate the details of the $S$-boxes. Generally, though there has been lots of work concentrating on the design and cryptanalysis of $S$-boxes [13], most cryptanalytic results by using impossible differential, integral and zero correlation linear cryptanalysis are independent of the choices of the $S$-boxes. If we choose some other $S$-boxes in a cipher, the corresponding cryptanalytic results will remain almost the same.

Along with the growing of the list of cryptanalytic tools, the question whether there are direct links or any connections among different tools has drawn much attention of the cryptographic research community, since such relations can be used to compare the effectiveness of different tools as well as to improve cryptanalytic results on block ciphers.

Efforts to find and build the links among different cryptanalytic techniques were initiated by Chabaud and Vaudenay in [14], where a theoretical link between differential and linear cryptanalysis was presented. After that, many attempts have been made to establish further relations among various cryptanalytic tools. In [15], Sun *et al.* proved that from an algebraic view, integral cryptanalysis can be seen as a special case of the interpolation attack. In [16], Leander stated that statistical saturation distinguishers are averagely equivalent to multidimensional linear distinguishers. In [17], Bogdanov *et al.* showed that an integral implies a zero correlation linear hull unconditionally, a zero correlation linear hull indicates an integral distinguisher under certain conditions, and a zero correlation linear hull is actually a special case of multidimensional linear distinguishers. In [18], Blondeau and Nyberg further analyzed the link between differential and linear cryptanalysis and demonstrated some new insights on this link to make it more applicable in practice. They established new formulas between the probability of truncated differentials and the correlation of linear hulls. This link was later applied in [19] to provide an exact expression of the bias of a differential-linear approximation. Moreover, they claimed that the existence of a zero correlation linear hull is equivalent to the existence of an impossible differential in some specific cases. As shown in [20], this link is usually not practical for most known impossible differential or zero correlation linear distinguishers, since the sum of the dimensions of input and output of each distinguisher is always the block size of the cipher, which means if the dimension parameter for one type is small, it should be infeasible large for the other type. Blondeau *et al.* proposed a practical relation between these two distinguishers for Feistel-type and Skipjack-type ciphers and showed some equivalence between impossible differentials and zero correlation linear hulls with respect to Feistel-type and Skipjack-type ciphers. In [21], Blondeau and Nyberg gave the link between truncated differential and multidimensional linear approximation, and then applied this link to explore the relations between the complexities of chosen-plaintext and known-plaintext distinguishing/key recovery attacks of differential and linear types. Moreover, they showed that statistical saturation cryptanalysis is indeed equivalent to truncated differential cryptanalysis, which could be used to estimate the data requirement of the statistical saturation key recovery attack.

**Contributions.** Although there have been intriguing results with respect to the relations among some important cryptanalytic approaches, the link between impossible differential cryptanalysis and integral cryptanalysis is still missing. In this paper, we aim to explore the link between these two cryptanalytic methods. Since the fundamental step in statistical cryptanalysis of block ciphers is to construct effective distinguishers, we focus on building the links among impossible differential, zero correlation linear and integral cryptanalysis from the aspect of distinguishers. Our main contributions are as follows (see Fig.1).

1. To characterize what "being independent of the choices of $S$-boxes" means, we propose the definition of *structure* $\mathcal{E}$, which is a set containing some ciphers that are "similar" to each other. Then, by introducing the *dual structure* $\mathcal{E}^{\perp}$, we prove that $a \to b$ is an impossible differential of $\mathcal{E}$ if and only if it is a zero correlation linear hull of $\mathcal{E}^{\perp}$. More specifically, let $P^T$ and $P^{-1}$ denote the transpose and inverse of $P$ respectively. Then for a Feistel structure with $SP$-type round functions where $P$ is invertible, denoted as $\mathcal{F}_{SP}$, constructing an $r$-round zero correlation linear hull is equivalent to constructing an impossible differential of $\mathcal{F}_{SP^T}$, which is the same structure as $\mathcal{F}_{SP}$ with $P^T$ instead of $P$; For an SPN structure $\mathcal{E}_{SP}$, constructing an $r$-round zero correlation linear hull of $\mathcal{E}_{SP}$ is equivalent to constructing an impossible differential of $\mathcal{E}_{S(P^{-1})^T}$, which is the same structure as $\mathcal{E}_{SP}$ with $(P^{-1})^T$ instead of $P$. Based on this result, we find 8-round zero correlation linear hulls of Camellia without $FL/FL^{-1}$ layer and 4-round zero correlation linear hulls of ARIA.
2. We show that the automatic search tool, presented by Wu and Wang in Indocrypt 2012, could find all impossible differentials of a cipher that are independent of the choices of the $S$-boxes. This can be used in provable security of block ciphers against impossible differential cryptanalysis.
3. We find that a zero correlation linear hull always implies the existence of an integral distinguisher, which means the conditions used for deriving integral distinguisher from zero correlation linear hull in [17] can be removed. This observation also provides a novel way for constructing integral distinguisher and converting known plaintext attacks to chosen plaintext attacks. Meanwhile, we observe that the statement "*integral unconditionally implies zero correlation linear hull*" in [17] is correct only under the definition that integral property is a balanced vectorial boolean function, while it does not hold for the general case. For example, up to date we cannot use the integral distinguisher for 4-round AES (with extra MixColumns) [4, 8] to construct a zero correlation linear hull.
4. Following the results given above, we build the link between impossible differential cryptanalysis and integral cryptanalysis, i.e., an $r$-round impossible differential of a structure $\mathcal{E}$ always implies the existence of an $r$-round integral distinguisher of $\mathcal{E}^{\perp}$. Moreover, in the case that $\mathcal{E}^{\perp} = A_2\mathcal{E}A_1$ where $A_1$ and $A_2$ are linear transformations, we could get direct links between impossible differential cryptanalysis and integral cryptanalysis of $\mathcal{E}$. Specifically, an $r$-round impossible differential of SPN structure which adopts bit permutation as the linear layer, always leads to an $r$-round integral distinguisher.
5. We improve the integral distinguishers of Feistel structures by 1 round, build a 24-round integral distinguisher of CAST-256, and present a 12-round integral distinguisher of SMS4 which is 2-round longer than previously best known ones and an 8-round integral distinguisher of Camellia without $FL/FL^{-1}$ layser which is 2-round longer than the best known ones which are independent with the choices of the $S$-box. These distinguishers could not be obtained by the known methods for constructing integral distinguishers or by using the link given in [17]. As an example, the best known key recovery attack on reduced round CAST-256 in non-weak key model is given to show the effectiveness of the newly constructed distinguishers.

In [18] and [21], the sum of the dimensions of input and output differences (masks) of an impossible differential (zero correlation linear hull) is always the block size of the cipher, therefore, the link between impossible differential and zero correlation linear hull is usually not practical. This constraint has been removed in this paper as well as in [20]. Compared with [20], our paper takes more complicated structures into account and exploits more details of the round function, thus leading to a more practical and applicable link between impossible differential and zero correlation linear cryptanalysis.

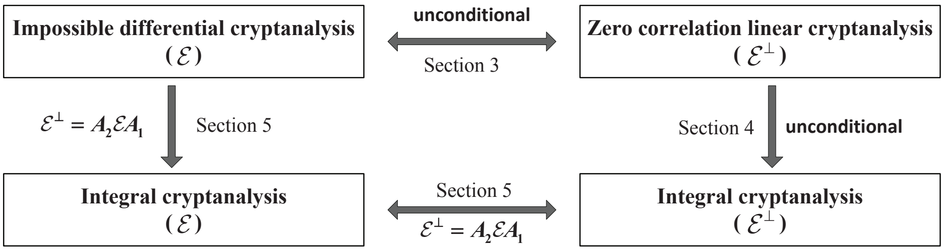

**Fig. 1.** Links among Impossible Differential, Integral and Zero Correlation Linear Cryptanalysis, where $\mathcal{E}$ is a structure and $\mathcal{E}^\perp$ is the dual structure of $\mathcal{E}$, $A_1$ and $A_2$ are linear transformations applied before the input and after the output of $\mathcal{E}$.

**Organization.** The remainder of this paper is organized as follows. Sec. 2 introduces the notations and concepts that will be used throughout the paper. In Sec. 3, we establish the new links between impossible differential cryptanalysis and zero correlation linear cryptanalysis. Sec. 4 shows the refined link between integral cryptanalysis and zero correlation linear cryptanalysis. The link between impossible differential cryptanalysis and integral cryptanalysis is presented in Sec. 5. Then in Sec. 6, we give some examples to show the effectiveness of the newly established links in constructing new distinguishers of block ciphers. Finally, Sec. 7 concludes this paper.

## 2 Preliminaries

### 2.1 Boolean Functions

This section recalls the notations and concepts [22] which will be used throughout this paper. Let $\mathbb{F}_2$ denote the finite field with two elements, and $\mathbb{F}_2^n$ be the vector space over $\mathbb{F}_2$ with dimension $n$. Let $a = (a_1, \ldots, a_n), b = (b_1, \ldots, b_n) \in \mathbb{F}_2^n$. Then

$$a \cdot b \triangleq a_1 b_1 \oplus \cdots \oplus a_n b_n$$

denotes the *inner product* of $a$ and $b$. Note that the inner product of $a$ and $b$ can be written as $ab^T$ where $b^T$ stands for the *transpose* of $b$ and the multiplication is defined as matrix multiplication. Given a function $G : \mathbb{F}_2^n \to \mathbb{F}_2$, the *correlation* of $G$ is defined by

$$c(G(x)) \triangleq \frac{\#\{x \in \mathbb{F}_2^n | G(x) = 0\} - \#\{x \in \mathbb{F}_2^n | G(x) = 1\}}{2^n} = \frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} (-1)^{G(x)}.$$

Given a vectorial function $H : \mathbb{F}_2^n \to \mathbb{F}_2^k$, the *correlation* of the linear approximation for a $k$-bit output mask $b$ and an $n$-bit input mask $a$ is defined by

$$c(a \cdot x \oplus b \cdot H(x)) \triangleq \frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} (-1)^{a \cdot x \oplus b \cdot H(x)}.$$

If $c(a \cdot x \oplus b \cdot H(x)) = 0$, then $a \to b$ is called a *zero correlation linear hull* of $H$[4]. This definition can be extended as follows: Let $A \subseteq \mathbb{F}_2^n$, $B \subseteq \mathbb{F}_2^k$. If for all $a \in A$, $b \in B$, $c(a \cdot x \oplus b \cdot H(x)) = 0$, then $A \to B$ is called a *zero correlation linear hull* of $H$. In the case that $H$ is a permutation on $\mathbb{F}_2^n$, for any $b \neq 0$, $c(b \cdot H(x)) = 0$ and for any $a \neq 0$, $c(a \cdot x) = 0$. We call $0 \to b$ and $a \to 0$ trivial zero correlation linear hulls of $H$ where $a \neq 0$ and $b \neq 0$. Let $A \subseteq \mathbb{F}_2^n$. If the size of the set

$$H_A^{-1}(y) \triangleq \{x \in A | H(x) = y\}$$

is independent of $y \in \mathbb{F}_2^k$, we say $H$ is *balanced on* $A$. Specifically, if $A = \mathbb{F}_2^n$, we say $H$ is a *balanced function*. If the sum of all images of $H$ is 0, i.e.

$$\sum_{x \in \mathbb{F}_2^n} H(x) = 0,$$

we say $H$ has an *integral-balanced (zero-sum)* property[3]. Let $\delta \in \mathbb{F}_2^n$ and $\Delta \in \mathbb{F}_2^k$. The differential probability of $\delta \to \Delta$ is defined as

$$p(\delta \to \Delta) \triangleq \frac{\#\{x \in \mathbb{F}_2^n | H(x) \oplus H(x \oplus \delta) = \Delta\}}{2^n}.$$

If $p(\delta \to \Delta) = 0$, then $\delta \to \Delta$ is called an *impossible differential* of $H$[1, 2]. Let $A \subseteq \mathbb{F}_2^n$, $B \subseteq \mathbb{F}_2^k$. If for all $a \in A$ and $b \in B$, $p(a \to b) = 0$, $A \to B$ is called an *impossible differential* of $H$.

