Periodic Table of Cryptanalysis: Geometric Approach with Different Bases

Kai Hu[1,5,6], Chi Zhang[2], Chengcheng Chang[1,5,6], Jiashu Zhang, Meiqin Wang[3,1,5,6] and Thomas Peyrin[4]

[1] School of Cyber Science and Technology, Shandong University, Qingdao, Shandong, China.
kai.hu@sdu.edu.cn, chengcheng.chang@mail.sdu.edu.cn, joshua020827@163.com,
[2] School of Mathematics, Shandong University, Jinan, Shandong 250100, China.
zhangchi010301@gmail.com
[3] Quancheng Laboratory, Jinan 250103, China.
mqwang@sdu.edu.cn
[4] Nanyang Technological University, Singapore.
thomas.peyrin@ntu.edu.sg
[5] Key Laboratory of Cryptologic Technology and Information Security, Ministry of Education, Shandong University, Jinan, China.
[6] State Key Laboratory of Cryptography and Digital Economy Security, Shandong University, Qingdao, 266237, China.

Abstract. In the past three decades, we have witnessed the creation of various cryptanalytic attacks. However, relatively little research has been done on their potential underlying connections. The geometric approach, developed by Beyne in 2021, shows that a cipher can be viewed as a linear operation when we treat its input and output as points in an induced *free vector space*. By performing a change of basis for the input and output spaces, one can obtain various transition matrices. Linear, differential, and (ultrametic) integral attacks have been well reinterpreted by Beyne's theory in a unified way.

Thus far, the geometric approach always uses the same basis for the input and output spaces. We observe here that this restriction is unnecessary and allowing different bases makes the geometric approach more flexible and able to interpret/predict more attack types. Given some set of bases for the input and output spaces, a family of basis-based attacks is defined by combining them, and all attacks in this family can be studied in a unified automatic search method.

We revisit three kinds of bases from previous geometric approach papers and extend them to four extra ones by introducing new rules when generating new bases. With the final seven bases, we can obtain 7^{2d} different basis-based attacks in the d-th order spaces, where the *order* is defined as the number of messages used in one sample during the attack.

We then provide four examples of applications of this new framework. First, we show that by choosing a better pair of bases, Beyne and Verbauwhede's ultrametric integral cryptanalysis can be interpreted as a single element of a transition matrix rather than as a linear combination of elements. This unifies the ultrametric integral cryptanalysis

with the previous linear and quasi-differential attacks. Second, we revisit the multiple-of-n property with our refined geometric approach and exhibit new multiple-of-n distinguishers that can reach more rounds of the SKINNY-64 cipher than the state-of-the-art. Third, we study the multiple-of-n property for the first-order case, which is similar to the subspace trail but it is the divisibility property that is considered. This leads to a new distinguisher for 11-round-reduced SKINNY-64. Finally, we give a closed formula for differential-linear approximations without any assumptions, even confirming that the two differential-linear approximations of SIMECK-32 and SIMECK-48 found by Hadipour *et al.* are deterministic independently of concrete key values. We emphasize that all these applications were not possible before.

Keywords: Cryptanalysis, Geometric Approach, Automatic Search, Transition Matrix

1 Introduction

A secure symmetric-key primitive (block cipher, stream cipher, cryptographic permutation, *etc*) is expected to have an indistinguishable behavior from an idealized one. In practice, whether the primitive meets this expectation is tested by cryptanalysis: confidence is brought about by the continuous analysis performed by the community. There are many attack techniques in the toolbox of cryptanalysts, such as differential [10], linear [26], and integral [22] cryptanalysis, as well as some combinatorial ones such as differential-linear attacks [23]. After creating a cipher, designers and third-party cryptanalysts test its resistance against all state-of-the-art cryptanalysis methods, and it is deemed secure only if it resists all of them with sufficient security margin.

One issue with this process is that there are too many different types of attacks and testing all of them is a very tedious task. In addition, being secure against all known attacks is not foolproof against potential new attacks. A well-known example is the boomerang attack on COCONUT98 [34], which was designed to be secure against differential and linear cryptanalysis, but was quickly broken by this newly introduced technique. Sometimes, even for well-studied ciphers, some unexpected properties are uncovered many years after their publication. At Eurocrypt 2017, a structural property [17] (later named the multiple-of-8 property) was found for the Advanced Encryption Standard (AES) [15]. This was surprising as AES has been carefully studied for almost 30 years, yet this simple property remained undiscovered. Furthermore, although some cryptanalysis methods have been widely used to evaluate the security of cryptographic algorithms, they may not yet be fully understood. For example, before 2015, integral cryptanalysis was already one of the most mature cryptanalytic methods. However, Todo's discovery of the division property [32,33] and the following parity set [13] and monomial prediction [19] revealed a close relationship between integral analysis and Boolean functions of cryptographic algorithms – an evident connection that had not been truly utilized in integral analysis.

More recently, Beyne and Verbauwhede applied the geometric approach [4] to integral analysis [7], significantly deepening the community's understanding of integral analysis and division property once again. This suggests that the current understanding of cryptanalysis methods remains relatively shallow.

A possible explanation for this situation is that cryptanalysis persists to be a task heavily based on the experience of cryptanalysts. Although great progress has been achieved in the past four decades, it is fair to say that the community still knows little about the underlying principles and interconnections of various cryptanalytical methods. Usually, new attacks are found based on the good intuition of the cryptanalysts rather than on some systematic methods. If a unified theory could be developed to describe all (or a large family of) attacks and could be used to discover new ones, it would be extremely beneficial for the advancement of the field. Recently, the *geometric approach* proposed by Beyne [5] has shown the potential to bring about an interesting change in cryptanalysis. This technique has been successfully used to reinterpret linear [4], (quasi-*d*-)differential [6,35] and integral cryptanalysis [7], overcoming many difficulties that could not be solved by classical methods. This theory also proposed a new attack called the ultrametric integral cryptanalysis [8] and attempted to describe the divisibility property of the weight of a ciphertext monomial, which was previously impossible.

The key point of the geometric approach is to linearize a symmetric-key cipher by viewing its input and output spaces as free vector spaces. Treating ciphers as linear mappings brings great convenience and deep insight into cryptanalysis, as researchers have accumulated a wide range of knowledge and many tools in linear algebra.

In the following, we assume the plaintext and ciphertext spaces of a cipher $\mathcal{E} : \mathbb{F}_2^n \to \mathbb{F}_2^n$ are $(\mathbb{F}_2^n)^{\otimes d} = \mathbb{F}_2^n \otimes \mathbb{F}_2^n \otimes \cdots \otimes \mathbb{F}_2^n$, where $\otimes$ represents the tensor product, and $(\mathbb{F}_2^n)^{\otimes d}$ is called a d-th order space which will be formally defined in Definition 5. Choosing a field $\mathbb{K}$ and regarding all vectors in $(\mathbb{F}_2^n)^{\otimes d}$, denoted by $(\delta_u, 0 \leq u < 2^{dn})$, as a set of bases, a free vector space can be induced as

$$\mathbb{K}[(\mathbb{F}_2^n)^{\otimes d}] = \left\{ \sum_u k_u \delta_u : k_u \in \mathbb{K}, u = 0, 1, \ldots, 2^{dn} - 1 \right\}.$$

The pushforward operator $\mathcal{T}^{\mathcal{E}}$ is induced from the cipher $\mathcal{E}$, which is a linear mapping that sends a vector of $\mathbb{K}[(\mathbb{F}_2^n)^{\otimes d}]$ to another in the same space. Here, $(\delta_u, 0 \leq u < 2^{dn})$ plays the role of the standard basis, under which the corresponding matrix of $\mathcal{T}^{\mathcal{E}}$ is uniquely determined, denoted by $T^{\mathcal{E}}$, which is called the transition matrix of $\mathcal{E}$. When we choose a different basis for $\mathcal{T}^{\mathcal{E}}$, denoted by $(\beta_0, \beta_1, \ldots, \beta_{2^{dn}-1})$, with the change-of-basis matrix being F that satisfies

$$(\delta_0, \delta_1, \ldots, \delta_{2^{dn}-1}) = (\beta_0, \beta_1, \ldots, \beta_{2^{dn}-1})F,$$

the corresponding matrix of $\mathcal{T}^{\mathcal{E}}$ becomes another matrix under the new basis $(\beta_0, \beta_1, \ldots, \beta_{2^{dn}-1})$, denoted by $A^{\mathcal{E}}$, that is *similar* to $T^{\mathcal{E}}$, *i.e.*, $A^{\mathcal{E}} = FT^{\mathcal{E}}F^{-1}$. This process can also be performed in a dual way by considering the dual space of $\mathbb{K}[(\mathbb{F}_2^n)^{\otimes d}]$.

With a new proper basis, Beyne found that the element of $A^{\mathcal{E}}$ would be related to some attacks. For example, when $d = 1$ and the new basis is chosen as $(\chi_u, u = 0, 1, \ldots, 2^n - 1)$, where $\chi_u = [(-1)^{u^\top x}, 0 \leq x < 2^n]$ is a column vector with 2^n length ($u^\top x$ representing the inner product of u and x), the element at the u-th column and v-th row of $A^{\mathcal{E}}$ is

$$A^{\mathcal{E}}_{v,u} = 2^{-n} \sum_{x \in \mathbb{F}_2^n} (-1)^{u^\top x + v^\top \mathcal{E}(x)}, \tag{1}$$

which corresponds to a linear approximation of $\mathcal{E}$ with input and output masks being u and v respectively.

From another perspective, for every pair (v, u), $A^{\mathcal{E}}_{v,u}$ is a *statistic* on a set of inputs and outputs, so it provides an opportunity for cryptanalysts to check if this statistic follows the same distribution as a random permutation. In linear cryptanalysis, for example, the expected value of $A^{\mathcal{R}}_{v,u}$ in Equation (1) of a random permutation $\mathcal{R}$ is zero, so if $A^{\mathcal{E}}_{v,u}$ significantly deviates from 0, we can distinguish $\mathcal{E}$ from $\mathcal{R}$. To make the distinguishing process easier, an attacker prefers to choose a pair (v, u) with the largest possible distance between $A^{\mathcal{E}}_{v,u}$ and zero, which corresponds to the process of finding a good pair of input and output masks in linear cryptanalysis. In theory, we can also check the variances or any other values to do the distinguishing attacks, as long as it is useful to be compared with a random permutation.

With various bases and different d, various transition matrices can be obtained. Their elements can be regarded as different statistics of inputs and outputs, providing opportunities (in theory) to compare the cipher and a random permutation. The linear [4], (quasi-d-)differential [6,35], integral attacks [7], and ultrametric integral attacks [8], all follow a similar philosophy. In each of these previous applications of geometric theory, the same kind of basis[7] is always chosen for the input and output space. Such a same-basis configuration works perfectly except for the ultrametric integral cryptanalysis: this attack describes the divisibility of a monomial value in ciphertext. Given a cipher $\mathcal{E} : \mathbb{F}_2^n \to \mathbb{F}_2^m$, the divisibility property is defined as

$$\sum_{x \preceq u} \tau(\mathcal{E}^v(x)) \equiv 0 \mod 2^t. \tag{2}$$

where the partial order $x \preceq u$ means that every coordinate of x is lesser than or equal to the corresponding coordinate of u, $\mathcal{E}^v(x)$ is a multiple of coordinates of $\mathcal{E}(x)$ according to the support of v whose values are 0 or 1, and the function $\tau(\cdot)$ is just changing 0 or 1 from $\mathbb{F}_2$ to $\mathbb{Q}$. When $t = 1$, this is simply the zero-sum property studied by integral cryptanalysis. To study this property, Beyne and Verbauwhede chose the basis as $(\mu_u, 0 \leq u < 2^n)$, where $\mu_u =$

[7] In this paper, when we write same basis/different bases, by default we mean same kind of basis/different kinds of bases. For example, linear bases for $\mathbb{K}[\mathbb{F}_2^n]$ and $\mathbb{K}[\mathbb{F}_2^m]$ are the same (kind of) basis, although they are bases for different spaces.

$[(-1)^{\mathsf{wt}(u\oplus x)}\tau(u^x), 0 \le x < 2^n]$ [8] ($\mathsf{wt}(x)$ represents the Hamming weight of x). The corresponding transition matrix element is

$$A^{\mathcal{E}}_{v,u} = \sum_{x \preceq u} (-1)^{\mathsf{wt}(u\oplus x)} \tau(\mathcal{E}^v(x))$$

By comparing Equations (2) and (1), one can observe that they are not the same, thus Equation (1) cannot be used to study Equations (2) directly. This is because Equation (2) does not equal any single element of $A^{\mathcal{E}}$, which makes the description of the ultrametric integral cryptanalysis different from the other applications of the geometric approach, and more techniques are required to describe this attack.

Our contributions. We first remark that the restriction on the input/output bases to be the same is not necessary. By allowing different bases for the input and output spaces, the geometric approach will be more flexible and will contain more attacks (we emphasize that previous geometric approach papers did not exclude this possibility at all). Given a pair of (same or different) bases, one can obtain the corresponding transition matrix. The elements of the matrix are statistics that provide an opportunity for an attacker to examine whether the cipher's input and output samples follow the same distribution as a random permutation.

This way, given a set of t different bases, t^2 attacks can be naturally defined by them. We call these attacks a family of basis-based attacks defined on the t bases. This paper first recalls three bases used in previous geometric approach papers, then introduces three rules to generate four new bases from these three known bases. To characterize the number of messages in a sample used in an attack, we also define the *order* of a space and an attack. Bases that are used in higher-order attacks can be generated from first-order bases by the tensor product. Finally, from the seven bases used in our work, 7^{2d} attacks are obtained for the d-th order cases, including many known and unknown attacks. This has significantly enlarged the scope of the geometric approach.

A direct benefit of allowing different bases is that the geometric approach can now be applied to describe combinatorial attacks such as differential-linear cryptanalysis [23]. Choosing the basis used in the quasi-differential cryptanalysis [6] for the input space, and the basis used in the linear cryptanalysis [4] (actually, a variant of this basis), we derive a closed formula for the differential-linear approximation without any independence assumption. Automatic search tools are also developed in a natural way to calculate/approximate the exact differential-linear basis. By enumerating all trails, we managed to confirm that two differential-linear approximations of `SIMECK` variants recently found by Hadipour *et al.* are key-independently deterministic [18].

Three more applications are provided as examples to show the effectiveness of our refined geometric approach.