We recall the following property of balanced boolean functions: a function $G : \mathbb{F}_2^n \to \mathbb{F}_2$ is balanced if and only if $c(G(x)) = 0$.

### 2.2 Block Ciphers

**Feistel Ciphers.** An $r$-round Feistel cipher $E$ is defined as follows:

Let $(L_0, R_0) \in \mathbb{F}_2^{2n}$ be the input of $E$. Iterate the following transformation $r$ times:

$$\begin{cases} L_{i+1} = F_i(L_i) \oplus R_i \\ R_{i+1} = L_i \end{cases} \quad 0 \le i \le r-1,$$

where $L_i, R_i \in \mathbb{F}_2^n$. The output of the $r$-th iteration is defined as the output of $E$. In this paper, we will focus on the case that $F_i$'s are SP-type functions which will be defined in the following.

**SPN Ciphers.** The SPN structure is widely used in constructing cryptographic primitives. It iterates some SP-type round functions to achieve confusion and diffusion. Specifically, the SP-type function $f : \mathbb{F}_2^{s\times t} \to \mathbb{F}_2^{s\times t}$ used in this paper is defined as follows:

Assume the input $x$ is divided into $t$ pieces $x = (x_0, \ldots, x_{t-1})$, and each of the $x_i$'s is an $s$-bit word. Then apply the nonlinear transformation $S_i$ to $x_i$ and let $y = (S_0(x_0), \ldots, S_{t-1}(x_{t-1})) \in \mathbb{F}_2^{s\times t}$. At last, apply a linear transformation $P$ to $y$, and $Py$ is the output of $f$.

The following strategies are popular in designing the diffusion layer $P$ of a cipher:

(1) $P$ is a bit-wise permutation of $\mathbb{F}_2^{s\times t}$ as in PRESENT [23]. PRESENT is an SPN block cipher with block length 64-bit. It is a lightweight block cipher primarily designed for hardware constrained environments such as RFID tags and sensor networks. PRESENT adopts bit permutation as the diffusion layer $P$, which can be defined as a permutation matrix $P = (P_{i,j})_{64\times 64}$:

$$P_{i,j} = \begin{cases} 1 & \text{if } j = 16i \bmod 63 \\ 0 & \text{otherwise} \end{cases} .$$

(2) Each bit of $Py$ is a sum of some bits of $y$ as in PRINCE [24]. PRINCE is a lightweight block cipher with block size 64-bit. The core component of PRINCE is PRINCE$_{core}$ which adopts a 12-round SPN structure. Firstly, we will define $SR$ and $M'$ as follows:

$SR$ behaves like the ShiftRows in AES and permutes the 16 nibbles of PRINCE in the following way.

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

Therefore it is also a permutation of 64 bits and we could write $SR$ as a permutation matrix in $\mathbb{F}_2^{64\times 64}$.

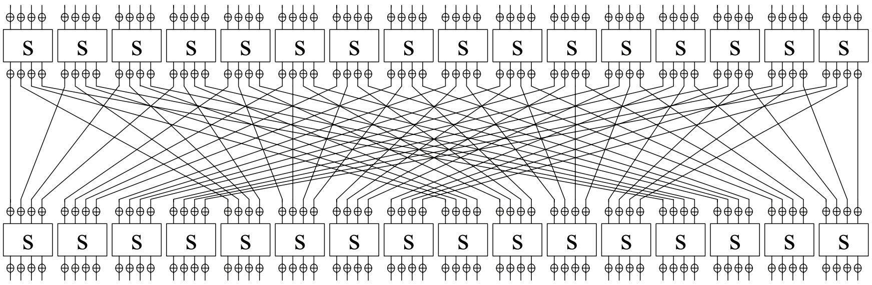

**Fig. 2.** Round-function of PRESENT

To construct $M'$, we first define

$$\hat{M}^{(0)}=\begin{pmatrix}M_0 & M_1 & M_2 & M_3\\ M_1 & M_2 & M_3 & M_0\\ M_2 & M_3 & M_0 & M_1\\ M_3 & M_0 & M_1 & M_2\end{pmatrix},\quad \hat{M}^{(1)}=\begin{pmatrix}M_1 & M_2 & M_3 & M_0\\ M_2 & M_3 & M_0 & M_1\\ M_3 & M_0 & M_1 & M_2\\ M_0 & M_1 & M_2 & M_3\end{pmatrix}$$

where

$$M_0=\begin{pmatrix}0&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix},M_1=\begin{pmatrix}1&0&0&0\\0&0&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix},M_2=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&0&1\end{pmatrix},M_3=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&0\end{pmatrix},$$

and then we define $M'=\mathrm{diag}(\hat{M}^{(0)},\hat{M}^{(1)},\hat{M}^{(1)},\hat{M}^{(0)})$, which is a $64\times 64$ block diagonal matrix.

$M'$ is used as the linear transformation of the middle round. The transformations $M=SR\circ M'$ and $M^{-1}$ are used before and after the middle round, respectively.

(3) Each word of $Py$ is a sum of some words of $y$ as in Camellia [25] and ARIA [26]. The block cipher Camellia was recommended in the NESSIE block cipher portfolio in 2003 and selected as a new international standard by ISO/IEC in 2005. It adopts the Feistel structure with invertible SP-type round functions if not taking into account the $FL/FL^{-1}$ layer. The linear transformation $P$ could be written as follows:

$$P=\begin{pmatrix}E&0&E&E&0&E&E&E\\E&E&0&E&E&0&E&E\\E&E&E&0&E&E&0&E\\0&E&E&E&E&E&E&0\\E&E&0&0&0&E&E&E\\0&E&E&0&E&0&E&E\\0&0&E&E&E&E&0&E\\E&0&0&E&E&E&E&0\end{pmatrix}$$

where $E$ and $0$ denote $8\times 8$ identity and zero matrices, respectively.

ARIA is a 128-bit block cipher established as a Korean Standard by the Ministry of Commerce, Industry and Energy in 2004. The diffusion layer of ARIA can be written as:

$$P = \begin{pmatrix}
0 & 0 & 0 & E & E & 0 & E & 0 & E & E & 0 & 0 & 0 & E & E & 0 \\
0 & 0 & E & 0 & 0 & E & 0 & E & E & E & 0 & 0 & E & 0 & 0 & E \\
0 & E & 0 & 0 & E & 0 & E & 0 & 0 & 0 & E & E & E & 0 & 0 & E \\
E & 0 & 0 & 0 & 0 & E & 0 & E & 0 & 0 & E & E & 0 & E & E & 0 \\
E & 0 & E & 0 & 0 & E & 0 & 0 & E & 0 & 0 & E & 0 & 0 & E & E \\
0 & E & 0 & E & E & 0 & 0 & 0 & 0 & E & E & 0 & 0 & 0 & E & E \\
E & 0 & E & 0 & 0 & 0 & 0 & E & 0 & E & E & 0 & E & E & 0 & 0 \\
0 & E & 0 & E & 0 & 0 & E & 0 & E & 0 & 0 & E & E & E & 0 & 0 \\
E & E & 0 & 0 & E & 0 & 0 & E & 0 & 0 & E & 0 & 0 & E & 0 & E \\
E & E & 0 & 0 & 0 & E & E & 0 & 0 & 0 & 0 & E & E & 0 & E & 0 \\
0 & 0 & E & E & 0 & E & E & 0 & E & 0 & 0 & 0 & 0 & E & 0 & E \\
0 & 0 & E & E & E & 0 & 0 & E & 0 & E & 0 & 0 & E & 0 & E & 0 \\
0 & E & E & 0 & 0 & 0 & E & E & 0 & E & 0 & E & E & 0 & 0 & 0 \\
E & 0 & 0 & E & 0 & 0 & E & E & E & 0 & E & 0 & 0 & E & 0 & 0 \\
E & 0 & 0 & E & E & E & 0 & 0 & 0 & E & 0 & E & 0 & 0 & E & 0 \\
0 & E & E & 0 & E & E & 0 & 0 & E & 0 & E & 0 & 0 & 0 & 0 & E
\end{pmatrix}.$$

where $E$ and 0 are the $8 \times 8$ identity and zero matrices, respectively.

(4) Each word of $Py$, seen as an element of some extension fields of $\mathbb{F}_2$, is a linear combination of some other words of $y$ as in the AES. In the following, we will use the matrix expression of finite fields to show how to write the linear layer of AES as a $128 \times 128$ binary matrix:

Since ShiftRows is a permutation on 16 bytes, it is also a permutation on 128 bits. Therefore, as in the discussion above, we can represent ShiftRows as a permutation matrix $M_{SR}$ in $\mathbb{F}_2^{128\times128}$. Let $\mathbb{F}_{2^8} = \mathbb{F}_2[x]/ < f(x) >$ where $\mathbb{F}_2[x]$ is the polynomial ring over $\mathbb{F}_2$, $f(x) = x^8 + x^4 + x^3 + x + 1 \in \mathbb{F}_2[x]$ is the defining polynomial of $\mathbb{F}_{2^8}$. Then $1 = (00000001) \in \mathbb{F}_{2^8}$ can be written as the $8 \times 8$ identity matrix $E$, $2 = (00000010) \in \mathbb{F}_{2^8}$ can be written as the following $8 \times 8$ matrix:

$$M_2 = \begin{pmatrix}
0&0&0&0&0&0&0&1 \\
1&0&0&0&0&0&0&1 \\
0&1&0&0&0&0&0&0 \\
0&0&1&0&0&0&0&1 \\
0&0&0&1&0&0&0&1 \\
0&0&0&0&1&0&0&0 \\
0&0&0&0&0&1&0&0 \\
0&0&0&0&0&0&1&0
\end{pmatrix}$$

and the matrix representation of $3 = (00000011)$ is $M_3 = E \oplus M_2$. If we substitute 1, 2 and 3 in MixColumns by $E$, $M_2$ and $M_3$, respectively, we get a $128 \times 128$ binary matrix $M_{MC}$ and the linear layer of AES can be written as $M_{MC}M_{SR}$ which is a $128 \times 128$ matrix over $\mathbb{F}_2$.