[8] In [8], this basis is equivalently written as $\sum_{x \preceq u} (-1)^{\mathsf{wt}(u\oplus x)} \delta_x$.

In Section 4, we revisit Beyne and Verbauwhede's ultrametric integral cryptanalysis. With a better choice of bases, this attack can be described as a simpler mix-basis attack, where the statistic derived from the bases corresponds exactly to a single element of the transition matrix. Thus, we can focus more on the tracing of trails from all transition matrices, rather than the linear combinations of the divisibility property of different input vectors. In Section 5, we apply our refined geometric approach to the multiple-of-n property, a generalized property of the multiple-of-8 property for the 5-round AES [17]. This property reached only 5 rounds for SKINNY-64 before this paper, but our automatic search method derived from the geometric approach easily extends its length to 10 rounds. Finally, in Section 6, we study the multiple-of-n property as a first-order attack. This is naturally similar to the original multiple-of-n property. We find a new distinguisher for SKINNY-64 that reaches 11 rounds, which is already of the same length as the integral distinguishers. The applications in Sections 5 and 6 provide an example of how to study the same property for different orders.

Paper organization. The remaining paper is organized as follows. Section 2 introduces the notations and recalls some background knowledge. In Section 3, we describe our main contribution of allowing two different bases in the geometric approach. The following four sections give four examples of applications of how to use the refined geometric approach. Section 8 concludes the paper.

2 Preliminaries

2.1 Notations

This paper strictly distinguishes $+$ and $\oplus$ where $x \oplus y = x + y \bmod 2$. A column vector is written as $[x_0, x_1, \ldots, x_{n-1}]$, if the vector can be generated by enumerating some variable, we will also use a simplified version as $[x_i, 0 \le i < n]$. The row vector is represented by $[x_0, x_1, \ldots, x_{n-1}]^\top$. We use double-struck uppercase letters to represent various sets, such as $\mathbb{Q}$ for rational numbers, and $\mathbb{G}$ for a group, *etc.* Ciphers or their parts and transforms are written with calligraphic upper letters such as $\mathcal{E}$ and $\mathcal{F}$. We use normal uppercase letters to represent matrices such as T, A *etc.* An element of A at the intersection of the u-th column and the v-th row is represented by $A_{v,u}$ where $A_{0,0}$ is at the top left. We also call $A_{v,u}$ the (v,u)-element of A. If the (v,u)-element of A is the value of a function $f_u(v)$, we write A as $A = [f_u(v)]_{v,u}$, where the first subscript (v) is for the index of the row and the second (u) is for the column. Writing n column vectors together leads to a matrix; we will interchangeably use $A = [f_u(v)]_{v,u}$ and $A = (f_0, f_1, \ldots, f_{n-1})$ where $f_u = [f_u(v), 0 \le v < n]$. For $x \in \mathbb{F}_2^n$, we use x_i to represent the i-th bit of x and x_0 is the most significant bit. For two vectors $a, b \in \mathbb{F}_2^n$, $a \succeq b$ means $a_i \ge b_i$ for all i. Similarly, $a \preceq b$ means $a_i \le b_i$ for all i.

We introduce several functions that have been extensively used in previous geometric approach papers and will play important roles in this one.

Function 1 (Weight function $\mathsf{wt}(\cdot)$) $\mathsf{wt}(\cdot) : \mathbb{F}_2^n \to \mathbb{N}$, *for any* $x \in \mathbb{F}_2^n$, $\mathsf{wt}(x)$ *is the Hamming weight of* x.

Function 2 (Correlation function $(-1)^{u^\top(\cdot)}$) *Let $u \in \mathbb{F}_2^n$, we define $(-1)^{u^\top(\cdot)} : \mathbb{F}_2^n \to \mathbb{Q}$ as*

$$(-1)^{u^\top x} = \begin{cases} 1 & \text{if } u^\top x = 0, \\ -1 & \text{if } u^\top x = 1, \end{cases}$$

where $u^\top x$ means the inner product of u and x, *i.e.*, $u^\top x = \sum_{0 \le i < n} u_i x_i \bmod 2$.

Function 3 (Dirac delta function $\delta_u(\cdot)$) *Let $u \in \mathbb{F}_2^n$, we define $\delta_u(\cdot) : \mathbb{F}_2^n \to \mathbb{Q}$ as*

$$\delta_u(x) = \begin{cases} 1 & \text{if } x = u, \\ 0 & \text{otherwise.} \end{cases}$$

Remark. Later in this paper, the notation "δ_u" (rather than "$\delta_u(\cdot)$") is also used to represent the unit vector where the u-th element is 1. This interpretation is natural when we express it as $\delta_u = [\delta_u(x), 0 \le x < 2^n]$.

Function 4 (Power function $(\cdot)^u$) *Let $u \in \mathbb{F}_2^n$, we define $(\cdot)^u : \mathbb{F}_2^n \to \mathbb{Q}$ as*

$$x^u = \begin{cases} 1 & \text{if } x \succeq u, \\ 0 & \text{otherwise.} \end{cases}$$

Note that in [8] and many previous papers, x^u is defined as a value in $\mathbb{F}_2$. To transform x^u into numbers in $\mathbb{Q}$, Beyne and Verbauwhede introduced an embedding function $\tau : \mathbb{F}_2 \to \mathbb{Q}$ where $\tau(x) = x$. However, since this paper only works in $\mathbb{Q}$, we default use x^u as a number in $\mathbb{Q}$ to omit the notation τ for the sake of a simpler description.

Function 5 (Exponential function $u^{(\cdot)}$) *Let $u \in \mathbb{F}_2^n$, we define $u^{(\cdot)} : \mathbb{F}_2^n \to \mathbb{Q}$ as*

$$u^x = \begin{cases} 1 & \text{if } x \preceq u, \\ 0 & \text{otherwise.} \end{cases}$$

Similarly to the power function, we also omit the τ function and regard u^x by default as a rational number.

2.2 Brief Introduction to Beyne's Geometric Approach

The following contents are mainly summarized from [5]. Mathematical background knowledge can be found in some textbooks, such as [29]. Assume that there exists a cipher $\mathcal{E}$ that takes an element of $\mathbb{F}_2^n$ to $\mathbb{F}_2^m$:

$$\mathcal{E} : \mathbb{F}_2^n \to \mathbb{F}_2^m.$$

Now, $\mathcal{E}$ is a function that connects two spaces. Beyne introduced a way to use linear algebra techniques to analyze the properties of $\mathcal{E}$ [5]. There are two methods for linearizing $\mathcal{E}$. The first method is using the free vector space concept.

Namely, let $\mathbb{K}$ be any field, we denote $\mathbb{K}[\mathbb{F}_2^n]$ as a vector space over $\mathbb{K}$, the elements of this vector space are $\sum_i a_i u_i$ where $a_i \in \mathbb{K}$, $u_i \in \mathbb{F}_2^n$. It is easy to show that the dimension of $\mathbb{K}[\mathbb{F}_2^n]$ is 2^n, the cardinality of $\mathbb{F}_2^n$. Then, $\mathcal{E}$ will induce a map $\mathcal{T}^{\mathcal{E}}$ that connects $\sum_i a_i u_i$ to $\sum_i b_i \mathcal{E}(u_i)$, and one can observe that $\mathcal{T}^{\mathcal{E}}$ is a linear map from $\mathbb{K}[\mathbb{F}_2^n]$ to $\mathbb{K}[\mathbb{F}_2^m]$.

The other method is through the linear function space. Denote $\mathbb{K}[\mathbb{F}_2^n]^\vee$ all linear functions from $\mathbb{F}_2^n$ to $\mathbb{K}$, then $\mathbb{K}[\mathbb{F}_2^n]^\vee$ forms a vector space naturally and its dimension is also 2^n. Actually, $\mathbb{K}[\mathbb{F}_2^n]$ and $\mathbb{K}[\mathbb{F}_2^n]^\vee$ are dual to each other. Therefore, $\mathcal{E}$ can induce another linear map $(\mathcal{T}^{\mathcal{E}})^\vee$ between $\mathbb{K}[\mathbb{F}_2^m]^\vee$ to $\mathbb{K}[\mathbb{F}_2^n]^\vee$ through the pullback function. For any function $f \in \mathbb{K}[\mathbb{F}_2^m]^\vee$, $(\mathcal{T}^{\mathcal{E}})^\vee(f)$ is a linear function in $\mathbb{K}[\mathbb{F}_2^n]^\vee$, defined as follows: for any $x \in \mathbb{F}_2^n$, $(\mathcal{T}^{\mathcal{E}})^\vee(f)(x) = f(\mathcal{E}(x))$.

Since $\mathcal{T}^{\mathcal{E}} : \mathbb{K}[\mathbb{F}_2^n] \longrightarrow \mathbb{K}[\mathbb{F}_2^m]$ and $(\mathcal{T}^{\mathcal{E}})^\vee : \mathbb{K}[\mathbb{F}_2^m]^\vee \longrightarrow \mathbb{K}[\mathbb{F}_2^n]^\vee$ are both linear, their matrix representation are determined after fixing the bases for the input and output spaces, called *transition matrix*. Different cryptanalytic theories are obtained by expressing cryptanalytic properties with respect to different pairs of dual bases for $\mathbb{K}[\mathbb{F}_2^n]$ and $\mathbb{K}[\mathbb{F}_2^n]^\vee$.

Definition 1 (Dual basis). *Let $\mathbb{V}$ be a linear space over a field $\mathbb{K}$. A basis for $\mathbb{V}$, denoted by β_u, and a basis for $\mathbb{V}^\vee$, denoted by $\beta_u^\vee$, is called a pair of dual bases if*

$$\beta_u^\vee(\beta_v) = \begin{cases} 0 & u \neq v \\ 1 & u = v \end{cases}$$

The matrix reflects the nature of the cipher, and each element of the matrix represents a statistic describing a cryptanalytic property.

Lemma 1. *Let $\mathbb{V}$ and $\mathbb{W}$ be two linear spaces on $\mathbb{K}$ of dimension n and m, respectively. Fix bases $(\alpha_u, 0 \leq u < n)$ and $(\beta_v, 0 \leq v < m)$ for $\mathbb{V}$ and $\mathbb{W}$, respectively. Denote by $\mathcal{T}$ the linear map from $\mathbb{V}$ to $\mathbb{W}$, and denote by T the matrix of $\mathcal{T}$ corresponding to the bases α_u and β_v. Then, the element (v, u) of T is equal to $\beta_v^\vee(\mathcal{T}(\alpha_u))$.*

Elements of the transition matrix of a cipher determined by a pair of dual bases can be seen as statistics. In Section 2.3, we recall applications of the geometric approach to linear [4], quasi-differential [6], quasi-d-differential [35], and ultrametric integral cryptanalysis [8].

Before we go on, we need to introduce some basic rules of the tensor product $\otimes$, which is important to understand geometry theory.

Basis calculation rules of $\otimes$. The tensor product can be defined in different ways (see *e.g.*. [5, Section 2.2.3]. Here, we introduce the basis-dependent definition. Let $V^{(i)}$ have a set of basis $\beta_u^{(i)}, 0 \leq u < |V^{(i)}|$, the tensor product of $V^{(1)}, \ldots, V^{(n)}$ is defined as

$$\bigotimes_i V^{(i)} = V^{(1)} \otimes V^{(2)} \otimes \cdots \otimes V^{(n)} = \mathrm{Span}\left\{ \beta_{u^{(1)}}^{(1)} \otimes \beta_{u^{(1)}}^{(2)} \otimes \cdots \otimes \beta_{u^{(n)}}^{(n)} : \text{for all } u^{(i)} \right\}.$$

When $V^{(i)}$ are the same as V, $\bigotimes_i V^{(i)}$ is also written as $V^{\otimes d}$.

Let $v^{(i)} = \sum_{u^{(i)}} c^{(i)}_{u^{(i)}} \beta^{(i)}_{u^{(i)}}$ be a vector of $V^{(i)}$, the tensor product of $v^{(1)}, v^{(2)}, \ldots, v^{(n)}$ is calculated as

$$\bigotimes_i v^{(i)} = v^{(1)} \otimes v^{(2)} \otimes \cdots \otimes v^{(n)} = \sum_{u^{(1)}} \cdots \sum_{u^{(n)}} \left(\prod_{i=1}^{n} c^{(i)}_{u^{(i)}} \right) \beta^{(1)}_{u^{(1)}} \otimes \beta^{(2)}_{u^{(1)}} \otimes \cdots \otimes \beta^{(n)}_{u^{(n)}}.$$

Let $\mathcal{L}^{(i)} : \mathbb{V}^{(i)} \longrightarrow \mathbb{W}^{(i)}$, the tensor product of $\bigotimes_i \mathcal{L}^{(i)}$ is defined as

$$\bigotimes_i \mathcal{L}^{(i)} : \bigotimes_i \mathbb{V}^{(i)} \longrightarrow \bigotimes_i \mathbb{W}^{(i)}, \quad \bigotimes_i v^{(i)} \mapsto \bigotimes_i \mathcal{L}^{(i)}(v^{(i)}).$$

The tensor product of corresponding matrices of these linear maps is

$$\left(\bigotimes_i L^{(i)} \right)_{\bigotimes_i v^{(i)}, \bigotimes_i u^{(i)}} = \left(\bigotimes_i L^{(i)} \right)_{v^{(1)}||\cdots||v^{(n)}, u^{(1)}||\cdots||v^{(n)}} = \prod_i L^{(i)}_{v^{(i)}, u^{(i)}},$$

where $L^{(i)}$ is the matrix of $\mathcal{L}^{(i)}$.