Generally, no matter which linear transformation a cipher adopts, it is always linear over $\mathbb{F}_2$. Therefore, $P$ can always be written as a multiplication by a matrix which leads to the following definition:

**Definition 1.** *Let $P$ be a linear transformation over $\mathbb{F}_2^m$ for some positive integer $m$. The matrix representation of $P$ over $\mathbb{F}_2$ is called the* primitive representation *of $P$.*

### 2.3 Structure and Dual Structure

In many cases, when constructing impossible differentials and zero correlation linear hulls, we are only interested in detecting whether there is a difference (mask) of an $S$-box or not regardless of the value of this difference (mask). For example, the truncated impossible differential and zero correlation linear hull of AES in [4, 27] and Camellia in [28, 29]. In other words, if these ciphers adopt some other $S$-boxes, these distinguishers still hold. This leads to the following definition:

**Definition 2.** *Let $E : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be a block cipher with bijective $S$-boxes as the basic non-linear components.*

(1) *A* structure $\mathcal{E}^E$ *on $\mathbb{F}_2^n$ is defined as a set of block ciphers $E'$ which is exactly the same as $E$ except that the $S$-boxes can take all possible bijective transformations on the corresponding domains.*
(2) *Let $a, b \in \mathbb{F}_2^n$. If for any $E' \in \mathcal{E}^E$, $a \to b$ is an impossible differential (zero correlation linear hull) of $E'$, $a \to b$ is called an impossible differential (zero correlation linear hull) of $\mathcal{E}^E$.*

**Note.** In the definition of $\mathcal{E}^E$, if $E$ uses bijective $S$-boxes, then the $S$-boxes in $\mathcal{E}^E$ should be bijective. However, if $S$-boxes used in $E$ are not necessarily bijective, then $\mathcal{E}^E$ could be defined as a set of block ciphers $E'$ which is exactly the same as $E$ except that the $S$-boxes can take all possible transformations on the corresponding domains. As discussed above, the truncated impossible differentials and zero correlation linear hulls of AES and Camellia found so far are actually the impossible differentials and zero correlation linear hulls of $\mathcal{E}^{\text{AES}}$ and $\mathcal{E}^{\text{Camellia}}$.

**Definition 3.** *Let $\mathcal{F}_{SP}$ be a* Feistel structure *with $SP$-type round function, and let the primitive representation of the linear transformation be $P$. Let $\sigma$ be the operation that exchanges the left and right halves of a state. Then the* dual structure $\mathcal{F}_{SP}^{\perp}$ *of $\mathcal{F}_{SP}$ is defined as $\sigma\mathcal{F}_{P^TS}\sigma$.*

*Let $\mathcal{E}_{SP}$ be an* SPN structure *with primitive representation of the linear transformation being $P$. Then the* dual structure $\mathcal{E}_{SP}^{\perp}$ *of $\mathcal{E}_{SP}$ is defined as $\mathcal{E}_{S(P^{-1})^T}$.*

## 3 Links between Impossible Differential and Zero Correlation Linear Cryptanalysis

In this section, we will show the equivalence between impossible differentials and zero correlation linear hulls of a structure, which will be used to establish the link between impossible differential and integral cryptanalysis in Sec.5.

**Theorem 1.** *$a \to b$ is an $r$-round impossible differential of $\mathcal{F}_{SP}$ if and only if it is an $r$-round zero correlation linear hull of $\mathcal{F}_{SP}^{\perp}$.*

*Proof.* The proof can be divided into the following two parts (See Fig.3):

**Part (I)** In this part, we prove that for $(\delta_0, \delta_1) \to (\delta_r, \delta_{r+1})$, if one can find $E \in \mathcal{F}_{SP}^{\perp}$ such that $c((\delta_0, \delta_1) \cdot x \oplus (\delta_r, \delta_{r+1}) \cdot E(x)) \neq 0$, then one can find $E' \in \mathcal{F}_{SP}$ such that $p((\delta_1, \delta_0) \to (\delta_{r+1}, \delta_r)) > 0$.

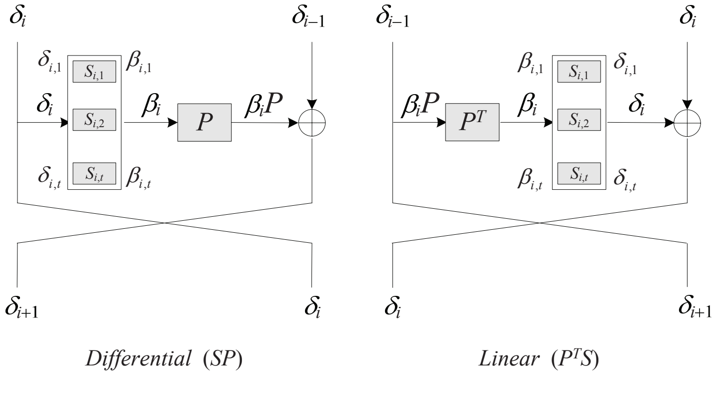

**Fig. 3.** Differential Propagation of $\mathcal{F}_{SP}$ and Linear Propagation of $\mathcal{F}_{SP}^{\perp}$

Assume that $(\delta_0, \delta_1) \to (\delta_r, \delta_{r+1})$ is a linear hull with non-zero correlation for some $E \in \mathcal{F}_{SP}^{\perp}$, and the input to the round function could be divided into $t$ pieces, each of which is an $s$-bit word. Then there exists a linear characteristic with non-zero correlation:

$$(\delta_0, \delta_1) \to \cdots (\delta_{i-1}, \delta_i) \to \cdots \to (\delta_r, \delta_{r+1}),$$

where $\delta_i \in (\mathbb{F}_2^s)^t$. In this characteristic, let the output mask of $S_i = (S_{i,1}, \ldots, S_{i,t})$ be $\delta_i = (\delta_{i,1}, \ldots, \delta_{i,t}) \in (\mathbb{F}_2^s)^t$, and let the input mask of $S_i$ be $\beta_i = (\beta_{i,1}, \ldots, \beta_{i,t}) \in (\mathbb{F}_2^s)^t$. Since for $\gamma \neq \beta_i P$, $c(\gamma \cdot x \oplus \beta_i \cdot (xP^T)) = 0$, $\delta_{i+1} = \delta_{i-1} \oplus \beta_i P$.

In the following, for any $(x_L, x_R) = (x_{L,1}, \ldots, x_{L,t}, x_{R,1}, \ldots, x_{R,t}) \in (\mathbb{F}_2^s)^t \times (\mathbb{F}_2^s)^t$, we will construct an $r$-round cipher $E_r \in \mathcal{F}_{SP}$, such that $E_r(x_L, x_R) \oplus E_r(x_L \oplus \delta_1, x_R \oplus \delta_0) = (\delta_{r+1}, \delta_r)$.

If $r = 1$, for $j \in \{1, \ldots, t\}$: if $\delta_{1,j} = 0$, we can define $S_{1,j}$ as any possible transformation on $\mathbb{F}_2^s$, and if $\delta_{1,j} \neq 0$, we can define

$$S_{1,j}(x_{L,j}) = x_{L,j}, \quad S_{1,j}(x_{L,j} \oplus \delta_{1,j}) = x_{L,j} \oplus \beta_{1,j},$$

then for $E_1 \in \mathcal{F}_{SP}$ which adopts such $S$-boxes,

$$E_1(x_L, x_R) \oplus E_1(x_L \oplus \delta_1, x_R \oplus \delta_0) = (\delta_0 \oplus \beta_1 P, \delta_1) = (\delta_2, \delta_1).$$

Suppose that we have constructed $E_{r-1}$ such that $E_{r-1}(x_L, x_R) \oplus E_{r-1}(x_L \oplus \delta_1, x_R \oplus \delta_0) = (\delta_r, \delta_{r-1})$. Denote by $(y_L, y_R) = (y_{L,1}, \ldots, y_{L,t}, y_{R,1}, \ldots, y_{R,t})$ the output of $E_{r-1}(x_L, x_R)$. Then in the $r$-th round, if $\delta_{r,j} = 0$, we can define $S_{r,j}$ as any possible transformation on $\mathbb{F}_2^s$, otherwise, define $S_{r,j}$ as follows:

$$S_{r,j}(y_{L,j}) = y_{L,j}, \quad S_{r,j}(y_{L,j} \oplus \delta_{r,j}) = y_{L,j} \oplus \beta_{r,j}.$$

Therefore $E_r(x_L, x_R) \oplus E_r(x_L \oplus \delta_1, x_R \oplus \delta_0) = (\delta_{r-1} \oplus \beta_r P, \delta_r) = (\delta_{r+1}, \delta_r)$.

**Part (II)** In this part, we prove that for $(\delta_1, \delta_0) \to (\delta_{r+1}, \delta_r)$, if one can find some $E \in \mathcal{F}_{SP}$ such that $p((\delta_1, \delta_0) \to (\delta_{r+1}, \delta_r)) > 0$, one can find some $E' \in \mathcal{F}_{SP}^{\perp}$ such that $c((\delta_0, \delta_1) \cdot x \oplus (\delta_r, \delta_{r+1}) \cdot E'(x)) \neq 0$.

Assume that $(\delta_1, \delta_0) \to (\delta_{r+1}, \delta_r)$ is a differential of $E \in \mathcal{F}_{SP}$. Then there exists a differential characteristic with positive probability:

$$(\delta_1, \delta_0) \to \cdots (\delta_{i+1}, \delta_i) \to \cdots \to (\delta_{r+1}, \delta_r),$$

where $\delta_i \in (\mathbb{F}_2^s)^t$. In this characteristic, the input difference of $S_i = (S_{i,1}, \ldots, S_{i,t})$ is $\delta_i = (\delta_{i,1}, \ldots, \delta_{i,t}) \in (\mathbb{F}_2^s)^t$, and let the output difference of $S_i$ be $\beta_i = (\beta_{i,1}, \ldots, \beta_{i,t}) \in (\mathbb{F}_2^s)^t$, then $\delta_{i+1} = \delta_{i-1} \oplus (\beta_i P)$.

Taking the following fact into consideration: for $(\delta_{i,j}, \beta_{i,j})$, where $\delta_{i,j} \neq 0$, there always exists an $s \times s$ binary matrix $M_{i,j}$ such that $\beta_{i,j} = \delta_{i,j} M_{i,j}^T$, then for $S_{i,j}(x) = xM_{i,j}$, $c(\beta_{i,j} \cdot x \oplus \delta_{i,j} \cdot S_{i,j}(x)) = 1$.