2.3 Geometric Approach to Linear, Quasi-Differential, Quasi-*d*-Differential and Ultrametric Integral Cryptanalysis

For two functions $\chi, \phi \in \mathbb{K}[\mathbb{G}]^\vee$, their inner product is denoted by $\langle \chi, \phi \rangle$. When $\chi, \phi \in \mathbb{K}[\mathbb{G}]^\vee$ is orthogonal, the inner product can be written as (note that this paper works over $\mathbb{Q}$, so we do not use $\overline{\phi(x)}$ in the following)

$$\langle \chi, \phi \rangle = \sum_{x \in \mathbb{G}} \chi(x)\phi(x).$$

Now when a function space is equipped with such inner product, we have a natural definition of an orthogonal basis:

Definition 2. *A basis $\phi_u, 0 \leq u \leq |\mathbb{G}|$ is called an orthogonal basis if ϕ_u satisfies the following property:*

$$\langle \phi_u, \phi_v \rangle = \begin{cases} \langle \phi_u, \phi_u \rangle & \text{if } u{=}v \\ 0 & \text{otherwise} \end{cases}$$

Definition 3 (Linear isomorphism with $(\phi_u, 0 \leq u < |\mathbb{G}|)$). *There is a linear isomorphism $\mathcal{F}$ from $\mathbb{K}[\mathbb{G}]^\vee$ to its dual space $\mathbb{K}[\mathbb{G}]^{\vee\vee}$. The behaviour of this linear isomorphism is determined by its effect on all the bases:*

$$\mathcal{F}(f) : \mathcal{F}(f)(\phi_u) = \langle f, \phi_u \rangle / \langle \phi_u, \phi_u \rangle$$

This linear isomorphism is useful to construct the dual basis. Given orthogonal basis $\phi_u \in \mathbb{K}[\mathbb{G}]^\vee$, $\mathcal{F}(\phi_u)$ is the dual of ϕ_u in $\mathbb{K}[\mathbb{G}]^{\vee\vee}$, since

$$[\mathcal{F}(\phi_u)](\phi_v) = \langle \phi_u, \phi_v \rangle / \langle \phi_u, \phi_u \rangle = \begin{cases} 1 & \text{if } u = v \\ 0 & \text{otherwise .} \end{cases}$$

Next, we introduce how to derive the transition matrix of a cipher under a pair of bases. In this paper, we will consider the combination of different attacks by combining different bases. Therefore, by default, we choose $\mathbb{K} = \mathbb{Q}$ to induce the corresponding free vector space in this paper. For linear, quasidifferential and quasi-d-differential cryptanalysis, we use the pullback operator, *i.e.*, for $\mathcal{E} : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2^m$, the corresponding transition matrices for $(A^{\mathcal{E}})^\vee : \mathbb{Q}[\mathbb{F}_2^m]^\vee \longrightarrow \mathbb{Q}[\mathbb{F}_2^n]^\vee$ are calculated first, then $A^{\mathcal{E}}$ is obtained as the transpose of $(A^{\mathcal{E}})^\vee$. For ultrametric integral cryptanalysis, we use the pushforward operator to derive $A^{\mathcal{E}}$ directly. In fact, both methods work for each case. We just follow the choices of the original papers.

Linear cryptanalysis. Beyne applied the geometric approach to linear cryptanalysis in [4]. The group $(\mathbb{F}_2^n, \oplus)$ is abelian; therefore, all the characters $\chi_u(x) = (-1)^{u^\top x}$ for $u \in \mathbb{F}_2^n$ form a set of orthogonal bases. With $(\chi_u, 0 \le u < 2^n)$ and $(\chi_v, 0 \le v < 2^m)$ being the bases for $\mathbb{Q}[\mathbb{F}_2^n]^\vee$ and $\mathbb{Q}[\mathbb{F}_2^m]^\vee$, respectively, the linear isomorphism in Definition 3 is just the famous Fourier transform. Thus, the pullback matrix of $\mathcal{E}$, *i.e.*, the matrix of $(\mathcal{A}^{\mathcal{E}})^\vee$ (denoted by $(A^{\mathcal{E}})^\vee$) is calculated with the Lemma 1.

$$\begin{aligned}(A^{\mathcal{E}})^\vee_{v,u} =&(\chi_v)^\vee\left((\mathcal{A}^{\mathcal{E}})^\vee(\chi_u)\right) = \mathcal{F}(\chi_v)\left((\mathcal{A}^{\mathcal{E}})^\vee(\chi_u)\right) = \mathcal{F}(\chi_v)\left(\chi_u(\mathcal{E}(x))\right)\\ =&\frac{1}{\langle\chi_u,\chi_u\rangle}\sum_{x\in\mathbb{F}_2^n}\chi_v(x)\chi_u(\mathcal{E}(x)) = 2^{-n}\sum_{x\in\mathbb{F}_2^n}(-1)^{v^\top x\oplus u^\top(\mathcal{E}(x))}\end{aligned}$$

The matrix of the pushforward is the transpose of $(A^{\mathcal{E}})^\vee$, thus

$$A^{\mathcal{E}}_{v,u} = 2^{-n}\sum_{x\in\mathbb{F}_2^n}(-1)^{u^\top x\oplus v^\top(\mathcal{E}(x))},$$

which is the statistic used in the linear cryptanalysis [26].

Quasi-d-differential cryptanalysis. Beyne and Rijmen introduced the quasidifferential cryptanalysis [6], which was later generalized to the quasi-d-differential cryptanalysis [35]. For the quasi-d-differential cryptanalysis (when $d = 1$, it is the basis for the quasi-differential cryptanalysis), the group is $((\mathbb{F}_2^n)^{\otimes d}, \oplus)$ and the basis for $\mathbb{Q}[(\mathbb{F}_2^n)^{\otimes d}]^\vee$ is chosen as

$$\beta_{u_0,\ldots,u_{d-1}}(x_0, x_1, \ldots, x_{d-1}) = (-1)^{{u_0}^\top x_0}\prod_{1\le i<d}\delta_{u_i}(x_0\oplus x_i).$$

Note that all the bases here are orthogonal, so the linear isomorphism $\mathcal{F}$ is constructed similarly to the Fourier transform as

$$\begin{aligned}\mathcal{F}(f) : \mathcal{F}(f)(\beta_{u_0,\ldots u_{d-1}}) =&\langle f, \beta_{u_0,\ldots u_{d-1}}\rangle/\langle\beta_{u_0,\ldots u_{d-1}}, \beta_{u_0,\ldots u_{d-1}}\rangle\\ =&\frac{1}{2^{-n}}\sum_{x_0,\ldots,x_{d-1}\in(\mathbb{F}_2^n)^{\otimes d}} f(x)\beta_{u_0,\ldots u_{d-1}}(x_0,\ldots,x_{d-1}).\end{aligned}$$

Therefore, choosing $\beta_{u_0,\dots u_{d-1}}$ for both the input and output function spaces, $((v_0,\dots,v_{d-1}),(u_0,\dots,u_{d-1}))$ element of the matrix $(A^{\mathcal{E}})^\vee$ which is the pullback operator $(\mathcal{A}^{\mathcal{E}})^\vee$ is calculated as

$$\begin{aligned}
&(A^{\mathcal{E}})^\vee_{(v_0,\dots,v_{d-1}),(u_0,\dots,u_{d-1})} = \beta^\vee_{v_0,\dots,v_{d-1}}\left((\mathcal{A}^{\mathcal{E}})^\vee(\beta_{u_0,\dots u_{d-1}})\right)\\
&= \mathcal{F}(\beta_{v_0,\dots,v_{d-1}})\left((\mathcal{A}^{\mathcal{E}})^\vee(\beta_{u_0,\dots u_{d-1}})\right) = \mathcal{F}(\beta_{v_0,\dots,v_{d-1}})\left(\beta_{u_0,\dots u_{d-1}}(\mathcal{E}(x))\right)\\
&= 2^{-n}\sum_{x_0,\dots,x_{d-1}\in(\mathbb{F}_2^n)^{\otimes d}\mathcal{E}(x_0)\oplus\mathcal{E}(x_0\oplus v_i)=u_i,1\le i<d} (-1)^{v_0^\top x_0\oplus u_0^\top(\mathcal{E}(x_0))}
\end{aligned}$$

Again, the matrix of the pushforward is the transpose of $(A^{\mathcal{E}})^\vee$, *i.e.*,

$$A^{\mathcal{E}}_{v,u} = 2^{-n}\sum_{\substack{x_0,\dots,x_{d-1}\in(\mathbb{F}_2^n)^{\otimes d}\\ \mathcal{E}(x_0)\oplus\mathcal{E}(x_0\oplus u_i)=v_i,1\le i<d}} (-1)^{u_0^\top x_0\oplus v_0^\top(\mathcal{E}(x_0))},$$

which is the statistic used in the quasi-d-differential cryptanalysis.

Ultrametric integral cryptanlysis. Beyne and Verbauwhede applied the geometric approach to integral cryptanalysis [7] and later introduced the ultrametric integral cryptanalysis [8]. The former works on $\mathbb{F}_2[\mathbb{F}_2^n]$, while the ultrametric integral cryptanalysis works on $\mathbb{Q}[\mathbb{F}_2^n]$, and the ultrametric integral cryptanalysis actually contains the cases of integral cryptanalysis. Since this paper will consider the combination of different attacks, we have to work on $\mathbb{Q}[\mathbb{F}_2^n]$. Thus, only the ultrametric integral cryptanalysis is introduced here.

Unlike previous linear and differential cryptanalysis, Beyne and Verbauwhede derived the transition matrix for the pushforward operator directly rather than from the transpose of the pullback operator. To show it, we also used their method to get the transition matrix. In this case, they constructed the basis for $\mathbb{Q}[\mathbb{F}_2^n]$, with the help of the basis of $\mathbb{Q}[\mathbb{F}_2^n]^\vee$. The basis of $\mathbb{Q}[\mathbb{F}_2^n]^\vee$ is chosen as

$$(\mu_u)^\vee : \mathbb{F}_2^n \longrightarrow \mathbb{Q}, \quad (\mu_u)^\vee(x) = x^u = \begin{cases} 1 & x \succeq u,\\ 0 & otherwise.\end{cases}$$

All these bases form the characters of the monoid $(\mathbb{F}_2^n,\wedge)$. Its dual basis in $\mathbb{Q}[\mathbb{F}_2^n]$ can be obtained by solving a set of linear equations, which is

$$\mu_u = \sum_{x\preceq u}(-1)^{\mathrm{wt}(x+u)}\delta_x = [(-1)^{\mathrm{wt}(x+u)}u^x, 0\le x<2^n]$$

Thus, the matrix of $\mathcal{A}^{\mathcal{E}} : \mathbb{Q}[\mathbb{F}_2^n] \longrightarrow \mathbb{Q}[\mathbb{F}_2^n]$ is

$$A^{\mathcal{E}}_{v,u} = \mu_v^\vee(\mathcal{A}^{\mathcal{E}}(\mu_u)) = \sum_{x\preceq u}(-1)^{\mathrm{wt}(x+u)}\mathcal{E}^v(x).$$

2.4 Propagation of Transition Matrix, Measurement and Automatic Search

Calculating the element (v,u) of a transition matrix $A^{\mathcal{E}}$ is challenging, as the sizes of such transition matrices are extremely large. However, the transition matrices enjoy the following property.

Theorem 1 (Propagation of transition matrices [6,4]). *The transition matrix of* $\mathcal{E} : \mathbb{F}_2^n \to \mathbb{F}_2^n$ *satisfies:*

(1) If $\mathcal{E} = \mathcal{E}_{s_0} || \cdots || \mathcal{E}_{s_{m-1}}$, $A^{\mathcal{E}} = \bigotimes_{i=0}^{m-1} A^{\mathcal{E}_i}$.
(2) if $\mathcal{E} = \mathcal{E}_{r-1} \circ \cdots \circ \mathcal{E}_1 \circ \mathcal{E}_0$, $A^{\mathcal{E}} = A^{\mathcal{E}_{r-1}} \cdot \cdots \cdot A^{\mathcal{E}_1} \cdot A^{\mathcal{E}_0}$.

According to Theorem 1, if $\mathcal{E} = \mathcal{E}_{r-1} \circ \mathcal{E}_{r-2} \circ \cdots \circ \mathcal{E}_0$ we have

$$A^{\mathcal{E}}_{u_r,u_0} = \sum_{u_{r-1},u_{r-2},\ldots,u_2} \prod_{i=0}^{r-1} A^{\mathcal{E}}_{u_{i+1},u_i}. \tag{3}$$

$A^{\mathcal{E}}_{u_r,u_0}$ is equal to the sum of *correlations* of all trails with input and output being u_0 and u_r, respectively.

Definition 4 (Trail and correlation [4]). *In Equation* (3)*,* $(u_0, u_1, \ldots, u_r)$ *is called the trail of the corresponding attack.* $\prod_{i=0}^{r-1} A^{\mathcal{E}_i}_{u_{i+1},u_i}$ *is called the correlation of this trail.*

Therefore, $A^{\mathcal{E}}_{u_r,u_0}$ is the sum of all the correlations of the trails that connect u_0 and u_r. The search for a trail or the enumeration of trails has been extensively studied in previous articles related to automatic search, such as [27,30], which can be and have been reused in a natural way for the search of trails in the geometric approach [5]. In Appendix B, we give a high-level description of the current automatic search methods.

To use the statistic $A^{\mathcal{E}}_{u_r,u_0}$ for a distinguishing attack, two types of *measures* are known. The first is to use the value of $A^{\mathcal{E}}_{u_r,u_0}$ as the correlation in linear attacks or the probability in differential attacks. The values for a cipher and a random permutation are expected to follow different statistical distributions that can be distinguished with some samples. The second is not to approximate the real value of $A^{\mathcal{E}}_{u_r,u_0}$, but to know if $A^{\mathcal{E}}_{u_r,u_0}$ is a multiple of a certain number (divisibility property). For example, as Beyne and Verbauwhede recently showed, the zero-sum property in integral cryptanalysis is equivalent to saying that the weight of the output Boolean function under some input sets is a multiple of 2. Thus, it is natural to consider whether the weight is also a multiple of 2^v ($v \geq 2$). This measurement has been well studied in [8]. $A^{\mathcal{E}}_{u_r,u_0} \equiv 0 \bmod 2^v$ is equivalent to saying $|A^{\mathcal{E}}_{u_r,u_0}|_2 \leq 2^{-v}$ where $|x|_2$ represents the 2-adic absolute value of x.

Since the number of trails can be too large to exhaust, in most cases, only one or a small percentage of trails that have the most significant correlations can be searched and used. These trails are called dominant trails [5]. In the first measurement, the sum of dominant trails cannot ensure that the approximation is always sound. Yet, in the second measurement, due to the ultrametric triangle inequality $|x + y|_2 \leq \max\{|x|_2, |y|_2\}$, the correlation of the dominant trails can bound the summed correlations of all trails.

3 Geometric Approach While Allowing Different Bases

This section introduces the main contribution of this paper. The geometric approach does not set the restriction that it has to use two same bases for input and output spaces. However, for certain reasons, all previous related papers have used the same basis. We demonstrate that using different bases for the input and output spaces is entirely feasible and can enhance the geometric approach significantly.

Besides, we choose to use a simpler way to write Beyne's geometric theories where we avoid advanced (for us) mathematical conceptions such as the dual basis representation. To understand this section, readers are only required to have some knowledge of linear transforms, change-of-basis operations, and the basic calculation rules of the tensor product in Section 2.2. This way of presenting is not new and has naturally been mentioned in previous geometric approach papers (but always mixed with other ways of writing).

3.1 Using Different Basis for Input and Output Spaces

Consider $\mathcal{E} : \mathbb{F}_2^n \to \mathbb{F}_2^m$. An attack on $\mathcal{E}$ can use samples $(p_0, p_1, \ldots, p_{d-1}) \in (\mathbb{F}_2^n)^d$ and corresponding $(c_0, c_1, \ldots, c_{d-1}) \in (\mathbb{F}_2^m)^d$ for a distinguishing or key-recovery attack. For the distinguishing attack, a *statistic* is calculated from $(p_0, p_1, \ldots, p_{d-1})$ and $(c_0, c_1, \ldots, c_{d-1})$ and reflects some statistical properties of the cipher. If this statistic of the target cipher follows a different probability distribution from a random function, we can perform a distinguishing attack with some computational resources. The number of sample components, *i.e.*, d, is an important information about the attack, defining the order of the input and output spaces. We call this number *the order of an attack*.

Definition 5 (The order of an attack). *The number d of $(\mathbb{F}_2^n)^d$ is called the order of the space $(\mathbb{F}_2^n)^d$. An attack that uses samples in a d-th order space is called the d-th order attack. Equivalently, we can consider $(\mathbb{F}_2^n)^d$ as $(\mathbb{F}_2^n)^{\otimes d}$ whose elements are $p_0 \otimes p_1 \otimes \cdots \otimes p_{d-1} \in (\mathbb{F}_2^n)^{\otimes d}$. A d-th order attack on $\mathcal{E} : \mathbb{F}_2^n \to \mathbb{F}_2^m$ uses samples like $p_0 \otimes p_1 \otimes \cdots \otimes p_{d-1} \in (\mathbb{F}_2^n)^{\otimes d}$ and $c_0 \otimes c_1 \otimes \cdots \otimes c_{d-1} \in (\mathbb{F}_2^m)^{\otimes d}$ to compute the corresponding statistic.*

Notation trick. For sake of convenience, in the following we will say that a d-th order attack on $\mathcal{E} : \mathbb{F}_2^n \to \mathbb{F}_2^m$ is an attack on $\mathcal{E}^{\otimes d} : (\mathbb{F}_2^n)^{\otimes d} \to (\mathbb{F}_2^m)^{\otimes d}$ (but in the cases when the order is clear, "$\otimes d$" might be omitted for simplicity). For example, differential cryptanalysis is a typical second-order attack as it uses a pair of messages, so we say that the differential attack is on $\mathcal{E}^{\otimes 2} : (\mathbb{F}_2^n)^{\otimes 2} \to (\mathbb{F}_2^m)^{\otimes 2}$. Furthermore, note that in a d-th ($d > 1$) order attack, in fact we are more interested in the difference between the i-th ($i > 1$) and the first component of the sample. Thus, we consider the input and output of $\mathcal{E}^{\otimes d}$ as

$$p_0 \otimes (p_0 \oplus p_1) \otimes (p_0 \oplus p_{d-1}) \text{ and } c_0 \otimes (c_0 \oplus c_1) \otimes (c_0 \oplus c_{d-1}),$$

which is similar to the setting of polytopic attacks [31]. In the following, by default we will use $(x_0, x_1, \ldots, x_{d-1})$ to represent the input of an attack where x_i $(i > 1)$ is the difference between the original element i-th and the first, and we will use $(\mathcal{E}(x_0), \mathcal{E}(x_1), \ldots, \mathcal{E}(x_{d-1}))$ to represent the output where $\mathcal{E}(x_i)$ $(i > 1)$ also represents the difference of the i-th ciphertext and the first. For example, in the differential attack, (x_0, x_1) and $(\mathcal{E}(x_0), \mathcal{E}(x_1))$ are, respectively, the input and output pairs where x_1 is the input difference and $\mathcal{E}(x_1)$ is the output difference. Such a trick can simplify the notation.

Consider a d-th space $(\mathbb{F}_2^n)^{\otimes d}$ and a field $\mathbb{K} = \mathbb{Q}$, a free vector space can be induced as $\mathbb{Q}[(\mathbb{F}_2^n)^{\otimes d}]$, where elements in $(\mathbb{F}_2^n)^{\otimes d}$ are a set of bases of $\mathbb{Q}[(\mathbb{F}_2^n)^{\otimes d}]$. This set of bases can be viewed as the standard basis consisting of 2^{dn} unit vectors, and the u-th unit vector can be written as $\delta_u = [\delta_u(x), 0 \leq x < 2^{dn}]$. All the 2^{dn} bases are written together like a matrix, denoted by $[\delta_u(x)]_{x,u}$ (recall Section 2.1).

Since $\mathbb{Q}[(\mathbb{F}_2^n)^{\otimes d}]$ is a linear space, it is possible to choose another set of bases consisting of 2^{dn} linearly independent vectors. Suppose that the u-th basis is denoted by $\alpha_u = [\alpha_u(x), 0 \leq x < 2^{dn}]$, thus $(\alpha_u, 0 \leq u < 2^{nd}) = [\alpha_u(x)]_{x,u}$ is the new set of basis. The relationship between the two sets of basis is connected by a change-of-basis matrix *i.e.*,

$$(\delta_0, \delta_1, \ldots, \delta_{2^{dn}-1}) = (\alpha_0, \alpha_1, \ldots, \alpha_{2^{dn}-1})P,$$

where $P \in \mathbb{Q}^{2^{dn} \times 2^{dn}}$.

For $\mathcal{E}^{\otimes d} : (\mathbb{F}_2^n)^{\otimes d} \longrightarrow (\mathbb{F}_2^m)^{\otimes d}$, a linear map $\mathcal{T}^{\mathcal{E}^{\otimes d}}$ is induced that maps a standard basis of $\mathbb{Q}[(\mathbb{F}_2^n)^{\otimes d}]$ to a standard basis of $\mathbb{Q}[(\mathbb{F}_2^m)^{\otimes d}]$, *i.e.*, $\mathcal{T}^{\mathcal{E}^{\otimes d}}(\delta_u) = \delta_{\mathcal{E}^{\otimes d}(u)}$. Under the standard basis, the corresponding matrix of $\mathcal{T}^{\mathcal{E}^{\otimes d}}$ is denoted by $T^{\mathcal{E}^{\otimes d}}$ with $T^{\mathcal{E}^{\otimes d}}_{v,u} = \delta_v(\mathcal{E}^{\otimes d}(\delta_u))$.

After doing the change-of-basis for the input space with a new set of basis $[\alpha_u(x)]_{x,u}$, *i.e.*,

$$(\delta_0, \delta_1, \ldots, \delta_{2^{nd}-1}) = (\alpha_0, \alpha_1, \ldots, \alpha_{2^{nd}-1})I,$$

and for the output space with a new set of basis $[\beta_u(x)]_{x,u}$, *i.e.*,

$$(\delta_0, \delta_1, \ldots, \delta_{2^{nd}-1}) = (\beta_0, \beta_1, \ldots, \beta_{2^{nd}-1})O,$$

where I and O are the corresponding change-of-basis matrices. A new transition matrix of $\mathcal{E}^{\otimes d}$ can be deduced as

$$A^{\mathcal{E}^{\otimes d}} = O\ T^{\mathcal{E}^{\otimes d}}\ I^{-1},$$

Note that the element (v, u) of a matrix can be obtained by multiplying to the left a row unit vector $\delta_v^\top$ and multiplying to the right a column unit vector δ_u, thus the (v, u)-element of $A^{\mathcal{E}^{\otimes d}}$ is calculated by

$$A^{\mathcal{E}^{\otimes d}}_{v,u} = {\delta_v}^\top\ A^{\mathcal{E}^{\otimes d}}\ \delta_u = \delta_v^\top\ (O\ T^{\mathcal{E}^{\otimes d}}\ I^{-1})\ \delta_u = \underbrace{(\delta_v^\top\ O)}_{v\text{-th row of } O}\quad T^{\mathcal{E}^{\otimes d}}\quad \underbrace{(I^{-1}\ \delta_u)}_{u\text{-th column of } I^{-1}}$$

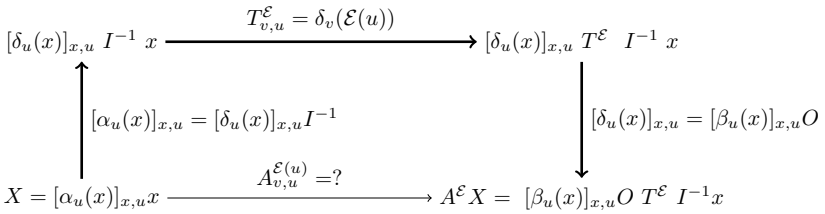

Fig. 1: The illustration of the geometric approach on a cryptanalysis with two different bases. Note $I^{-1} = [\alpha_u(x)]_{x,u}$ and $O = [\beta_u(x)]^{-1}_{x,u}$. Given $X = [\alpha_u(x)]_{x,u}x$ where x is the coordinate. After the change of basis operation in the input space, it becomes a vector represented by $[\delta_u(x)]_{x,u}$, and then transformed by $T^{\mathcal{E}}$ to $[\delta_u(x)]_{x,u}T^{\mathcal{E}}I^{-1}x$. After the change of basis operation in the output space, it becomes to the final form under the basis $[\beta_u(x)]_{x,u}$ whose coordinate is $OT^{\mathcal{E}}I^{-1}x$.

The process is shown in Figure 1. Note that I^{-1} is just the matrix $[\alpha_u(x)]_{x,u}$, thus the u-th column of I^{-1} is $[\alpha_u(x), 0 \le x < 2^{dn}]$. The v-th row of O is based on the shape of $[\beta_u(x)]_{x,u}$, as $O = [\beta_u(x)]^{-1}_{x,u}$ (the inverse of $[\beta_u(x)]_{x,u}$). Assume that $[\beta_u(x)]^{-1}_{x,u}$ can also be written in a compact form denoted by $[\beta^\star_u(x)]_{x,u}$, the v-th row of O is then $[\beta^\star_y(v), 0 \le y < 2^{dn}]^\top$. Thus (a step-by-step explanation of the calculation process is give in Appendix C),

$$
\begin{aligned}
A^{\mathcal{E}^{\otimes d}}_{v,u} &= [\beta^\star_y(v), 0 \le y < 2^{dn}]^\top\ T^{\mathcal{E}^{\otimes d}}\ [\alpha_u(x), 0 \le x < 2^{dn}] \\
&= \left[\sum_{y \in (\mathbb{F}_2^n)^{\otimes d}} \beta^\star_y(v)\delta_y(\mathcal{E}(0)), \ldots, \sum_{y \in (\mathbb{F}_2^n)^{\otimes d}} \beta^\star_y(v)\delta_y(\mathcal{E}(2^{dn}-1)) \right]^\top\ [\alpha_u(x), 0 \le x < 2^{dn}] \\
&= \left[\beta^\star_{\mathcal{E}^{\otimes d}(0)}(v), \ldots, \beta^\star_{\mathcal{E}^{\otimes d}(2^{dn}-1)}(v) \right]^\top\ [\alpha_u(x), 0 \le x < 2^{dn}] \\
&= \sum_{x \in (\mathbb{F}_2^n)^{\otimes d}} \beta^\star_{\mathcal{E}^{\otimes d}(x)}(v)\ \alpha_u(x)
\end{aligned}
\tag{4}
$$

Depending on whether the two bases for the input and output spaces are the same, we divide attacks into two kinds.

Definition 6 (Same-basis and mix-basis attack). *An attack on*

$$\mathcal{E}^{\otimes d} : (\mathbb{F}_2^n)^{\otimes d} \to (\mathbb{F}_2^m)^{\otimes d}$$

is called a same-basis attack if the bases chosen for the input and output spaces are the same; otherwise, a mix-basis attack.

This partition is crucial for calculating the propagation matrix of the composite function. Most modern ciphers are constructed from smaller component functions, so computing the whole transition matrix of the cipher should handle the propagation properties of the transition matrices of its components.

Consider a d-th order attack on $\mathcal{E}^{\otimes d} = \mathcal{E}_2^{\otimes d} \circ \mathcal{E}_1^{\otimes d} \circ \mathcal{E}_0^{\otimes d}$, *i.e.*, $\mathcal{E}^{\otimes d}$ divided into three parts. When we choose the same basis for the input and output spaces of $\mathcal{E}^{\otimes d}$, the attack is a same-basis one. According to Theorem 1, the transition matrix of $\mathcal{E}^{\otimes d}$ is the product of the transition matrices of $\mathcal{E}_2^{\otimes d}$ and $\mathcal{E}_1^{\otimes d}$ and $\mathcal{E}_0^{\otimes d}$. However, things are a bit more complicated for a mix-basis attack because different bases are used for different part ciphers. We have the following proposition.

Proposition 1 (Propagation of the mix-basis transition matrices). *For $\mathcal{E}^{\otimes d} = \mathcal{E}_2^{\otimes d} \circ \mathcal{E}_1^{\otimes d} \circ \mathcal{E}_0^{\otimes d}$, suppose that we select $[\alpha_u(x)]_{x,u}$ for the change-of-basis for the input space of $\mathcal{E}^{\otimes d}$ (it is also the input space of $\mathcal{E}_0^{\otimes d}$) and $[\beta_u(x)]_{x,u}$ for the output space of $\mathcal{E}^{\otimes d}$ (it is also the output space of $\mathcal{E}_2^{\otimes d}$). Denote the transition matrix of $\mathcal{E}^{\otimes d}$ under the two bases by $A^{\mathcal{E}^{\otimes d}}$. Then we have*

$$A^{\mathcal{E}^{\otimes d}} = A^{\mathcal{E}_2^{\otimes d}} A^{\mathcal{E}_1^{\otimes d}} A^{\mathcal{E}_0^{\otimes d}}$$

where $A^{\mathcal{E}_0^{\otimes d}}$ is the transition matrix of $\mathcal{E}_0^{\otimes d}$ under the same input and output bases $[\alpha_u(x)]_{x,u}$, $A^{\mathcal{E}_1^{\otimes d}}$ is the transition matrix of $\mathcal{E}_1^{\otimes d}$ under the input basis $[\alpha_u(x)]_{x,u}$ and output basis $[\beta_u(x)]_{x,u}$, and $A^{\mathcal{E}_2^{\otimes d}}$ is the transition matrix of $\mathcal{E}_2^{\otimes d}$ under the same input and output bases $[\beta_u(x)]_{x,u}$.