Now we construct an $r$-round cipher $E_r \in \mathcal{F}_{SP}^{\perp}$ such that $c((\delta_0, \delta_1) \cdot x \oplus (\delta_r, \delta_{r+1}) \cdot E_r(x)) \neq 0$. If $r = 1$, let $S_{1,j}(x) = xM_{1,j}$ for $\delta_{1,j} \neq 0$ and any linear transformation on $\mathbb{F}_2^s$ otherwise. Then all operations in $E_1 \in \mathcal{F}_{SP}^{\perp}$ are linear over $\mathbb{F}_2$, which implies that there exists a $2st \times 2st$ binary matrix $M_1$ such that $E_1(x) = xM_1$, and

$$c((\delta_0, \delta_1) \cdot x \oplus (\delta_1, \delta_2) \cdot E_1(x)) = 1.$$

Assume that we have constructed $E_{r-1}(x) = xM_{r-1}$ with $M_{r-1}$ being a $2st \times 2st$ binary matrix such that

$$c((\delta_0, \delta_1) \cdot x \oplus (\delta_{r-1}, \delta_r) \cdot E_{r-1}(x)) = 1,$$

and we can define $S_{r,j}(x)$ in the $r$-th round similarly, then $E_r(x) = xM_r$ for some $2st \times 2st$ binary matrix $M_r$, and

$$c((\delta_0, \delta_1) \cdot x \oplus (\delta_r, \delta_{r+1}) \cdot E_r(x)) = 1,$$

which ends our proof. □

Similarly, we can prove the following theorem:

**Theorem 2.** *$a \to b$ is an $r$-round impossible differential of $\mathcal{E}_{SP}$ if and only if it is an $r$-round zero correlation linear hull of $\mathcal{E}_{SP}^{\perp}$.*

Definition 2 implies that the "impossibility" of an impossible differential of a structure can be caused only by a differential $\delta_1 \to \delta_2$ where either $\delta_1 = 0$ or $\delta_2 = 0$ (but not both) over an invertible S-box, or by a differential $0 \to \delta_2$ over a non-invertible S-box. Otherwise, according to the proof of Theorem 1, we can always find an S-box such that $\delta_1 \to \delta_2$ is a possible differential. Therefore, we have the following corollary:

**Corollary 1.** *The method presented in [7] finds all impossible differentials of $\mathcal{F}_{SP}$ and $\mathcal{E}_{SP}$.*

As a matter of fact, this Corollary can be used in the provable security of block ciphers against impossible differential cryptanalysis, since with the help of this Corollary, the longest impossible differentials of a given structure could be given.

In case $P$ is invertible, according to the definition of equivalent structures given in [30], we have

$$\mathcal{F}_{P^T S} = \left((P^T)^{-1}, (P^T)^{-1}\right) \mathcal{F}_{SP^T} \left(P^T, P^T\right), \tag{1}$$

which indicates:

**Corollary 2.** *Let $\mathcal{F}_{SP}$ be a Feistel structure with $SP$-type round function, and let the primitive representation of the linear transformation be $P$. If $P$ is invertible, finding zero correlation linear hulls of $\mathcal{F}_{SP}$ is equivalent to finding impossible differentials of $\mathcal{F}_{SP^T}$.*

*Example 1.* **(8-Round Zero Correlation Linear Hull of Camellia Without $FL/FL^{-1}$)** Let Camellia* denote the cipher which is exactly the same as Camellia without $FL/FL^{-1}$ layer except that $P^T$ is used instead of $P$. Then we find that, for example:

$$((0,0,0,0,0,0,0,0),(0,0,0,0,a,0,0,0)) \to ((0,0,0,0,0,0,0,h),(0,0,0,0,0,0,0,0))$$

is an 8-round impossible differential of Camellia*, where $a$ and $h$ denote any non-zero values. Therefore, we could derive an 8-round zero correlation linear distinguisher of Camellia without $FL/FL^{-1}$ layer as shown below:

$$((a,a,0,0,a,0,a,a),(0,0,0,0,0,0,0,0)) \to ((0,0,0,0,0,0,0,0),(h,0,0,h,0,h,h,h)).$$

Furthermore, if $\mathcal{F}_{SP} = \mathcal{F}_{SP^T}$ and $\mathcal{E}_{SP} = \mathcal{E}_{S(P^{-1})^T}$, the following corollary holds:

**Corollary 3.** *For a Feistel structure $\mathcal{F}_{SP}$ with $SP$-type round function, if $P$ is invertible and $P = P^T$, there is a one-to-one correspondence between impossible differentials and zero correlation linear hulls.*

*For an SPN structures $\mathcal{E}_{SP}$, if $P^T P = E$, $a \to b$ is an impossible differential if and only if it is a zero correlation linear hull.*

*Example 2.* **(4-Round Zero Correlation Linear Hull of ARIA)** Since the linear layer $P$ of ARIA satisfies $P^T P = E$, any impossible differential of $\mathcal{E}^{\text{ARIA}}$ is automatically a zero correlation linear hull of $\mathcal{E}^{\text{ARIA}}$. Therefore, the impossible differentials of 4-round ARIA shown in [28] are also zero correlation linear hulls of 4-round ARIA.

**Notes.**

1. In the proof of Theorem 1, the $S$-boxes we constructed are not necessarily bijective. If we add the bijective condition, Theorem 1 still holds. Since for a bijective $S$-box, if the correlation is non-zero, $\delta_{1,j} \neq 0$ implies $\beta_{1,j} \neq 0$. Therefore, in Part(I) of the proof, we can further define $S_{1,j}$ as

$$S_{1,j}(x) = \begin{cases} x_{L,j} \oplus \delta_{1,j} & x = x_{L,j} \oplus \beta_{1,j}, \\ x_{L,j} \oplus \beta_{1,j} & x = x_{L,j} \oplus \delta_{1,j}, \\ \quad x & \text{others}, \end{cases}$$

and a similar definition can also be given to $S_{r,j}$. In this case, the $S$-boxes are invertible. Moreover, for a bijective $S$-box, if the differential probability is positive, $\delta_{i,j} \neq 0$ implies $\beta_{i,j} \neq 0$, thus in Part (II) of the proof, we can always find a non-singular binary matrix $M_{i,j}$ such that $\beta_{i,j} = \delta_{i,j} M_{i,j}^T$.

2. Theorem 1 and 2 show some links between impossible differentials and zero correlation linear hulls of a structure $\mathcal{E}$ and the corresponding dual structure $\mathcal{E}^{\perp}$. However, it doesn't mean that, for example, an impossible differential of a cipher $E \in \mathcal{E}$ indicates a zero correlation linear hull of another cipher $E' \in \mathcal{E}^{\perp}$. This follows from the difference between the definitions of an impossible differential and a zero correlation linear hull of a cipher and a structure, respectively.

## 4 Links between Integral and Zero Correlation Linear Cryptanalysis

Firstly, we will give two foundational statements that give links between integral cryptanalysis and zero correlation linear cryptanalysis:

**Lemma 1.** *Let $A$ be a subspace of $\mathbb{F}_2^n$, $A^{\perp} = \{x \in \mathbb{F}_2^n | a \cdot x = 0, a \in A\}$ be the dual space of $A$ and $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be a function on $\mathbb{F}_2^n$. For any $\lambda \in \mathbb{F}_2^n$, $T_\lambda : A^{\perp} \to \mathbb{F}_2^n$ is defined as $T_\lambda(x) = F(x \oplus \lambda)$, then for any $b \in \mathbb{F}_2^n$,*

$$\sum_{a \in A} (-1)^{a \cdot \lambda} c(a \cdot x \oplus b \cdot F(x)) = c(b \cdot T_\lambda(x)).$$

*Proof.*

$$\sum_{a \in A} (-1)^{a \cdot \lambda} c(a \cdot x \oplus b \cdot F(x)) = \sum_{a \in A} (-1)^{a \cdot \lambda} \frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} (-1)^{a \cdot x \oplus b \cdot F(x)}$$

$$= \frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} (-1)^{b \cdot F(x)} \sum_{a \in A} (-1)^{a \cdot (\lambda \oplus x)} = \frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} (-1)^{b \cdot F(x)} |A| \delta_{A^{\perp}}(\lambda \oplus x)$$

$$= \frac{1}{|A^{\perp}|} \sum_{y \in A^{\perp}} (-1)^{b \cdot T_\lambda(y)} = c(b \cdot T_\lambda(x)),$$

where $\delta_{A^{\perp}}(x) = \begin{cases} 1 & x \in A^{\perp} \\ 0 & x \notin A^{\perp} \end{cases}$. □

The second statement is as follows:

**Lemma 2.** *Let $A$ be a subspace of $\mathbb{F}_2^n$, $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$, and let $T_\lambda : A^{\perp} \to \mathbb{F}_2^n$ be defined as $T_\lambda(x) = F(x \oplus \lambda)$ where $\lambda \in \mathbb{F}_2^n$. Then for any $b \in \mathbb{F}_2^n$,*

$$\frac{1}{2^n} \sum_{\lambda \in \mathbb{F}_2^n} (-1)^{b \cdot F(\lambda)} c(b \cdot T_\lambda(x)) = \sum_{a \in A} c^2(a \cdot x \oplus b \cdot F(x)).$$

*Proof.*

$$\sum_{a \in A} c^2(a \cdot x \oplus b \cdot F(x)) = \sum_{a \in A} \frac{1}{2^n} \sum_{x \in \mathbb{F}_2^n} (-1)^{a \cdot x \oplus b \cdot F(x)} \frac{1}{2^n} \sum_{\lambda \in \mathbb{F}_2^n} (-1)^{a \cdot \lambda \oplus b \cdot F(\lambda)}$$

$$= \frac{1}{2^{2n}} \sum_{x \in \mathbb{F}_2^n} \sum_{\lambda \in \mathbb{F}_2^n} (-1)^{b \cdot F(x) \oplus b \cdot F(\lambda)} \sum_{a \in A} (-1)^{a \cdot x \oplus a \cdot \lambda}$$

$$= \frac{1}{2^{2n}} \sum_{x \in \mathbb{F}_2^n} \sum_{\lambda \in \mathbb{F}_2^n} (-1)^{b \cdot F(x) \oplus b \cdot F(\lambda)} |A| \delta_{A^{\perp}}(x \oplus \lambda)$$

Let $\theta = x \oplus \lambda$. Since $|A| \times |A^{\perp}| = 2^n$, we have

$$\begin{aligned}
&\frac{1}{2^{2n}} \sum_{x \in \mathbb{F}_2^n} \sum_{\lambda \in \mathbb{F}_2^n} (-1)^{b \cdot F(x) \oplus b \cdot F(\lambda)} |A| \delta_{A^{\perp}}(x \oplus \lambda) \\
&= \frac{|A|}{2^{2n}} \sum_{\theta \oplus \lambda \in \mathbb{F}_2^n} \sum_{\lambda \in \mathbb{F}_2^n} (-1)^{b \cdot F(\theta \oplus \lambda) \oplus b \cdot F(\lambda)} \delta_{A^{\perp}}(\theta) = \frac{1}{2^n |A^{\perp}|} \sum_{\lambda \in \mathbb{F}_2^n} (-1)^{b \cdot F(\lambda)} \sum_{\theta \oplus \lambda \in \mathbb{F}_2^n} (-1)^{b \cdot F(\theta \oplus \lambda)} \delta_{A^{\perp}}(\theta) \\
&= \frac{1}{2^n} \sum_{\lambda \in \mathbb{F}_2^n} (-1)^{b \cdot F(\lambda)} \frac{1}{|A^{\perp}|} \sum_{\theta \in A^{\perp}} (-1)^{b \cdot F(\theta \oplus \lambda)} = \frac{1}{2^n} \sum_{\lambda \in \mathbb{F}_2^n} (-1)^{b \cdot F(\lambda)} c(b \cdot T_{\lambda}(x)).
\end{aligned}$$

□

The authors of [17] concluded that an integral distinguisher implies a zero correlation linear hull. However, for general integral distinguishers, $c(b \cdot T_{\lambda}(x))$ may not necessarily be 0, hence the conclusion that integral unconditionally implies zero correlation linear hull in [17] is correct only under their definition of integral while it may not hold for general ones.

From Lemma 1, we can deduce the following:

**Corollary 4.** *Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^n$ be a function on $\mathbb{F}_2^n$, and let $A$ be a subspace of $\mathbb{F}_2^n$ and $b \in \mathbb{F}_2^n \setminus \{0\}$. Suppose that $A \to b$ is a zero correlation linear hull of $F$, then for any $\lambda \in \mathbb{F}_2^n$, $b \cdot F(x \oplus \lambda)$ is balanced on $A^{\perp}$.*

This Corollary states that if the input masks of a zero correlation linear hull form a subspace, then a zero correlation linear hull implies an integral distinguisher. Furthermore, the condition that input masks form a subspace could be removed, which leads to the following result:

**Theorem 3.** *A nontrivial zero correlation linear hull of a block cipher always implies the existence of an integral distinguisher.*

*Proof.* Assume that $A \to B$ is a non-trivial zero correlation linear hull of a block cipher $E$. Then we can choose $0 \neq a \in A, 0 \neq b \in B$, such that $\{0, a\} \to b$ is also a zero correlation linear hull of $E$.

Since $V = \{0, a\}$ forms a subspace on $\mathbb{F}_2$, according to Corollary 4, $b \cdot E(x)$ is balanced on $V^{\perp}$. This implies an integral distinguisher of $E$. □

Moreover, in the proof of Theorem 3, we can always assume that $0 \in A$. Then

1. If $A$ forms a subspace, an integral distinguisher can be constructed from $A \to b$;
2. If $A$ does not form a subspace, we can choose some $A_1 \subset A$ such that $A_1$ forms a subspace, then an integral distinguisher can be constructed from $A_1 \to b$.

It was stated in [17] that a zero correlation linear hull indicates the existence of an integral distinguisher under certain conditions, while Theorem 3 shows that these conditions can be removed. This results in a more applicable link between zero correlation linear cryptanalysis and integral cryptanalysis.

It can be seen that Theorem 3 also gives us a new approach to find integral distinguishers of block ciphers. More specifically, an $r$-round zero correlation linear hull can be used to construct an $r$-round integral distinguisher. Interestingly, Theorem 3 can also be used to convert known plaintext attacks to chosen plaintext attacks.

## 5 Links between Impossible Differential and Integral Cryptanalysis

According to the links given in the previous sections, we establish a link between impossible differential cryptanalysis and integral cryptanalysis:

**Theorem 4.** *Let $\mathcal{E} \in \{\mathcal{F}_{SP}, \mathcal{E}_{SP}\}$. Then impossible differential of $\mathcal{E}$ always implies the existence of an integral of $\mathcal{E}^{\perp}$.*

In case $\mathcal{E}^{\perp} = A_2\mathcal{E}A_1$ where $A_1$ and $A_2$ are linear transformations, we get the direct links between impossible differential and integral cryptanalysis:

**Corollary 5.** *Let $\mathcal{F}_{SP}$ be a Feistel structure with $SP$-type round function, and let the primitive representation of the linear transformation be $P$. If $P$ is invertible and there exists a permutation $\pi$ on $t$ elements such that for any $(x_0, \dots, x_{t-1}) \in \mathbb{F}_2^{s\times t}$,*

$$P(x_0, \dots, x_{t-1}) = \pi^{-1}P^T\pi(x_0, \dots, x_{t-1}),$$

*then for $\mathcal{F}_{SP}$, an impossible differential always implies the existence of an integral distinguisher.*

*Proof.* Let $\pi$ be a permutation on $(x_0, \dots, x_{t-1}) \in \mathbb{F}_2^{s\times t}$. Since

$$S_i \circ \pi(x_0, \dots, x_{t-1}) = \pi \circ S_i(x_0, \dots, x_{t-1}),$$

in the case that $P$ is invertible, we have

$$\mathcal{F}_{P^TS} = \left((\pi P^T)^{-1}, (\pi P^T)^{-1}\right)\mathcal{F}_{S(\pi^{-1}P^T\pi)}\left(\pi P^T, \pi P^T\right).$$

Therefore, the impossible differential of $\mathcal{F}_{SP}$ implies a zero correlation linear hull of $\mathcal{F}_{P^TS}$, which implies a zero correlation linear hull of $\mathcal{F}_{S(\pi^{-1}P^T\pi)} = \mathcal{F}_{SP}$, which in turn an integral distinguisher of $\mathcal{F}_{SP}$. □

*Example 3.* SNAKE(2) is a Feistel cipher proposed by Lee and Cha at JW-ISC'97, please refer to [31, 32] for details. According to [30], the round function of SNAKE(2) can be seen as an SP-type one with the primitive presentation of the matrix being defined as

$$P = \begin{pmatrix} E & E & E & E \\ E & 0 & E & E \\ E & 0 & 0 & E \\ E & 0 & 0 & 0 \end{pmatrix},$$

where $E$ and 0 are the identity and zero matrices of $\mathbb{F}_2^{8\times 8}$, respectively. Let

$$\pi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Then we have $P = \pi^{-1}P^T\pi$, therefore, an impossible differential of SNAKE(2), which is independent of the details of the $S$-boxes, always implies the existence of an integral distinguisher of SNAKE(2). Denote by $C_r$ the output of $r$-round SNAKE(2), and let $a \to b$ be an impossible differential of $r$-round SNAKE(2). Then $\left((\pi P^T)^{-1}, (\pi P^T)^{-1}\right)b \cdot C_r$ is balanced when the input takes all values in $((\pi P^T, \pi P^T)a)^{\perp}$.

**Corollary 6.** *Let $\mathcal{E}_{SP}$ be an SPN structure with the primitive representation of the linear transformation being $P$. If $P^TP = diag(Q_1, \dots, Q_t)$, where $Q_i \in \mathbb{F}_2^{s\times s}$, then for $\mathcal{E}_{SP}$, an impossible differential always implies the existence of an integral distinguisher.*

*Proof.* Firstly, according to Theorem 4, if $P^TP = E$, an impossible differential of $\mathcal{E}_{SP}$ always implies the existence of an integral.

Secondly, for the $S$-layer of $\mathcal{E}_{SP}$, if we substitute $S$ by applying $Q_i$ to the $i$-th $S$-box, according to definition 2, the structure stays identical. Since

$$P \circ (\text{diag}(Q_1, \dots, Q_t) \circ S) = (P \circ \text{diag}(Q_1, \dots, Q_t)) \circ S,$$

an SPN structure $\mathcal{E}_{SP}$ is equivalent to an SPN structure $\mathcal{E}_{S(P\circ\text{diag}(Q_1,\dots,Q_t))}$.

Based on the above two points, we can get the conclusion. □

To show applications of these links, we recall that, an $n \times n$ matrix $P$ is called *orthogonal* if and only if $P^T P = E$, where $E$ is the $n \times n$ identity matrix.

*Example 4.* We can check that, $SR$ and $M'$ used in PRINCE are orthogonal matrices, therefore

$$M^T M = (SR \circ M')^T (SR \circ M') = E,$$

where $E$ is the $64 \times 64$ identity matrix. So all the linear layers used in different rounds of PRINCE are orthogonal based on which we could conclude that any $r$-round impossible differential of PRINCE which is independent of the choices of the $S$-boxes implies the existence of an $r$-round integral distinguisher.

*Example 5.* Since the linear layer $P$ of ARIA is both symmetric and involutional, e.g. $P = P^{-1} = P^T$, any impossible differential of ARIA which is independent of the choices of $S$-boxes implies the existence of an integral distinguisher.

*Example 6.* We can check that $P$ used in PRESENT satisfies $P = (P^{-1})^T$, therefore, an impossible differential, which is independent of the details of the $S$-boxes, always leads to the existence of an integral distinguisher. In fact, since a permutation matrix $P$ is always orthogonal, we have the following Corollary:

**Corollary 7.** *For an SPN structure which adopts bit permutation as the diffusion layer, an $r$-round impossible differential always implies the existence of an $r$-round integral distinguisher.*

## 6 New Integral Distinguishers of Block Ciphers/Structures

### 6.1 New Integral Distinguishers for Feistel Structures

**Equivalence between $r$-Round Impossible Differential and Zero Correlation Linear Hull of Feistel Structures.** Let $\mathcal{E}_r$ be an $r$-round Feistel structure $\mathcal{F}_{SP}$. In the case that $P$ is the identity transformation, we get $\mathcal{E}_r^{\perp} = \sigma \mathcal{E}_r \sigma$, from which we can conclude that, $(a_L, a_R) \to (b_L, b_R)$ is an impossible differential of $\mathcal{E}_r$ if and only if $(a_R, a_L) \to (b_R, b_L)$ is a zero correlation linear hull of $\mathcal{E}_r$. If the round functions are not necessarily bijective, we obtain equivalence between the following two statements:

1. For any $a \neq 0$, $b \neq a$, $(0, a) \to (b, 0)$ is an impossible differential of $\mathcal{E}_3$;
2. For any $a \neq 0$, $b \neq a$, $(a, 0) \to (0, b)$ is a zero correlation linear hull of $\mathcal{E}_3$;

If the round functions are bijective, we obtain equivalence between 5-round impossible differentials and zero correlation linear hulls of Feistel structures:

(1) For any $a \neq 0$, $(0, a) \to (a, 0)$ is an impossible differential of $\mathcal{E}_5$;
(2) For any $a \neq 0$, $(a, 0) \to (0, a)$ is a zero correlation linear hull of $\mathcal{E}_5$. □

**New Integral Distinguishers of Feistel Structures.** So far the longest integral distinguisher known for a Feistel structure with bijective round functions counts 4 rounds, and the longest integral distinguisher for a Feistel structure with general round functions counts 2 rounds. We improve these distinguishers by 1 round using Theorem 3.

**Proposition 1.** *Let $\mathcal{E}_r$ be an $r$-round Feistel structure defined on $\mathbb{F}_2^{2n}$. Then*

1. *If the $F_i$'s are bijective, then for any $c \in \mathbb{F}_2^n$, $c \neq 0$, $c \cdot R_5$ is balanced on $\{(0,0), (c,0)\}^{\perp}$ with respect to $\mathcal{E}_5$.*
2. *If the $F_i$'s are not necessarily bijective, then let $\{\alpha_0, \ldots, \alpha_{n-1}\}$ be a base of $\mathbb{F}_2^n$ over $\mathbb{F}_2$. Then $\alpha_{n-1} \cdot R_3$ is balanced on $\{(0, \sum_{i=0}^{n-2} c_i \alpha_i) | c_i \in \mathbb{F}_2\}^{\perp}$ with respect to $\mathcal{E}_3$.*

As a matter of fact, for any $c \in \mathbb{F}_2^n$, $c \neq 0$, $(c, 0) \to (0, c)$ is a zero correlation linear hull of $\mathcal{E}_5$. Thus according to Theorem 3, we can construct an integral distinguisher of $\mathcal{E}_5$, i.e., let $(L_0, R_0)$ take all values in $\{(0, 0), (c, 0)\}^\perp$, then $c \cdot R_5$ is balanced.

Specifically, let $c = (1, 1, \ldots, 1) \in \mathbb{F}_2^n$. Then we have

$$\{(0,0),(c,0)\}^\perp = \{((x_1,\ldots,x_n),(x_{n+1},\ldots,x_{2n}))|x_i \in \mathbb{F}_2, \sum_{t=}^{n} x_t = 0\}.$$

Let $R_5 = (R_{5,1}, \ldots, R_{5,n})$. Then we can derive that $\sum_{i=1}^n R_{5,i}$ is balanced on $\{(0, 0), (c, 0)\}^\perp$.