Proof. An illustration for this proof is provided in Figure 2. According to Equation 3.1,

$$A^{\mathcal{E}_0^{\otimes d}} = [\alpha_u(x)]_{x,u}^{-1}\ T^{\mathcal{E}_0^{\otimes d}}\ [\alpha_u(x)]_{x,u},$$

$$A^{\mathcal{E}_1^{\otimes d}} = [\beta_u(x)]_{x,u}^{-1}\ T^{\mathcal{E}_1^{\otimes d}}\ [\alpha_u(x)]_{x,u},$$

and

$$A^{\mathcal{E}_2^{\otimes d}} = [\beta_u(x)]_{x,u}^{-1}\ T^{\mathcal{E}_2^{\otimes d}}\ [\beta_u(x)]_{x,u}.$$

Therefore,

$$\begin{aligned}
&A^{\mathcal{E}^{\otimes d}} = [\beta_u(x)]_{x,u}^{-1}\ T^{\mathcal{E}^{\otimes d}}\ [\alpha_u(x)]_{x,u} \\
&= [\beta_u(x)]_{x,u}^{-1}\ T^{\mathcal{E}_2^{\otimes d}}\ [\beta_u(x)]_{x,u} \cdot [\beta_u(x)]_{x,u}^{-1}\ T^{\mathcal{E}_1^{\otimes d}}\ [\alpha_u(x)]_{x,u} \cdot [\alpha_u(x)]_{x,u}^{-1}\ T^{\mathcal{E}_0^{\otimes d}}\ [\alpha_u(x)]_{x,u}. \\
&= [\beta_u(x)]_{x,u}^{-1}\ T^{\mathcal{E}_2^{\otimes d}}\ T^{\mathcal{E}_1^{\otimes d}}\ T^{\mathcal{E}_0^{\otimes d}} [\alpha_u(x)]_{x,u}.
\end{aligned}$$

□

Corollary 1. *Suppose $\mathcal{E}^{\otimes d} = \mathcal{E}_{r-1}^{\otimes d} \circ \mathcal{E}_{r-2}^{\otimes d} \circ \cdots \circ \mathcal{E}_0^{\otimes d}$. Choose $r+1$ bases $[\alpha_u(x)]_{x,u}^{(i)}, 0 \leq i < r+1$, denote the transition matrix of $\mathcal{E}_i^{\otimes d}$ under the input basis $[\alpha_u(x)]_{x,u}^{(i)}$ and output basis $[\alpha_u(x)]_{x,u}^{(i+1)}$ by $A^{\mathcal{E}_i^{\otimes d}}$. Therefore, the transition matrix of $\mathcal{E}^{\otimes d}$ under the input basis $[\alpha_u(x)]_{x,u}^{(0)}$ and output space $[\alpha_u(x)]_{x,u}^{(r)}$ can be calculated by*

$$A^{\mathcal{E}^{\otimes d}} = A^{\mathcal{E}_{r-1}^{\otimes d}} A^{\mathcal{E}_{r-2}^{\otimes d}} \cdots A^{\mathcal{E}_0^{\otimes d}}.$$

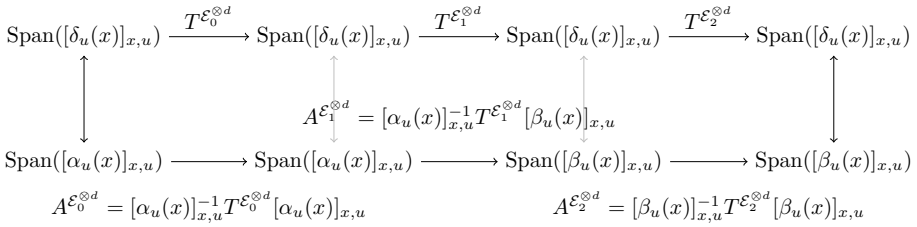

Fig. 2: The illustration of the proof for Proposition 1. Span($[\alpha_u(x)]_{x,u}$) represents that the space is spanned from $[\alpha_u(x)]_{x,u}$, and vectors are expressed as a linear expression of $[\alpha_u(x)]_{x,u}$.

3.2 Basis of First Order Spaces and Attacks

In this subsection, we enumerate several bases for the first-order spaces that have been used in previous geometric theory and introduce rules to generate new bases based on these existing bases.

Linear cryptanalysis. In [4], Beyne introduced geometric theory for the first time and applied it to linear cryptanalysis. The basis he chose for the linear cryptanalysis can be represented by

Basis 1 (Linear basis [4]) $\left[(-1)^{u^\top x}\right]_{x,u}$ (5)

Quasi-differential cryptanalysis. In [6], Beyne and Rijmen introduced the quasi-differential attack. The differential cryptanalysis is a second-order attack whose input and output spaces are second-order spaces. Thus, Beyne and Rijmen chose two bases for the two component spaces. The first basis is for the value, which is just the linear basis as Equation (5). The second basis is for the difference, which is the standard basis.

Basis 2 (Standard basis [6]) $[\delta_u(x)]_{x,u}$

Ultrametric integral cryptanalysis. In [8], Beyne and Verbauwhede introduced the ultrametric integral cryptanalysis to study the divisibility property. Note that in [7], Beyne and Verbauwhede also introduced the algebraic transition matrix for the integral attacks, but that attack works in $\mathbb{F}_2[\mathbb{F}_2^n]$, which is difficult to be combined with other basis whose space is based on $\mathbb{Q}$, so we omit it in this paper.

Basis 3 (Ultrametric integral basis [8]) $\left[(-1)^{\mathrm{wt}(u \oplus x)} u^x\right]_{x,u}$

Next, we introduce several rules that can generate new bases based on existing ones. These rules follow a simple fact that any 2^{nd} linearly independent vectors can serve as a set of basis for a d-th order space.

Table 1: Seven bases of the first order space concluded from previous geometric approach papers and induced from Rules 1, 2 and 3. The usage of their effects of these bases for the input and output have been shown in Equation (4).

Index	**Basis**	**Effect of input** $\alpha_u(x)$	**Effect of output** $\beta^{\star}_{\mathcal{E}(x)}(v)$
0	$[\delta_u(x)]_{x,u}$	$\delta_u(x)$	$\delta_{\mathcal{E}(x)}(v)$
1	$[(-1)^{u^\top x}]_{x,u}$	$(-1)^{u^\top x}$	$2^{-n}(-1)^{\mathcal{E}(x)^\top v}$
2	$[2^{-n}(-1)^{u^\top x}]_{x,u}$	$2^{-n}(-1)^{u^\top x}$	$(-1)^{\mathcal{E}(x)^\top v}$
3	$[u^x]_{x,u}$	u^x	$(-1)^{\mathrm{wt}(v\oplus\mathcal{E}(x))}\mathcal{E}^v(x)$
4	$[(-1)^{\mathrm{wt}(u\oplus x)}u^x]_{x,u}$	$(-1)^{\mathrm{wt}(u\oplus x)}u^x$	$\mathcal{E}^v(x)$
5	$[x^u]_{x,u}$	x^u	$(-1)^{\mathrm{wt}(v\oplus\mathcal{E}(x))}v^{\mathcal{E}(x)}$
6	$[(-1)^{\mathrm{wt}(u\oplus x)}x^u]_{x,u}$	$(-1)^{\mathrm{wt}(u\oplus x)}x^u$	$v^{\mathcal{E}(x)}$

Rule 1 (Inverse) *If* $[\alpha_u(x)]_{x,u}$ *is a set of basis,* $[\alpha_u(x)]^{-1}_{x,u}$ *is a set of basis.*

Rule 2 (Transpose) *If* $[\alpha_u(x)]_{x,u}$ *is a set of basis,* $[\alpha_u(x)]^{\top}_{x,u}$ *is a set of basis.*

Rule 3 (Scale) *If* $[\alpha_u(x)]_{x,u}$ *is a set of basis,* $[k\alpha_u(x)]_{x,u}$ *is a set of basis, where* $k \neq 0$ *belongs to the corresponding field, in this paper the field is* $\mathbb{Q}$*.*

According to these three rules, we obtain four more bases.

Basis 4 (Inverse of linear basis) $[2^{-n}(-1)^{u^\top x}]_{x,u}$

It is easy to check $[2^{-n}(-1)^{u^\top x}]_{x,u} \cdot [(-1)^{u^\top x}]_{x,u} = \text{Identity}$.

Basis 5 (Inverse of ultrametric integral basis) $[u^x]_{x,u}$

It is easy to check $[u^x]_{x,u} \cdot [(-1)^{\mathsf{wt}(u\oplus x)}u^x]_{x,u} = \text{Identity}$.

Basis 6 (Transpose of ultrametric integral basis) $[(-1)^{\mathsf{wt}(u\oplus x)}x^u]_{x,u}$

Basis 7 (Inverse and transpose of ultrametric integral basis) $[x^u]_{x,u}$

When choosing specific bases for the input and output spaces, we can use Equation (4) to calculate the element in the corresponding transition matrix. The $\beta^{\star}_{\mathcal{E}^{\otimes d}(x)}(v)$ and $\alpha_u(x)$ calculated according to the matrix composed of the bases are called *effects*. We list them for the seven bases above in Table 1. These effects can help us quickly write the statistic *i.e.*, the element of the corresponding transition matrix, based on the chosen bases.

Combining these seven bases for the input and output spaces, 49 different attacks, including 7 same-basis and 42 mix-basis ones, are generated. We list them in Tables 6 and 7.

Remark. One may doubt if some of them can be called "attacks". For example, when choosing $[\delta_u(x)]_{x,u}$ for both input and output spaces, the statistic

$$A^{\mathcal{E}}_{v,u} = \sum_{x=u,\mathcal{E}(x)=v} 1$$

says nothing except $\mathcal{E}(u) = v$. Whether we should regard it as an attack depends on the definition of "attacks". On the one hand, considering $\mathcal{E}$ as a public permutation, knowing $\mathcal{E}(u) = v$ is indeed useful to distinguish $\mathcal{E}$ from a random permutation. On the other hand, when $\mathcal{E}$ is key-dependent, $A^{\mathcal{E}}_{v,u} = \sum_{x=u,\mathcal{E}(x)=v} 1$ means there is a deterministic invariant behavior of $\mathcal{E}$ independently of the key (practically, this statistic is always influenced by the secret key). Therefore, we still include such simple statistics as attacks. In addition, we note that some of the 49 attacks are actually identical, due to the bases $[(-1)^{u^\top x}]_{x,u}$ and $[2^{-n}(-1)^{u^\top x}]_{x,u}$. However, counting these duplicated attacks can bring convenience for us.

3.3 Basis of Higher Order Spaces and Attacks

For a d-th order attack, the input and output spaces are also d-th order. In theory, any 2^{dn} linearly independent vectors can serve as a set of bases and lead to a basis-based attack. However, a random basis is difficult to handle if it does not have a compact representation. Thus, inspired by the quasi-differential cryptanalysis [6], we generate a basis for higher-order spaces by the tensor product of first-order space bases.

Proposition 2 (Basis for $\mathbb{K}[(\mathbb{F}_2^n)^{\otimes d}]$). *For a d-th space $\mathbb{K}[(\mathbb{F}_2^n)^{\otimes d}]$, we choose bases for each of its d components, denoted by $[\alpha_u(x)]^{(i)}_{x,u}$. Then $\bigotimes_{0\le i<d}[\alpha_u(x)]^{(i)}_{x,u}$ is a basis of $\mathbb{K}[(\mathbb{F}_2^n)^{\otimes d}]$.*

Proof. This is from the calculation rules for the tensor product (see Section 2.2). Since $[\alpha_u(x)]^{(i)}_{x,u}$ spans to $\mathbb{K}[\mathbb{F}_2^n]$, $\bigotimes_{0\le i<d}[\alpha_u(x)]^{(i)}_{x,u}$ spans to $\bigotimes_{0\le i<d}\mathbb{K}[(\mathbb{F}_2^n)] = \mathbb{K}[(\mathbb{F}_2^n)^{\otimes d}]$.

To compute Equation (4), we need the inverses of the basis matrices. The proposition 3 gives a simple way of calculating.

Proposition 3 (Inverse of a higher order basis matrix). *Let $\bigotimes_{0\le i<d}[\alpha_u(x)]^{(i)}_{x,u}$ be a set of basis of $\mathbb{K}[(\mathbb{F}_2^n)^{\otimes d}]$. Then, $\bigotimes_{0\le i<d}\left([\alpha_u(x)]^{(i)}_{x,u}\right)^{-1}$ is the inverse of $\bigotimes_{0\le i<d}[\alpha_u(x)]^{(i)}_{x,u}$, which is also a set of bases of $\mathbb{K}[(\mathbb{F}_2^n)^{\otimes d}]$.*

Proof. This directly follows from the fact that the inverse matrix of $A \otimes B$ is $A^{-1} \otimes B^{-1}$.

Therefore, combining the seven first-order bases in Table 1, 7^d different d-th order bases are obtained. The 7^d different bases lead to 7^{2d} attacks including 7^d same-basis attacks and $7^{2d} - 7^d$ mix-basis attacks. Again, similar to the first-order case, not all of them look interesting, but we still see them as attacks to keep the theory intact. The effects of 49 bases for the second-order case are listed in Tables 8 and 9.

To quickly derive the statistic of an attack, we can use a similar method with first-order attacks. Either we can write all attacks according to the effects of the bases and check if some are interesting, or we can write the statistic we are interested in and see if there are proper bases that can lead to this attack.

3.4 Trail Search for Mix-Basis Attacks

Recalling Definition 1, trails are clustered to compute or approximate the transition matrix elements. For the mix-basis attacks following Proposition 1 and Corollary 1, transition matrices are calculated based on the corresponding input and output bases, which completely follows the same method as the same-basis attacks like the linear [4], quasi-differential attacks [6] and the ultramtric integral attacks [8].

In terms of the measurements, the case of mix-basis attacks is also the same as that of same-basis attacks. We can approximate the value of the statistic by adding the correlations of trails, or study the divisibility property by studying their 2-adic absolute values. In this paper, we do not have a specific rule for how to choose the measurements, but we try both to see if we can get interesting attacks.