### 6.2 24-Round Integral Distinguisher of CAST-256

The block cipher CAST-256 [33] was proposed as a first-round AES candidate, and we refer to [33] for details. Firstly, we recall the following zero correlation linear property given in [17].

*Property 1.* $(0, 0, 0, L_1) \to (0, 0, 0, L_2)$ is a zero correlation linear hull of the 24-round CAST-256 (from the 13-th round to the 36-th round of CAST-256), where $L_1 \neq 0$, $L_2 \neq 0$ and $L_1 \neq L_2$.

Let $L_1^* = \{(l_1, l_2, \ldots, l_{31}, 0)|l_i \in \mathbb{F}_2\}$ and $L_2 = (0, \ldots, 0, 1)$. Then we obtain a zero correlation linear hull $(0, 0, 0, L_1^*) \to (0, 0, 0, L_2)$ for the 24-round CAST-256. According to Theorem 3, we can get the following result:

**Proposition 2.** *Let $V = \{(x_1, x_2, x_3, 0^{31}y)|x_i \in \mathbb{F}_2^{32}, y \in \mathbb{F}_2\}$. If the input takes all values in $V$, and let the output of the 24-round be $(C_0, C_1, C_2, C_3) \in \mathbb{F}_2^{32\times 4}$ (from the 13-th round to 36-th round). Then $(0, \ldots, 0, 1)\cdot C_3$ is balanced.*

Based on this integral distinguisher, we present a key recovery attack on 28-round CAST-256 which is the best known attack on CAST-256 in the non-weak key model. The details of the attack are listed in Appendix A. Table 1 gives the summary of attacks on CAST-256 in the non-weak key model.

**Table 1.** Summary of Attacks on CAST-256 in the Non-Weak Key Model

| Type of Attack | Attacked Rounds | Key Size | Data Complexity | Time Complexity | Memory Complexity | Success Probability |
|---|---|---|---|---|---|---|
| Boomerang [34] | 16 | all | $2^{49.3}$ ACPC | - | - | - |
| Linear [35] | 24 | 192/256 | $2^{124.1}$ KP | $2^{156.52}$ Enc | - | - |
| Multidim. ZC [17] | 28 | 256 | $2^{98.8}$ KP | $2^{246.9}$ Enc | $2^{103.8}$ B [1] | 0.846 |
| Integral (Sec.5.2) | 28 | 256 | $2^{97}$ CP | $2^{239.19}$ Enc | $2^{102}$ B | 1 |

CP: Chosen plaintexts, KP: Known plaintexts,
ACPC: Adaptive chosen plaintexts and ciphertexts,
Enc: Encryptions, B: Bytes, -: Not given in the related paper.

### 6.3 12-Round Integral Distinguisher of SMS4

The SMS4 block cipher is designed by the Chinese government as part of their WAPI standard for wireless networks [36]. Up to date, the longest known integral distinguisher of SMS4 covers 10 rounds [37]. The details of SMS4 and the proof of the following Propositions are listed in Appendix B.

**Proposition 3.** *Let $V = \{v \in (\mathbb{F}_2^8)^4|HW(\mathcal{L}^T v) = 1\}$, where $HW(x_1, x_2, x_3, x_4) = \#\{x_i \neq 0, i = 1, 2, 3, 4\}$. For any $d \in V$, $(0, 0, 0, d) \to (d, 0, 0, 0)$ is a 12-round zero correlation linear hull of SMS4.*

**Proposition 4.** *Let* $V = \{v \in (\mathbb{F}_2^8)^4 | HW(\mathcal{L}^T v) = 1\}$, $V_d = \{w \in (\mathbb{F}_2^{32})^4 | (0,0,0,d) \cdot w = 0\}$, *and let* $(c_0, c_1, c_2, c_3)$ *be the output of* 12*-round SMS4. Then for any* $d \in V$, *when the input takes all possible values in* $V_d$, *we have*

$$\#\{d \cdot c_0 = 0\} = \#\{d \cdot c_0 = 1\}.$$

Note that most of the known integral distinguishers are independent of the choices of the $S$-boxes. However, the integral distinguisher presented above is highly related with the $S$-boxes, since for different $S$-boxes, we would find different zero correlation linear hulls which lead to different integral distinguishers of SMS4.

### 6.4 8-Round Integral Distinguisher of Camellia without $FL/FL^{-1}$ Layer

In [39], by using the division property, the author proposed a 6-round integral distinguisher of Camellia without $FL/FL^{-1}$ layers as the best known integral distinguisher which could be built without knowing the details of the $S$-box. Based on the 8-round zero correlation linear hull presented in Example 1, the integral distinguisher which is independent with the choices of $S$-box could be improved from 6-round to 8-round:

**Proposition 5.** *Let* $V$ *be defined as*

$$V = \{((x_1, \ldots, x_8), (x_9, \ldots, x_{16})) | x_1 \oplus x_2 \oplus x_5 \oplus x_7 \oplus x_8 = 0, x_i \in \mathbb{F}_2^8\}.$$

*For any* $h \in \mathbb{F}_2^8$, $h \neq 0$, $(h, 0, 0, h, 0, h, h, h) \cdot R_{i+8}$ *is balanced on* $V$ *with respect to* 8*-round Camellia without* $FL/FL^{-1}$ *layer.*

## 7 Conclusion

In this paper, we have investigated the link between impossible differential and integral cryptanalysis. To do this, we have introduced the concept of *structure* $\mathcal{E}$ and *dual structure* $\mathcal{E}^{\perp}$ and established the link in the following steps:

- We derived the relation between impossible differential of $\mathcal{E}$ and zero correlation linear hull of $\mathcal{E}^{\perp}$. We have shown that for a Feistel structure $\mathcal{F}_{SP}$ with $SP$-type round functions where $P$ is invertible, constructing a zero correlation linear hull of $\mathcal{F}_{SP}$ is equivalent to constructing an impossible differential of $\mathcal{F}_{SP^T}$, which is the same structure as $\mathcal{F}_{SP}$ with $P^T$ instead of $P$. For an SPN structure $\mathcal{E}_{SP}$, constructing a zero correlation linear hull of $\mathcal{E}_{SP}$ is equivalent to constructing an impossible differential of $\mathcal{E}_{S(P^{-1})^T}$, which is the same structure as $\mathcal{E}_{SP}$ with $(P^{-1})^T$ instead of $P$.
- We presented the relation between zero correlation linear hull and integral distinguisher of block ciphers. As proven in Sec.4, a zero correlation linear hull always implies the existence of an integral distinguisher, while such statement only holds under certain conditions in [17]. Meanwhile, we have observed that the statement "*integral unconditionally implies zero correlation linear hull*" in [17] is correct only under the definition that integral property is a balanced vectorial boolean function, while it does not hold for the general case (i.e., integral defined in [3] is a zero-sum property).
- We built the link between impossible differential of $\mathcal{E}$ and integral distinguisher of $\mathcal{E}^{\perp}$. We have demonstrated that an $r$-round impossible differential of $\mathcal{E}$ always leads to an $r$-round integral distinguisher of $\mathcal{E}^{\perp}$. In the case that $\mathcal{E}$ and $\mathcal{E}^{\perp}$ are linearly equivalent, we obtained some direct links between impossible differential and integral distinguisher of $\mathcal{E}$. Specifically, an $r$-round impossible differential of an SPN structure, which adopts bit permutation as the linear layer, always indicates the existence of an $r$-round integral distinguisher.

The results and links presented in this paper not only allow to achieve a better understanding and classifying of impossible differential cryptanalysis, integral cryptanalysis and zero correlation linear cryptanalysis, but also provide some new insights with respect to these cryptanalytic approaches as shown below:

[1] The original memory complexity of the attack in [17] ($2^{68}$ bytes) was underestimated since it did not take into account the memory requirement for storing the $2^{98.8}$ plaintext-ciphertext pairs.

- The automatic search tool presented by Wu and Wang in Indocrypt 2012 finds all impossible differentials of both Feistel structures with $SP$-type round functions and SPN structures, which is useful in provable security of block ciphers against impossible differential cryptanalysis.
- Our statement "*zero correlation linear hull always implies the existence of an integral distinguisher*" provides a novel way for constructing integral distinguisher of block ciphers and converting known plaintext attacks to chosen plaintext attacks. With this observation, we have improved the integral distinguishers of Feistel structures by 1 round, built a 24-round integral distinguisher of CAST-256, and proposed a 12-round integral distinguisher of SMS4 which is 2-round longer than previously best known ones and an 8-round integral distinguisher of Camellia without $FL/FL^{-1}$ laysser which is 2-round longer than the best known ones which are independent with the choices of the $S$-box. These distinguishers could not be obtained by the previously known methods for constructing integral distinguishers or by using the link given in [17]. Moreover, we have presented the best known key recovery attack on CAST-256 in non-weak key model to show that the new links can also be used to improve cryptanalytic results of some concrete ciphers.

By using the matrix representation given in [38], the concept of dual structure can be extended to generalized Feistel structures, and we can get similar results for these structures. Furthermore, we have focused on the links among the distinguishers used in impossible differential, integral and zero correlation linear cryptanalysis since distinguishers are the essential points in the evaluation of security margins of a block cipher against various cryptanalytic tools, and our results can be helpful in designing a block cipher from this point of view.

## References

1. L.R. Knudsen. DEAL — A 128-bit Block Cipher. Department of Informatics, University of Bergen, Norway. Technical report, 1998.
2. E. Biham, A. Biryukov, A. Shamir. Cryptanalysis of Skipjack Reduced to 31 Rounds Using Impossible Differentials. EUROCRYPT 1999, LNCS 1592, pp. 12–23, Springer-Verlag, 1999.
3. L.R. Knudsen, D. Wagner. Integral Cryptanalysis. FSE 2002, LNCS 2365, pp. 112–127, Springer–Verlag, 2002.
4. A. Bogdanov, V. Rijmen. Linear Hulls with Correlation Zero and Linear Cryptanalysis of Block Ciphers. Designs, Codes and Cryptography, 70(3), pp. 369–383, 2014.
5. J. Kim, S. Hong, J. Sung, S. Lee, J. Lim. Impossible Differential Cryptanalysis for Block Cipher Structures. Indocrypt 2003, LNCS 2904, pp. 82–96, 2003.
6. Y. Luo, X. Lai, Z. Wu, G. Gong. A Unified Method for Finding Impossible Differentials of Block Cipher Structures. Information Sciences, Volume 263, 1 April 2014, Pages 211–220.
7. S. Wu, M. Wang. Automatic Search of Truncated Impossible Differentials for Word-Oriented Block Ciphers. Indocrypt 2012, LNCS 7668, pp. 283–302, 2012.
8. J. Daemen, L. R. Knudsen, V. Rijmen. The Block Cipher Square. Fast Software Encryption 1997, LNCS 1267, pp. 149–165, Springer–Verlag, 1997.
9. S. Lucks. The Saturation Attack — A Bait for Twofish. Fast Software Encryption 2001, LNCS 2355, pp. 1–15, Springer–Verlag, 2002.
10. A. Biryukov, A. Shamir. Structural Cryptanalysis of SASAS. EUROCRYPT 2001, LNCS 2045, pp. 394–405, Springer–Verlag, 2001.
11. X. Lai. Higher Order Derivatives and Differential Cryptanalysis. Communications and Cryptography: Two Sides of One Tapestry, 227 (1994)
12. L.R. Knudsen. Truncated and Higher Order Differentials. Fast Software Encryption 1994, LNCS 1008, pp. 196–211. Springer, Heidelberg (1995)
13. S. Picek, L. Batina, D. Jakobović, B. Ege, M. Golub. S-box, SET, Match: A Toolbox for S-box Analysis. WISTP 2014, LNCS 8501, pp. 140–149, 2014.
14. F. Chabaud, S. Vaudenay. Links Between Differential and Linear Cryptoanalysis. EUROCRYPT 1994, LNCS 950, pp. 356–365, Springer-Verlag, 1995.
15. B. Sun, R. Li, L. Qu, C. Li. SQUARE Attack on Block Ciphers with Low Algebraic Degree. Science China Information Sciences 53(10), pp. 1988–1995, 2010.