4 Example Application I: A Simplified Ultrametric Integral Attack with a Better Basis

4.1 Revisiting the Ultrametric Integral Cryptanalysis from [8]

In [8], Beyne and Verbauwhede introduced the ultrametric integral cryptanalysis to describe the divisibility property. The divisibility property is a generalization of the integral property [22], interpolating between bits that sum to zero (divisibility by two) and saturated bits (divisibility by 2^{n-1} for 2^n inputs). Given $u \in \mathbb{F}_2^n$, Suppose $\mathbb{U} = \{y \in \mathbb{F}_2^n : y \preceq u, u \in \mathbb{F}_2^n\}$ is a structure of the plaintexts, the divisibility studies if

$$\sum_{y \in \mathbb{U}} \mathcal{E}^v(y) = \sum_{y \preceq u} \mathcal{E}^v(y) \equiv \mod 2^t. \tag{6}$$

To study it, Beyne and Verbauwhede chose the ultrametric integral basis as

$$[(-1)^{\mathsf{wt}(u \oplus x)} u^x]_{x,u}.$$

The change-of-basis matrix between the standard basis $[\delta_u(x)]_{x,u}$ is denoted by I, thus

$$[\delta_u(x)]_{x,u} = [(-1)^{\mathsf{wt}(u \oplus x)} u^x]_{x,u} I.$$

Each element in $\mathbb{U}$ can be expressed by a linear combination of bases in $[(-1)^{\mathsf{wt}(u\oplus x)}u^x]_{x,u} = (\mu_0, \mu_1, \ldots, \mu_{2^n-1})$, thus $\delta_{\mathbb{U}} = \sum_{y\in\mathbb{U}} \delta_y = \sum_{y\preceq u} \delta_y$ is

$$\delta_{\mathbb{U}} = \sum_{\nu\preceq u} 2^{\mathsf{wt}(u)-\mathsf{wt}(\nu)}\mu_\nu.$$

Let the transition matrix of $\mathcal{E}$ under the basis $[(-1)^{\mathsf{wt}(u\oplus x)}u^x]_{x,u}$ be $A^{\mathcal{E}}$, we have

$$A^{\mathcal{E}}_{v,\nu} = \delta_v^\top A^{\mathcal{E}} \delta_\nu = \delta_v^\top I\ T^{\mathcal{E}}\ I^{-1}\delta_\nu = (T^{\mathcal{E}}\mu_\nu)^v.$$

Thus, the corresponding summation of the ciphertext is

$$\sum_{y\preceq u} \mathcal{E}^v(y) = \sum_{y\in\mathbb{U}} (T^{\mathcal{E}}\delta_y)^v = (T^{\mathcal{E}}\delta_{\mathbb{U}})^v = \sum_{\nu\preceq u} 2^{\mathsf{wt}(u)-\mathsf{wt}(\nu)}(T^{\mathcal{E}}\mu_\nu)^v = \sum_{\nu\preceq u} 2^{\mathsf{wt}(u)-\mathsf{wt}(\nu)} A^{\mathcal{E}}_{v,\nu}.$$

There is equivalence between $\sum_{y\preceq u} \mathcal{E}^v(y) \equiv 0 \bmod 2^t$ $(t \le \mathsf{wt}(u))$ and $\left|\sum_{y\preceq u} \mathcal{E}^v(y)\right|_2 \le 2^{-t}$. According to the ultrametric triangle inequality of the 2-adic absolute value $|x+y|_2 \le \max\{|x|_2, |y|_2\}$,

$$\left|\sum_{y\preceq u} \mathcal{E}^v(y)\right|_2 = \left|\sum_{\nu\preceq u} 2^{\mathsf{wt}(u)-\mathsf{wt}(\nu)} A^{\mathcal{E}}_{v,\nu}\right|_2 \le \max_{\nu\preceq u} 2^{\mathsf{wt}(\nu)-\mathsf{wt}(u)} \left|A^{\mathcal{E}}_{v,\nu}\right|_2.$$

Thus, if we prove $\max_{\nu\preceq u} 2^{\mathsf{wt}(\nu)-\mathsf{wt}(u)} \left|A^{\mathcal{E}}_{v,\nu}\right|_2 \le 2^{-t}$, we verify that $\sum_{y\preceq u} \mathcal{E}^v(y) \equiv 0 \bmod 2^t$. For those ν satisfying $\mathsf{wt}(\nu) \le \mathsf{wt}(u) - t$, $\max_{\nu\preceq u} 2^{\mathsf{wt}(\nu)-\mathsf{wt}(u)} \left|A^{\mathcal{E}}_{v,\nu}\right|_2 \le 2^{-t}$ is already valid. For ν satisfying $\mathsf{wt}(\nu) > \mathsf{wt}(u) - t$, we need to verify that $\left|A^{\mathcal{E}}_{v,\nu}\right|_2$ is divisible by $2^{2t-\mathsf{wt}(u)}$ which can be done by searching for trails. That is, the divisibility in Equation (6) is studied in an indirect way. The reason is that the vector corresponding to the input set $\mathbb{U}$ is not any column index of the matrix derived from the ultrametric integral basis.

4.2 A Simplified Version of Ultrametric Integral Cryptanalysis

Using two different bases for the input and output, we can derive a matrix whose (v,u)-element is exactly $\sum_{y\preceq u} \mathcal{E}^v(y)$, *i.e.*, the transition matrix under the two bases is $A^{\mathcal{E}}$ that satisfies

$$A^{\mathcal{E}}_{v,u} = \sum_{y\preceq u} \mathcal{E}^v(y).$$

Note that $A^{\mathcal{E}}_{v,u} = \sum_{x\preceq u} \mathcal{E}^v(x) = \sum_{x\in\mathbb{F}_2^n} u^x \mathcal{E}^v(x)$. According to Table 1, if we want a term u^x, we can choose the basis

$$[u^x]_{x,u}$$

for the input space. For $\mathcal{E}^v(y)$, we can choose the basis

$$[(-1)^{\mathsf{wt}(u\oplus x)}u^x]_{x,u}$$

for the output space. The (v,u)-element of the transition matrix under these two bases is

$$A^{\mathcal{E}}_{v,u} = \sum_{x\in\mathbb{F}_2^n} u^x \mathcal{E}^v(x) = \sum_{x\preceq u} \mathcal{E}^v(y).$$

Since the bases for the input and output space are different, the simplified version of ultrametric integral cryptanalysis belongs to the mix-basis attacks. To characterize the propagation of the transition matrices, we divide an r-round cipher $\mathcal{E}$ into three parts

$$\mathcal{E} = \mathcal{E}_2 \circ \mathcal{E}_1 \circ \mathcal{E}_0,$$

where $\mathcal{E}_0$, $\mathcal{E}_1$ and $\mathcal{E}_2$ are three consecutive parts of $\mathcal{E}$ whose number of rounds are respectively r_0, r_1 and r_2 that satisfies $r_0 + r_1 + r_2 = r$.

For $\mathcal{E}_0$, the transition matrix is obtained in the same-base attack framework under the basis $[u^x]_{x,u}$ for the input and output spaces. Thus, the (v,u)-element of the transition $A^{\mathcal{E}_0}$ is

$$A^{\mathcal{E}_0}_{v,u} = \sum_{x\in\mathbb{F}_2^n} u^x \cdot (-1)^{\mathsf{wt}(v\oplus\mathcal{E}(x))}\mathcal{E}^v(x) = \sum_{x\preceq u}(-1)^{\mathsf{wt}(v\oplus\mathcal{E}(x))}\mathcal{E}^v(x). \tag{7}$$

For $\mathcal{E}_2$, the transition matrix is also obtained in the same-basis attack framework under the basis $[(-1)^{\mathsf{wt}(u\oplus x)}u^x]_{x,u}$ for both the input and output spaces. Thus, the (v,u)-element of the transition $A^{\mathcal{E}_2}$ is

$$A^{\mathcal{E}_2}_{v,u} = \sum_{x\in\mathbb{F}_2^n} (-1)^{\mathsf{wt}(u\oplus x)}u^x \cdot \mathcal{E}^v(x) = \sum_{x\preceq u}(-1)^{\mathsf{wt}(u\oplus x)}\mathcal{E}^v(x). \tag{8}$$

Note that $A^{\mathcal{E}_2}$ is just the transition matrix of the original ultrametric integral cryptanalysis derived by Beyne and Verbauwhede.

For $\mathcal{E}_1$, the transition matrix is derived from the same bases for $\mathcal{E}$, so the (v,u)-element of this transition matrix is

$$A^{\mathcal{E}_1}_{v,u} = \sum_{x\in\mathbb{F}_2^n} u^x \cdot \mathcal{E}^v(x) = \sum_{x\preceq u}\mathcal{E}^v(x). \tag{9}$$

Finally,

$$A^{\mathcal{E}} = A^{\mathcal{E}_2}A^{\mathcal{E}_1}A^{\mathcal{E}_0}.$$

The automatic search can be done with the same methods introduced in Section 2.4 and [8], with the $|\cdot|_2$ being the measurement. The targets of the original and our simplied ultrametric integral cryptanalysis are the same, our method cannot find more distinguishers than the original one. However, the simplified version does not require any more techniques in the automatic search. We give an example on how to use the automatic search model for the simplified ultrametric integral cryptanalysis and re-find the ultrametric integral distinguishers for 9-round `PRESENT` in Appendix D.

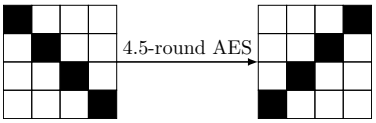

Fig. 3: An example of the multiple-of-8 property of the 5-round AES without the last MixColumn. The black squares denote the subspaces. Encrypt all plaintexts whose active bytes are in a diagonal, the number of ciphertext pairs whose active bytes in an anti-diagonal is always a multiple of 8 [17].

5 Example Application II: Multiple-of-n Property

The multiple-of-8 property of the AES was found for the first time by Grassi, Rechberger, and Rønjom [17], which is the first key-independent distinguisher on the 5-round AES. This property shows that there exist two linear spaces $\mathbb{V}$ and $\mathbb{W}$ of $\mathbb{F}_2^{128}$ satisfying: for any coset of $\mathbb{V}$, say $c \oplus \mathbb{V}$, the number of distinct pairs of elements $x, x', x \neq x'$ in $c + \mathbb{V}$ such that $\mathcal{E}(x)$ and $\mathcal{E}(x')$ belong to the same coset of $\mathbb{W}$ is always divisible of 8, where $\mathcal{E}$ is the 5-round AES in [17] (an example is given in Figure 3). The original proof by Grassi *et al.* was done by a detailed case-by-case method, and it was believed by Grassi *et al.* that the maximum branch number of the AES MixColumn matrix was crucial for this property. However, Boura, Canteaut, and Coggia later pointed out that the maximum branch number is not necessary for this property by proposing a compact general proof framework [14]. For different ciphers, the multiple-of-8 property may be generalized to multiple-of-n property. For example, `SKINNY` has multiple-of-2^{h-1} property, where $h \in \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14\}$ according to different subspace trails found by the methods of [24]. However, since this proof heavily relies on the subspace trail, the rounds for `SKINNY`'s multiple-of-n property still stop at 5 rounds.

5.1 Geometric Approach for Multiple-of-n Properties

In this subsection, we show how to apply the geometric approach to the multiple-of-n property. For a cipher $\mathcal{E} : \mathbb{F}_2^n \to \mathbb{F}_2^n$, we assume $c = 0$ for $c \oplus \mathbb{V}$ and consider plaintexts in $\{x \in \mathbb{F}_2^n : x \preceq u\}$ and ciphertexts in $\{\mathcal{E}(x) \in \mathbb{F}_2^n : \mathcal{E}(x) \preceq v\}$ for some u and v. Then we find a proper statistic for this property. Since the plaintexts are chosen in a subspace, and pairs are combined with these plaintexts, this should be a second-order attack. The statistic used in this attack can be

$$A^{\mathcal{E}}_{v_0||v_1,u_0||u_1} = \sum_{\substack{x_0 \preceq u_0, x_1 \preceq u_1 \\ \mathcal{E}(x_1) \preceq v_1}} v_0^{\mathcal{E}}(x_0) = \sum_{x_0 \in \mathbb{F}_2^n, x_1 \in \mathbb{F}_2^n} \underbrace{u_0^{x_0} u_1^{x_1}}_{\text{effect of input basis}} \underbrace{v_0^{\mathcal{E}(x_0)} v_1^{\mathcal{E}(x_1)}}_{\text{effect of output basis}} \tag{10}$$

with $v_0 = \mathbf{1}$ ($\mathbf{1}$ is a n-bit string with elements being 1) and $u_0 = u_1$. The $v_0 = \mathbf{1}$ condition is required because the output values in the multiple-of-n property are free, so no restrictions should be placed on them. Condition $u_0 = u_1$ is required as the two plaintexts should come from the same subspace.

The above statistic with the two conditions exactly counts the number of pairs that satisfy

1. the two inputs $(x_0, x_0 \oplus x_1)$ (recall again that x_1 represents the difference) are from a subspace $\{x \in \mathbb{F}_2^n : x \preceq u_0 = u_1\}$,
2. the difference of the outputs $\mathcal{E}(x_1)$ fall into a subspace $\{x \in \mathbb{F}_2^n : x \preceq v_1\}$ whatever $\mathcal{E}(x_0)$ is (as $\mathcal{E}(x_0) \preceq \mathbf{1}$).

By checking the effects of bases in Tables 8 and 9, the input basis can be chosen as $[u^x]_{x,u} \otimes [u^x]_{x,u}$, and the output basis chosen as $[(-)^{\mathsf{wt}(x \oplus u)} x^u]_{x,u} \otimes [(-)^{\mathsf{wt}(x \oplus u)} x^u]_{x,u}$.

Note that the multiple-of-n property is to count the number of distinct pairs while the nunber counted by the statistic $A^{\mathcal{E}}_{v_0||v_1,u_0||u_1}$ is the ordered pairs (*i.e.*, (a, b) and (b, a) are counted twice). Besides, the trivial pairs such as (a, a) is also counted. We have the following proposition.