16. G. Leander. On Linear Hulls, Statistical Saturation Attacks, PRESENT and a Cryptanalysis of PUFFIN. EUROCRYPT 2011, LNCS 6632, pp. 303–322, Springer-Verlag, 2011.
17. A. Bogdanov, G. Leander, K. Nyberg and M. Wang. Integral and Multidimensional Linear Distinguishers with Correlation Zero. ASIACRYPT 2012, LNCS 7658, pp. 244–261, Springer–Verlag, 2012.
18. C. Blondeau and K. Nyberg. New Links Between Differential and Linear Cryptanalysis. EUROCRYPT 2013, LNCS 7881, pp. 388–404, Springer–Verlag, 2013.
19. C. Blondeau, G. Leander, K. Nyberg. Differential-Linear Cryptanalysis Revisited. FSE 2014, to appear.
20. C. Blondeau, A. Bogdanov, M. Wang. On the (In)Equivalence of Impossible Differential and zero correlation Distinguishers for Feistel- and Skipjack-type Ciphers. ACNS 2014, LNCS 8479, pp. 271–288, 2014.
21. C. Blondeau, K. Nyberg. Links Between Truncated Differential and Multidimensional Linear Properties of Block Ciphers and Underlying Attack Complexities. EUROCRYPT 2014, LNCS 8441, pp. 165–182, 2014.
22. C. Carlet. Boolean Functions for Cryptography and Error Correcting Codes. Cambridge University Press, 2006.
23. A. Bogdanov, L. R. Knudsen, G. Leander, C. Paar, A. Poschmann, M. J. B. Robshaw, Y. Seurin, C. Vikkelsoe: PRESENT: An Ultra-Lightweight Block Cipher. CHES 2007, LNCS 4727, pp 450–466, 2007.
24. J. Borghoff, A. Canteaut, T. Güneysu, E. B. Kavun, M. Knezevic, L. R. Knudsen, G. Leander, V. Nikov, C. Paar, C. Rechberger, P. Rombouts, S. Thomsen, T. Yalçın. PRINCE — A Low-Latency Block Cipher for Pervasive Computing Applications - Extended Abstract. ASIACRYPT 2012. LNCS 7658, pp. 208–225, 2012.
25. K. Aoki, T. Ichikawa, M. Kanda, M. Matsui, S. Moriai, J. Nakajima and T. Tokita. Camellia: A 128–Bit Block Cipher Suitable for Multiple Platforms - Design and Analysis. SAC 2000, LNCS 2012, pp. 39–56, Springer–Verlag, 2000.
26. D. Kwon, J. Kim, S. Park, S.H. Sung etc. New Block Cipher: ARIA. ICISC 2003, LNCS 2971, pp.432–445, Springer-Verlag 2004.
27. H. Mala, M. Dakhilalian, V. Rijmen, M. Modarres-Hashemi. Improved Impossible Differential Cryptanalysis of 7-Round AES-128. INDOCRYPT 2010, LNCS 6498, pp. 282–291, Springer–Verlag, 2010.
28. W. Wu, W. Zhang, D. Feng. Impossible Differential Cryptanalysis of Round-Reduced ARIA and Camellia. Journal of Computer Science and Technology, 22(3), pp. 449–456, 2007.
29. A. Bogdanov, H. Geng, M. Wang, L. Wen, B. Collard. Zero Correlation Linear Cryptanalysis with FFT and Improved Attacks on ISO Standards Camellia and CLEFIA. SAC 2013, LNCS 8282, pp. 306–323.
30. L. Duo, C. Li, K. Feng. New Observation on Camellia. SAC 2005, LNCS 3897, pp. 51–64, Springer–Verlag, 2006.
31. C. Lee, Y. Cha. The Block Cipher: SNAKE with Provable Resistance against DC and LC Attacks. In Proceedings of 1997 Korea-Japan Joint Workshop on Information Security and Cryptology (JW–ISC'97), pp. 3–17, 1997.
32. S. Moriai, T. Shimoyama, T. Kaneko. Interpolation Attacks of the Block Cipher: SNAKE. FSE 1999, LNCS 1636, pp. 275–289, 1999.
33. First AES Candidate Conference. http://csrc.nist.gov/archive/aes/round1/conf1/aes1conf.htm.
34. D. Wagner. The Boomerang Attack. FSE 1999, LNCS 1636, pp. 156–170, Springer–Verlag, 1999.
35. M. Wang, X. Wang and C. Hu. New Linear Cryptanalytic Results of Reduced-Round of CAST-128 and CAST-256. SAC 2008, LNCS 5381, pp. 429–441. Springer–Verlag, 2009.
36. Specification of SMS4, Block Cipher for WLAN Products C SMS4 (in Chinese),
http://www.oscca.gov.cn/UpFile/200621016423197990.pdf
37. W. Zhang, B. Su, W. Wu, D. Feng. C. Wu. Extending Higher-Order Integral: An Efficient Unified Algorithm of Constructing Integral Distinguishers for Block Ciphers. ACNS 2012, LNCS 7341, pp. 117–134, Springer-Verlag, 2012.
38. T.P. Berger, M. Minier, G. Thomas. Extended Generalized Feistel Networks Using Matrix Representation. SAC 2013, LNCS 8282, pp. 289–305, 2014.
39. Yosuke Todo. Structural Evaluation by Generalized Integral Property. To appear in EUROCRYPT 2015.
http://eprint.iacr.org/2015/090

## Appendix A

The block cipher CAST-256 [33] was proposed as a first-round AES candidate. It is a 128-bit block cipher which adopts a generalized Feistel structure. CAST-256 supports variable key sizes, i.e., 128, 192 or 256-bit key size, and the number of rounds for all variants is 48.

Two types of round functions are used in CAST-256, i.e., the *forward quad-round* $Q(\cdot)$ and the *reverse quad-round* $\bar{Q}(\cdot)$. Let $I_i = (I_{i,1}, I_{i,2}, I_{i,3}, I_{i,4})$ denote the input of the $i$-th round of CAST-256, where $i \equiv 1$

(mod 4) and $I_{i,j} \in \mathbb{F}_2^{32}$, $1 \leq j \leq 4$. Then the forward quad-round $Q(I_i)$ is defined as consecutive application of 4 rounds as follows (See Fig.4):

$$I_{i+4,3} = I_{i,3} \oplus F_1(I_{i,4}, K_{R_1}^{(i)}, K_{M_1}^{(i)}),\ I_{i+4,2} = I_{i,2} \oplus F_2(I_{i+4,3}, K_{R_2}^{(i)}, K_{M_2}^{(i)}),$$

$$I_{i+4,1} = I_{i,1} \oplus F_3(I_{i+4,2}, K_{R_3}^{(i)}, K_{M_3}^{(i)}),\ I_{i+4,4} = I_{i,4} \oplus F_1(I_{i+4,1}, K_{R_2}^{(i)}, K_{M_2}^{(i)}).$$

Similarly, the reverse quad-round $\bar{Q}(I_i)$ is defined as:

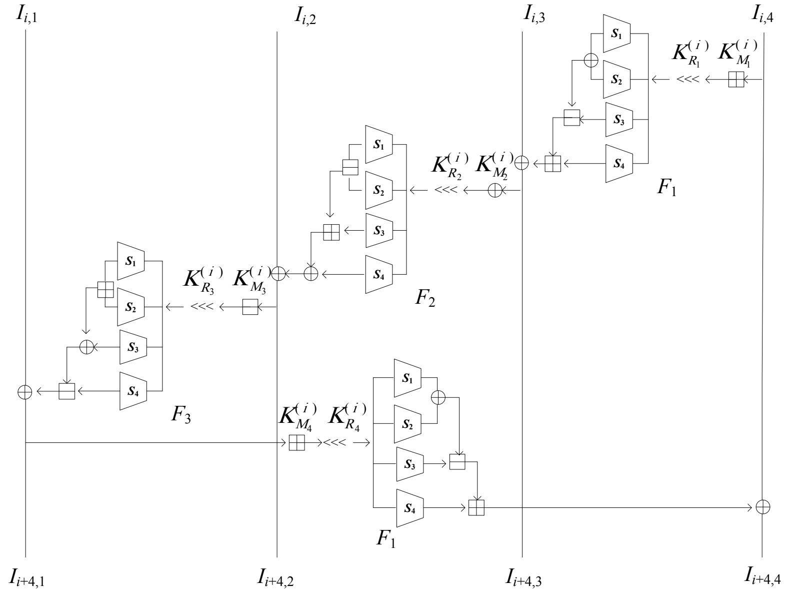

**Fig. 4.** Forward quad-round of CAST-256

$$I_{i+4,4} = I_{i,4} \oplus F_1(I_{i,1}, K_{R_4}^{(i)}, K_{M_4}^{(i)}),\ I_{i+4,1} = I_{i,1} \oplus F_3(I_{i,2}, K_{R_3}^{(i)}, K_{M_3}^{(i)}),$$

$$I_{i+4,2} = I_{i,2} \oplus F_2(I_{i,3}, K_{R_2}^{(i)}, K_{M_2}^{(i)}),\ I_{i+4,3} = I_{i,3} \oplus F_1(I_{i,4}, K_{R_1}^{(i)}, K_{M_1}^{(i)}).$$

where $K_R^{(i)} = \{K_{R_1}^{(i)}, K_{R_2}^{(i)}, K_{R_3}^{(i)}, K_{R_4}^{(i)}\} \in (\mathbb{F}_2^5)^4$ is the set of rotation keys for the $i$-th quad-round, and $K_M^{(i)} = \{K_{M_1}^{(i)}, K_{M_2}^{(i)}, K_{M_3}^{(i)}, K_{M_4}^{(i)}\} \in (\mathbb{F}_2^{32})^4$ is the set of masking keys for the $i$-th quad-round.

The encryption procedure for CAST-256 consists of 6 forward quad-rounds followed by 6 reverse quad-rounds, counting 48 rounds in total. Please refer to [33] for the details.

### Attacking 28-round CAST-256

With the help of the integral distinguisher presented in Proposition 2, we can mount an attack on the 28 rounds of CAST-256 (from the 13-th round to the 40-th round of CAST-256 and denoted as $E$) and recover the 148 subkey bits used in the last 4 rounds of $E$. The attack works as below.

**Step 1.** Collect a structure of plaintexts with the form $(x_1, x_2, x_3, 0^{31}y)$, where $x_i \in \mathbb{F}_2^{32}, y \in \mathbb{F}_2$ can take all possible values. Thus this structure consists of $2^{97}$ plaintexts. Ask for the encryption of this structure so as to get the corresponding ciphertexts.