Proposition 4. *When $u_0 = u_1 = u$, $v_0 = \mathbf{1}$, and $2^{\mathsf{wt}(u)-1} \equiv 0 \bmod n$,*

$$A^{\mathcal{E}}_{v_0||v_1,u_0||u_1} = \sum_{\substack{x_0 \preceq u_0, x_1 \preceq u_1 \\ \mathcal{E}(x_1) \preceq v_1}} v_0^{\mathcal{E}}(x_0) \equiv 0 \bmod 2n$$

is equivalent to

$$|\{(p_0, p_1) : p_0 \neq p_1, p_0 \preceq u, p_1 \preceq u, \mathcal{E}(p_0) \oplus \mathcal{E}(p_1) \preceq v_1\}| \equiv 0 \bmod n.$$

Proof. Among the $A^{\mathcal{E}}_{v_0||v_1,u_0||u_1}$ pairs, there are $2^{\mathsf{wt}(u)}$ trivial ones. After excluding the trivial ones, there are $A^{\mathcal{E}}_{v_0||v_1,u_0||u_1} - 2^{\mathsf{wt}(u)}$ non-trivial pairs. Thus, $(A^{\mathcal{E}}_{v_0||v_1,u_0||u_1} - 2^{\mathsf{wt}(u)})/2$ is the number of distinct pairs. Since $A^{\mathcal{E}}_{v_0||v_1,u_0||u_1} \equiv 0 \bmod 2n$, we have $A^{\mathcal{E}}_{v_0||v_1,u_0||u_1} = p \times 2n$ for a certain p. Therefore,

$$(A^{\mathcal{E}}_{v_0||v_1,u_0||u_1} - 2^{\mathsf{wt}(u)})/2 = (p \times 2n - 2^{\mathsf{wt}(u)})/2 = p \times n - 2^{\mathsf{wt}(u)-1}.$$

Thus, $2^{\mathsf{wt}(u)-1} \equiv 0 \bmod n$ leads to $(A^{\mathcal{E}}_{v_0||v_1,u_0||u_1} - 2^{\mathsf{wt}(u)})/2 \equiv 0 \bmod n$. □

Similar to ultrametric integral cryptanalysis [8], $A^{\mathcal{E}}_{v_0||v_1,u_0||u_1} \equiv 0 \mod 2^t$ is equivalent to $|A^{\mathcal{E}}_{v_0||v_1,u_0||u_1}|_2 \leq 2^{-t}$. This can be done by searching trails as described in Section 2.4.

5.2 Automatic Search for Multiple-of-n Properties: Application to SKINNY-64

Due to the page limitation, the specifications of SKINNY-64 are provided in Appendix A.2. Next, we introduce our search model. Consider r rounds of SKINNY-64, we first divide it into three parts, as $\mathcal{E} = \mathcal{E}_2 \circ \mathcal{E}_1 \circ \mathcal{E}_0$, where $\mathcal{E}_1$ contains only one layer of Sboxes. For $\mathcal{E}_1$, the same bases are applied to compute the

statistic and a corresponding matrix is obtained. For $\mathcal{E}_0$, the input and output bases are chosen as $[u^x]_{x,u} \otimes [u^x]_{x,u}$, so a same-basis attack is generated for $\mathcal{E}_0$. The statistic is

$$\begin{aligned} A^{\mathcal{E}_0}_{v_0||v_1,u_0||u_1} &= \sum_{x_0\in\mathbb{F}_2^n, x_1\in\mathbb{F}_2^n} u_0^{x_0} u_1^{x_1} (-1)^{\mathsf{wt}(v_0\oplus\mathcal{E}(x_0))} \mathcal{E}(x_0)^{v_0} (-1)^{\mathsf{wt}(v_1\oplus\mathcal{E}(x_1))} \mathcal{E}(x_1)^{v_1} \\ &= \sum_{\substack{x_0\preceq u_0, x_1\preceq u_1 \\ \mathcal{E}(x_0)\succeq v_0, \mathcal{E}(x_1)\succeq v_1}} (-1)^{\mathsf{wt}(v_0\oplus\mathcal{E}(x_0))} (-1)^{\mathsf{wt}(v_1\oplus\mathcal{E}(x_1))} \end{aligned} \tag{11}$$

For $\mathcal{E}_2$, $[(-1)^{\mathsf{wt}(u\oplus x)} x^u]_{x,u} \otimes [(-1)^{\mathsf{wt}(u\oplus x)} x^u]_{x,u}$ are chosen as the bases, thus another same-basis attack is derived for $\mathcal{E}_2$, where the statistic is

$$\begin{aligned} A^{\mathcal{E}_2}_{v_0||v_1,u_0||u_1} &= \sum_{x_0\in\mathbb{F}_2^n, x_1\in\mathbb{F}_2^n} (-1)^{\mathsf{wt}(u_0\oplus x_0)} u_0^{x_0} (-1)^{\mathsf{wt}(u_1\oplus x_1)} u_1^{x_1} \mathcal{E}(x_0)^{v_0} \mathcal{E}(x_1)^{v_1} \\ &= \sum_{\substack{x_0\succeq u_0, x_1\succeq u_1 \\ \mathcal{E}(x_0)\preceq v_0, \mathcal{E}(x_1)\preceq v_1}} (-1)^{\mathsf{wt}(u_0\oplus x_0)} (-1)^{\mathsf{wt}(u_1\oplus x_1)} \end{aligned} \tag{12}$$

According to Proposition 2.4, the product of the transition matrices of $\mathcal{E}_0, \mathcal{E}_1$ and $\mathcal{E}_2$ under corresponding bases is the transition matrix of $\mathcal{E}$ related to the statistic of Equation (10).

To use the automatic search, we need to generate the matrices for operations in $\mathcal{E}_0, \mathcal{E}_1$ and $\mathcal{E}_2$. For Sboxes in SC and LBox in the MC (the MixColumn operation of SKINNY-64 can be split into 16 parallel 4-bit Sboxes, which are called LBox, see Appendix A.2 for more details), given u_0, u_1, v_0, v_1, the corresponding correlation (Definition 4) is calculated with Equations (11), (10) and (12) for $\mathcal{E}_0$, $\mathcal{E}_1$ and $\mathcal{E}_2$, respectively. The 2-adic absolute values of the correlations of u_0, u_1, v_0, v_1 is calculated and described with the automatic search tool languagae, as a classical way.

For the bit-permutations like SR, a direct variable changes works. For AC, the affected bits can be seen as a 1-bit Sbox, thus the corresponding propagation rules can be derived. For ART, the round tweakey are seen as random constants, thus the elements in the corresponding matrices are tweakey-dependent. For $\mathcal{E}_0$ ($\mathcal{E}_2$), according to Equation (11) (Equation (12)), the transition matrix of ART is ($u_0||u_1$ is the index of columns, $v_0||v_1$ is the index of rows, k is the tweakey/constant)

$$A^{\mathsf{ART}} = \begin{bmatrix} (-1)^k & 0 & 0 & 0 \\ 0 & (-1)^k & 0 & 0 \\ k & 0 & 1 & 0 \\ 0 & k & 0 & 1 \end{bmatrix} \quad \left(A^{\mathsf{ART}} = \begin{bmatrix} (-1)^k & 0 & k & 0 \\ 0 & (-1)^k & 0 & k \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \right)$$

A^{AC} is a special case of A^{ART} where k is known. Propagation rules of A^{ART} are obtained when considering all possible key values since the key is unknown.

Table 2: The multiple-of-n property on SKINNY-64 from 6 to 10 rounds. "1" in the input/output means the corresponding bits of the input/output value-difference pairs can be any value.

Rnd.	Input/Output Value-Difference Pairs	Multiple -of-$2n$(-n)	Config.
6	(0fff′ffff′ffff′ffff, 0fff′ffff′ffff′ffff) ↓ (ffff′ffff′ffff′ffff, ffff′f0ff′ffff′ffff)	$2^{47}(2^{46})$	$2+1+3$
7	(0fff′ffff′ffff′ffff, 0fff′ffff′ffff′ffff) ↓ (ffff′ffff′ffff′ffff, ffff′f0ff′ffff′ffff)	$2^{42}(2^{41})$	$3+1+3$
8	(0fff′ffff′ffff′ffff, 0fff′ffff′ffff′ffff) ↓ (ffff′ffff′ffff′ffff, ffff′f0ff′ffff′ffff)	$2^{29}(2^{28})$	$4+1+3$
9	(0fff′ffff′ffff′ffff, 0fff′ffff′ffff′ffff) ↓ (ffff′ffff′ffff′ffff, ffff′f0ff′ffff′ffff)	$2^{17}(2^{16})$	$4+1+4$
10	(0fff′ffff′ffff′ffff, 0fff′ffff′ffff′ffff) ↓ (ffff′ffff′ffff′ffff, ffff′f0ff′ffff′ffff)	$2^{8}(2^{7})$	$4+1+5$

Results. The 10-round SKINNY-64 has a multiple-of-2^7 property, which has surpassed the previous best 5-round results. We list the results for SKINNY-64 from 6 to 10 rounds in Table 2

Restrictions of our method. There are two main restrictions of our automatic search method compared to previous ones [17,14]. The first is that we cannot set the inactive cell values as non-zero constants, currently they are always set as zero. The second is that the current search is limited to 4-bit cell ciphers such as SKINNY-64. For larger ones such as AES, the search needs to handle the propagations from a 16-bit vector to another. Current automatic search methods have usually difficulties to trace such large-scale propagations.

6 Example Application III: First Order Multiple-of-n Property

This section studies the divisibility property of messages that follow certain subspace trails. Such a property is very similar to the multiple-of-n property in Section 5, but we do it as a first-order attack. Concretely, we study the property that for $\mathcal{E} : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2^n$ and a pair of u and v, if there exists a number n that satisfies

$$|\{x \in \mathbb{F}_2^n : x \preceq u, \mathcal{E}(x) \preceq v\}| \equiv 0 \bmod n.$$

When $n \geq 2$, we obtain an interesting property of $\mathcal{E}$ as in the random case the property holds with a proportion of $1/n$.

Similar to the second-order multiple-of-n property, we study the following statistic:

$$A^{\mathcal{E}}_{v,u} = \sum_{x \preceq u, \mathcal{E}(x) \preceq v} 1 = \sum_{x \in \mathbb{F}_2^n} \underbrace{u^x}_{\substack{\text{effect of} \\ \text{input basis}}} \underbrace{v^{\mathcal{E}(x)}}_{\substack{\text{effect of} \\ \text{output basis}}}$$

Checking Table 1, the input basis is $[u^x]_{x,u}$, and the output basis is $[(-1)^{\mathsf{wt}(u \oplus x)} x^u]_{x,u}$, which is a mix-basis attack. The cipher will be divided into three parts where the middle round contains one layer of Sboxes. The first and last parts are two same-basis attacks with bases being $[u^x]_{x,u}$ and $[(-1)^{\mathsf{wt}(u \oplus x)} x^u]_{x,u}$, respectively. The two statistics are respectively

$$A^{\mathcal{E}_0}_{v,u} = \sum_{x \preceq u} (-1)^{\mathsf{wt}(\mathcal{E}_0(x) \oplus v)} \mathcal{E}(x)^v$$

and

$$A^{\mathcal{E}_2}_{v,u} = \sum_{x \succeq u} (-1)^{\mathsf{wt}(x \oplus u)} v^{\mathcal{E}(x)}$$

The automatic search model is constructed based on matrices of the three parts.

To study if $A^{\mathcal{E}}_{v,u}$ is divisible by 2^t, we check if $|A^{\mathcal{E}}_{v,u}|_2 \leq 2^{-t}$, this can be done with proving that there is no trail whose correlation is larger than 2^{-t}.

6.1 Application to SKINNY-64

Applying this statistic to SKINNY-64 is almost identical to Section 5. The Sboxes and Lboxes are described with 16×16 matrices. For each input and output pair (u, v), the corresponding 2-adic absolute value is recorded by variables. For key/constant XOR operations, we also regard them as 1-bit Sboxes. The transition matrices of the key XOR for $\mathcal{E}_0$ and $\mathcal{E}_2$ are respectively

$$A^{\mathsf{ART}} = \begin{bmatrix} (-1)^k & 0 \\ k & 1 \end{bmatrix} \text{(for } \mathcal{E}_0\text{)} \quad \text{and} \quad A^{\mathsf{ART}} = \begin{bmatrix} (-1)^k & k \\ 0 & 1 \end{bmatrix} \text{(for } \mathcal{E}_2\text{)}.$$

Results. The longest first order multiple-of-n property reaches 11 rounds for SKINNY-64, as shown in Table 3. The 11-round SKINNY-64 has a first order multiple-of-2 property, which has the same length as the longest integral distinguishers [16].

7 Example Application IV: Differential-Linear Cryptanalysis

Differential-linear (DL) cryptanalysis was originally proposed by Langford and Hellman in 1994 [23]. In this attack, a cipher $\mathcal{E}$ is decomposed into two subciphers as $\mathcal{E} = \mathcal{E}_1 \circ \mathcal{E}_0$, where a differential for $\mathcal{E}_0$ and a linear approximation for $\mathcal{E}_1$

Table 3: The first order multiple-of-n property of SKINNY-64 from 6 to 11 rounds. "1" in the input/output means the corresponding bits of the input/output value can be any value.

Rnd.	Input/Output Value	Multiple-of-n	Configure
6	ffff′ffff′ffff′fff0 → ffff′ffff′ffff′fff0	2^{47}	$3+1+2$
7	ffff′ffff′ffff′fff0 → ffff′ffff′ffff′fff0	2^{35}	$3+1+3$
8	ffff′ffff′ffff′fff0 → ffff′ffff′ffff′fff0	2^{26}	$4+1+3$
9	ffff′ffff′ffff′fff0 → ffff′ffff′ffff′fff0	2^{13}	$4+1+4$
10	ffff′ffff′ffff′fff0 → ffff′ffff′ffff′fff0	2^{5}	$4+1+5$
11	ffff′ffff′ffff′fff0 → ffff′f0ff′ffff′ffff	2^{1}	$3+1+7$

are considered. The bias of this DL approximation can be estimated accordingly under some independence assumptions.