**Step 2.** Initialize a counter $T$. Guess the value of the 148 subkey bits applied in the last 4 rounds of $E$, then do the following:

- For each of the $2^{97}$ ciphertexts obtained above, do the partial decryption to derive the value of $I_{37,4}$.
- Calculate the parity $(0,\dots,0,1)\cdot I_{37,4}$. If the parity is 0, increase $T$ by 1, and decrease $T$ by 1 otherwise.
- Check whether $T$ is equal to 0 or not. If yes, keep the guessed value as the correct subkey information, and discard it otherwise.

For a wrong key guess, the probability that it can pass the test in Step 2 is about $C_{2^{97}}^{2^{96}}/2^{2^{97}}$. Since there are $2^{148}$ possible values of the 148 subkey bits used in the last 4 rounds of $E$ and $2^{148}\times C_{2^{97}}^{2^{96}}/2^{2^{97}} < 2^{-256}$, it can be expected that only the correct subkey information will be kept after Step 2, and the success probability of this attack is approximately 1.

The data complexity of this attack is $2^{97}$ chosen plaintexts. The time complexity of this attack is mainly dominated by the partial decryptions in Step 2, thus it can be measured as $2^{97}\times 2^{148}\times 4/28 \approx 2^{239.19}$ 28-round CAST-256 encryptions. Moreover, the memory complexity of this attack is primarily owing to keeping $2^{97}$ plaintext-ciphertext pairs, accordingly, it can be estimated as $2^{97}\times 256/8 = 2^{102}$ bytes.

## Appendix B

SMS4 takes a 128-bit plaintext $P=(P_0,P_1,P_2,P_3)\in(\mathbb{F}_2^{32})^4$ as input, and 128-bit secret key $K$ which is used to derive the roundkeys used in different round.

Let $X_i=(X_{i,0},X_{i,1},X_{i,2},X_{i,3})\in(\mathbb{F}_2^{32})^4$, and set $X_0=P$, then, see Fig.5, for $i=1,2,\dots,32$,

$$\begin{cases} X_{i,0}=X_{i-1,1}\\ X_{i,1}=X_{i-1,2}\\ X_{i,2}=X_{i-1,3}\\ X_{i,3}=X_{i-1,0}\oplus F(X_{i-1,1}\oplus X_{i-1,2}\oplus X_{i-1,3}\oplus k_i)\end{cases}$$

where $k_i$ is the roundkey and $F:\mathbb{F}_2^{32}\to\mathbb{F}_2^{32}$ is defined as following:

Assume $M=(M_0,M_1,M_2,M_3)\in(\mathbb{F}_2^{8})^4$, $S$ be an $8\times 8$ bijective S-box and $M\lll_i$ be left rotation of $M$ by $i$ bits. Let

$$\mathcal{S}(M)=(S(M_0),S(M_1),S(M_2),S(M_3)),$$

and

$$\mathcal{L}(M)=M\oplus(M\lll_2)\oplus(M\lll_{10})\oplus(M\lll_{18})\oplus(M\lll_{24})$$

Then $F(M)=\mathcal{L}\circ\mathcal{S}(M)$.

To make the decryption identical to the encryption, there is a permutation after we get $X_{32}$, however, since it does not influence the properties introduced in this paper, details are omitted.

In the following, as shown in Fig.5, we will first construct a 12-round zero correlation linear hull of SMS4 which will be used to construct integral distinguishers.

Let the input mask of SMS4 be $(0,0,0,d)$, $d\neq 0$. Then to construct a characteristic with non-zero correlation, the output masks of the first, second and third round should be $(0,0,d,0)$, $(0,d,0,0)$ and $(d,0,0,0)$, respectively. For any $0\neq\alpha\in\mathbb{F}_2^{32}$, let

$$\mathcal{V}(\alpha)=\{\beta\in\mathbb{F}_2^{32}|c(\beta\cdot x\oplus\alpha\cdot F(x))\neq 0\},$$

then the output mask of the forth and fifth rounds are $(u_1,u_1,u_1,d)$, $u_1\in\mathcal{V}(d)$ and $(u_1\oplus v_1,u_1\oplus v_1,d\oplus v_1,u_1)$, $v_1\in\mathcal{V}(u_1)$, respectively. Therefore, the output mask of the sixth round is

$$(u_1\oplus v_1\oplus w_1,d\oplus v_1\oplus w_1,u_1\oplus w_1,u_1\oplus v_1)\qquad(2)$$

where $w_1\in\mathcal{V}(u_1\oplus v_1)$.

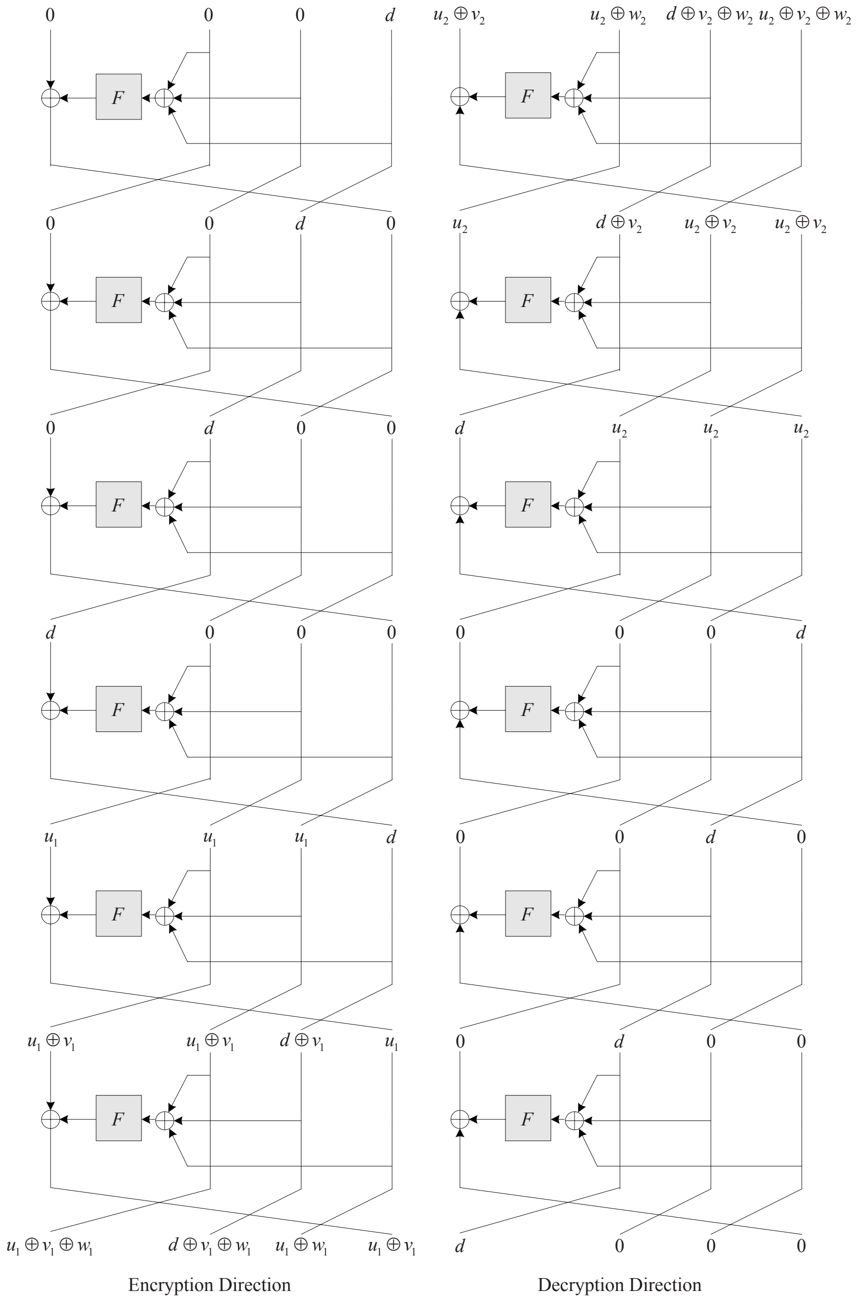
0 0 0 d
0 0 d 0
0 d 0 0
d 0 0 0
u1 u1 u1 d
u1 ⊕ v1 u1 ⊕ v1 d ⊕ v1 u1
u1 ⊕ v1 ⊕ w1 d ⊕ v1 ⊕ w1 u1 ⊕ w1 u1 ⊕ v1
u2 ⊕ v2 u2 ⊕ w2 d ⊕ v2 ⊕ w2 u2 ⊕ v2 ⊕ w2
u2 d ⊕ v2 u2 ⊕ v2 u2 ⊕ v2
d u2 u2 u2
0 0 0 d
0 0 d 0
0 d 0 0
d 0 0 0
F
Encryption Direction
Decryption Direction

**Fig. 5.** 12-Round Zero Correlation Linear Hull of SMS4

Let the output mask of the twelfth round be $(d, 0, 0, 0)$. Then to construct a characteristic with nonzero correlation, the output masks of the eleventh, tenth and ninth round are $(0, d, 0, 0)$, $(0, 0, d, 0)$ and $(0, 0, 0, d)$, respectively. The output mask of the eighth and seventh round are $(d, u_2, u_2, u_2)$, $u_2 \in \mathcal{V}(d)$, $(u_2, d \oplus v_2, u_2 \oplus v_2, u_2 \oplus v_2)$, $v_2 \in \mathcal{V}(u_2)$, respectively. Finally, we get the output mask of the sixth round

$$(u_2 \oplus v_2, u_2 \oplus w_2, d \oplus v_2 \oplus w_2, u_2 \oplus v_2 \oplus w_2) \tag{3}$$

where $w_2 \in \mathcal{V}(u_2 \oplus v_2)$. Therefore, we have

$$\begin{cases} u_1 \oplus v_1 \oplus w_1 = u_2 \oplus v_2 \\ d \oplus v_1 \oplus w_1 = u_2 \oplus w_2 \\ u_1 \oplus w_1 = d \oplus v_2 \oplus w_2 \\ u_1 \oplus v_1 = u_2 \oplus v_2 \oplus w_2 \end{cases}$$

which implies $w_1 = w_2 = 0$. Taking $w_1 \in \mathcal{V}(u_1 \oplus v_1)$ and $w_2 \in \mathcal{V}(u_2 \oplus v_2)$ into consideration, we have $u_1 = v_1$, $u_2 = v_2$ and $u_1 \oplus v_2 = u_2 \oplus v_1 = d$. Therefore, for $0 \neq d \in \mathbb{F}_2^{32}$, if we could not find $u_1$, such that $u_1 \in \mathcal{V}(d)$, $u_1 \oplus d \in \mathcal{V}(d)$, $u_1 \in \mathcal{V}(u_1)$ and $u_1 \oplus d \in \mathcal{V}(u_1 \oplus d)$, then $(0, 0, 0, d) \rightarrow (d, 0, 0, 0)$ is a zero correlation linear hull of12-round SMS4.

By exhaustive search, we have found many $d$'s such that $(0, 0, 0, d) \rightarrow (d, 0, 0, 0)$ is a 12-round zero correlation linear hull of SMS4 which is summarized in Proposition 3.