As pointed out in [9], experiments are required to verify the estimated bias when possible because the underlying assumptions may fail. A closed formula for the DL bias, from Blondeau, Leander and Nyberg [11], is given under the sole assumption that $\mathcal{E}_0$ and $\mathcal{E}_1$ are independent. Let $\varepsilon[\delta \xrightarrow{\mathcal{E}_0} \gamma]$ denote the correlation of a DL distinguisher over $\mathcal{E}_0$ with the input difference δ and output mask γ, and $c[\theta \xrightarrow{\mathcal{E}_1} \lambda]$ denote the linear correlation with input and output masks θ and λ, respectively, over $\mathcal{E}_1$. Then, based on the independence assumption between $\mathcal{E}_0$ and $\mathcal{E}_1$, a DL distinguisher over $\mathcal{E} = \mathcal{E}_1 \circ \mathcal{E}_0$ with input difference δ and output mask λ has the exact correlation

$$\varepsilon[\delta \xrightarrow{\mathcal{E}} \lambda] = \sum_{\gamma} \varepsilon[\delta \xrightarrow{\mathcal{E}_0} \gamma] c^2[\gamma \xrightarrow{\mathcal{E}_1} \lambda]. \tag{13}$$

Recently, more new methods to estimate the DL bias have been proposed, for example, Bar-On *et al.* proposed the Differential-Linear Connectivity Table (DLCT) [1], Liu *et al.* introduced the differential algebraic transform form (DATF) to approximate the bias [25], Hadipour, Derbez and Eichlseder generalized the DLCT to more rounds [18], and Peng *et al.* combined the truncated differential for a preciser estimation on the DL bias [28].

7.1 Closed Formula without Independence Assumption

Using our notations, the DL approximation over a cipher $\mathcal{E}$ with input difference δ and output mask λ can be described by the following statistic

$$A^{\mathcal{E}}_{v_0||v_1,u_0||u_1} = 2^{-n} \sum_{x_0 \in \mathbb{F}_2^n, x_1 = u_1} (-1)^{u_0^\top x_0 \oplus v_0^\top \mathcal{E}(x_0) \oplus v_1^\top \mathcal{E}(x_1)}. \tag{14}$$

Where $u_0 = v_0 = 0$, $u_1 = \delta$ and $v_1 = \lambda$. Indeed, after replacing u_0, v_0, u_1, v_1 with $0, 0, \delta, \lambda$, the above equation becomes

$$\varepsilon[\delta \xrightarrow{\mathcal{E}} \lambda] = A^{\mathcal{E}}_{0||\lambda,0||\delta} = 2^{-n} \sum_{x_0 \in \mathbb{F}_2^n, x_1 = \delta} (-1)^{v_1^\top \mathcal{E}(x_1)}.$$

By checking the effects in Tables 8 and 9, Equation (14) can be obtained with the geometric approach with the input basis $[(-1)^{u^\top x}]_{x,u} \otimes [\delta_u(x)]_{x,u}$ and output basis $[(-1)^{u^\top x}]_{x,u} \otimes [(-1)^{u^\top x}]_{x,u}$.

Therefore, we can treat the DL attacks as a mix-basis attack. We first divide $\mathcal{E}$ into three parts as $\mathcal{E} = \mathcal{E}_2 \circ \mathcal{E}_1 \circ \mathcal{E}_0$. For $\mathcal{E}_0$, the quasi-differential attack [6] is applied, and the transition matrix is denoted by $A^{\mathcal{E}_0}$. For $\mathcal{E}_1$, Equation (14) is used as the statistic, and the transition matrix is denoted by $A^{\mathcal{E}_1}$. For $\mathcal{E}_2$, the statistic derived with the same basis $[(-1)^{u^\top x}]_{x,u} \otimes [(-1)^{u^\top x}]_{x,u}$ for the input/output spaces is

$$A^{\mathcal{E}_2}_{v_0||v_1,u_0||u_1} = 2^{-2n} \sum_{x_0 \in \mathbb{F}_2^n, x_1 \in \mathbb{F}_2^n} (-1)^{u_0^\top x_0 \oplus v_0^\top \mathcal{E}(x_0) \oplus u_1^\top x_1 \oplus v_1^\top \mathcal{E}(x_1)}$$

The transition matrix of $\mathcal{E}$ with input basis $[(-1)^{u^\top x}]_{x,u} \otimes [\delta_u(x)]_{x,u}$ and output basis $[(-1)^{u^\top x}]_{x,u} \otimes [(-1)^{u^\top x}]_{x,u}$ is calculated by

$$A^{\mathcal{E}} = A^{\mathcal{E}_2} \circ A^{\mathcal{E}_1} \circ A^{\mathcal{E}_0}$$

Setting $u_0 = v_0 = 0$, $u_1 = \delta$ and $v_1 = \lambda$, we get

$$\varepsilon[\delta \xrightarrow{\mathcal{E}} \lambda] = A^{\mathcal{E}}_{0||\lambda,0||\delta} = \sum_{\theta_0||\theta_1,\gamma_0||\gamma_1} A^{\mathcal{E}_2}_{0||\lambda,\theta_0||\theta_1} \, A^{\mathcal{E}_1}_{\theta_0||\theta_1,\gamma_0||\gamma_1} \, A^{\mathcal{E}_0}_{\gamma_0||\gamma_1,0||\delta} \tag{15}$$

Equation (15) can be seen as the closed formula for the DL approximation correlation. Inherent from the geometric approach, such a formula holds without independence assumptions. Using the automatic search tools to trace all trails derived from Equation (15), we can get the exact correlation.

When treating $\mathcal{E}_1 \circ \mathcal{E}_0$ as a whole part and considering $\mathcal{E}_1 \circ \mathcal{E}_0$ and $\mathcal{E}_2$ as two independent parts, we can get the same Blondeau-Leander-Nyberg formula from Equation (15). The details is shown in Appendix E.

7.2 Automatic Search for DL Approximation

Like previous applications, it is easy to develop an automatic search model to look for DL distinguishers on a cipher. For a given input difference δ and an output mask λ, we can use trails to approximate Equation (14). If we can exhaust all possible trails, the sum of all trail correlations is the exact DL approximation.

In [18], Hadipour *et al.* extended the DLCT to cover more rounds to give an efficient and precise method to estimate the correlation of DL approximations. They applied the method to the block cipher `SIMECK` and obtained the currently

Table 4: Three deterministic DL approximations of SIMECK found by Hadipour *et al.* [18].

Cipher	Round	Input Diff	Output Mask	Cor.
SIMECK-32	7	00001000	00000400	1
SIMECK-48	8	000000020000	000000010000	1

best-known DL distinguishers. Among the DL distinguishers they found, there are two deterministic DL approximations, one for SIMECK-32 and one for SIMECK-48, as shown in Table 4. However, as Hadipour *et al.*'s model was set based on the classical assumption that the consecutive rounds are independent, it is difficult to know if these deterministic DL approximations hold for all the key values. This is actually a challenge for almost all classical cryptanalysis methods. The geometric approach, as shown in previous applications, inherently works well without independence assumptions, as long as we can exhaust all trails.

We set the automatic search tools for the two DL distinguishers. Our automatic search model is able to exhaust all trails for the three DL approximations, thus calculates out the exact correlations of them. According to our search results, the sum of correlations of trails with non-zero masks for the values (which means the concrete key values would affect the final correlation) is always zero. The sum of correlations of trails with zero masks for the values (which means the key values would not affect the final correlation) is finally 1. Therefore, we confirm that the two DL approximations have exactly 1 correlation, without being affected by the key bits.

8 Conclusion

This paper extends Beyne's geometric approach by allowing using two different bases for the input and output spaces. We utilized three previously known bases and generated four new ones according to some simple rules. Based on these seven bases, we defined a family of *basis-based* attacks. For a d-th order, the seven bases lead to 7^{2d} attacks. The basis-based attacks provide a systematic way to generate new ones rather than the classical intuitive method. Our extension makes the geometric approach more flexible and able to describe/predict more types of attacks. Inherent to the geometric approach, all basis-based attacks can be studied with a similar automatic search method. The core is to track the propagation trails and estimate the correlations according to certain measurements. We provided four example applications to show how to take some basis-based attacks into practice, including a simplified ultrametric integral cryptanalysis, multiple-of-n properties for the second-order and first-order attacks, and finally, the differential-linear attacks.

There are many future works. For example, one can explore how to quickly check all these basis-based attacks and identify the most threatening one for a

certain cipher. Besides, Corollary 1 is not really used in this paper, it is interesting to study how to find a "best basis chain" that can connect the bases for the input and output spaces of each round that brings the best dominant trail [8, Theorem 2.2], which can reduce the search burden significantly. Finally, it would be interesting to study more possibilities of the bases, in addition to the ones presented by this paper.

References

1. Bar-On, A., Dunkelman, O., Keller, N., Weizman, A.: DLCT: A new tool for differential-linear cryptanalysis. In: Ishai, Y., Rijmen, V. (eds.) Advances in Cryptology - EUROCRYPT 2019 - 38th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Darmstadt, Germany, May 19-23, 2019, Proceedings, Part I. Lecture Notes in Computer Science, vol. 11476, pp. 313–342. Springer (2019). https://doi.org/10.1007/978-3-030-17653-2_11, https://doi.org/10.1007/978-3-030-17653-2_11
2. Beaulieu, R., Shors, D., Smith, J., Treatman-Clark, S., Weeks, B., Wingers, L.: The SIMON and SPECK lightweight block ciphers. In: Proceedings of the 52nd Annual Design Automation Conference, San Francisco, CA, USA, June 7-11, 2015. pp. 175:1–175:6. ACM (2015). https://doi.org/10.1145/2744769.2747946, https://doi.org/10.1145/2744769.2747946
3. Beierle, C., Jean, J., Kölbl, S., Leander, G., Moradi, A., Peyrin, T., Sasaki, Y., Sasdrich, P., Sim, S.M.: The SKINNY family of block ciphers and its low-latency variant MANTIS. In: Robshaw, M., Katz, J. (eds.) Advances in Cryptology - CRYPTO 2016 - 36th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 14-18, 2016, Proceedings, Part II. Lecture Notes in Computer Science, vol. 9815, pp. 123–153. Springer (2016). https://doi.org/10.1007/978-3-662-53008-5_5, https://doi.org/10.1007/978-3-662-53008-5_5
4. Beyne, T.: A geometric approach to linear cryptanalysis. In: Tibouchi, M., Wang, H. (eds.) Advances in Cryptology - ASIACRYPT 2021 - 27th International Conference on the Theory and Application of Cryptology and Information Security, Singapore, December 6-10, 2021, Proceedings, Part I. Lecture Notes in Computer Science, vol. 13090, pp. 36–66. Springer (2021). https://doi.org/10.1007/978-3-030-92062-3_2, https://doi.org/10.1007/978-3-030-92062-3_2
5. Beyne, T.: A geometric approach to symmetric-key cryptanalysis (2023)
6. Beyne, T., Rijmen, V.: Differential cryptanalysis in the fixed-key model. In: Dodis, Y., Shrimpton, T. (eds.) Advances in Cryptology - CRYPTO 2022 - 42nd Annual International Cryptology Conference, CRYPTO 2022, Santa Barbara, CA, USA, August 15-18, 2022, Proceedings, Part III. Lecture Notes in Computer Science, vol. 13509, pp. 687–716. Springer (2022). https://doi.org/10.1007/978-3-031-15982-4_23, https://doi.org/10.1007/978-3-031-15982-4_23
7. Beyne, T., Verbauwhede, M.: Integral cryptanalysis using algebraic transition matrices. IACR Trans. Symmetric Cryptol. **2023**(4), 244–269 (2023). https://doi.org/10.46586/TOSC.V2023.I4.244-269, https://doi.org/10.46586/tosc.v2023.i4.244-269
8. Beyne, T., Verbauwhede, M.: Ultrametric integral cryptanalysis. IACR Cryptol. ePrint Arch. p. 722 (2024), https://eprint.iacr.org/2024/722
9. Biham, E., Dunkelman, O., Keller, N.: Enhancing differential-linear cryptanalysis. In: Zheng, Y. (ed.) Advances in Cryptology - ASIACRYPT 2002, 8th International

Conference on the Theory and Application of Cryptology and Information Security, Queenstown, New Zealand, December 1-5, 2002, Proceedings. Lecture Notes in Computer Science, vol. 2501, pp. 254–266. Springer (2002). https://doi.org/10.1007/3-540-36178-2_16, https://doi.org/10.1007/3-540-36178-2_16

10. Biham, E., Shamir, A.: Differential Cryptanalysis of the Full 16-Round DES. In: Brickell, E.F. (ed.) Advances in Cryptology - CRYPTO '92. LNCS, vol. 740, pp. 487–496. Springer (1992). https://doi.org/10.1007/3-540-48071-4_34, https://doi.org/10.1007/3-540-48071-4_34
11. Blondeau, C., Leander, G., Nyberg, K.: Differential-linear cryptanalysis revisited. J. Cryptol. **30**(3), 859–888 (2017). https://doi.org/10.1007/S00145-016-9237-5, https://doi.org/10.1007/s00145-016-9237-5
12. Bogdanov, A., Knudsen, L.R., Leander, G., Paar, C., Poschmann, A., Robshaw, M.J.B., Seurin, Y., Vikkelsoe, C.: PRESENT: an ultra-lightweight block cipher. In: Paillier, P., Verbauwhede, I. (eds.) Cryptographic Hardware and Embedded Systems - CHES 2007, 9th International Workshop, Vienna, Austria, September 10-13, 2007, Proceedings. Lecture Notes in Computer Science, vol. 4727, pp. 450–466. Springer (2007). https://doi.org/10.1007/978-3-540-74735-2_31, https://doi.org/10.1007/978-3-540-74735-2_31
13. Boura, C., Canteaut, A.: Another view of the division property. In: Robshaw, M., Katz, J. (eds.) Advances in Cryptology - CRYPTO 2016 - 36th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 14-18, 2016, Proceedings, Part I. Lecture Notes in Computer Science, vol. 9814, pp. 654–682. Springer (2016). https://doi.org/10.1007/978-3-662-53018-4_24, https://doi.org/10.1007/978-3-662-53018-4_24
14. Boura, C., Canteaut, A., Coggia, D.: A general proof framework for recent AES distinguishers. IACR Trans. Symmetric Cryptol. **2019**(1), 170–191 (2019). https://doi.org/10.13154/TOSC.V2019.I1.170-191, https://doi.org/10.13154/tosc.v2019.i1.170-191
15. Daemen, J., Rijmen, V.: AES and the Wide Trail Design Strategy. In: Knudsen, L.R. (ed.) Advances in Cryptology - EUROCRYPT 2002. LNCS, vol. 2332, pp. 108–109. Springer (2002). https://doi.org/10.1007/3-540-46035-7_7, https://doi.org/10.1007/3-540-46035-7_7
16. Derbez, P., Fouque, P.: Increasing precision of division property. IACR Trans. Symmetric Cryptol. **2020**(4), 173–194 (2020). https://doi.org/10.46586/TOSC.V2020.I4.173-194, https://doi.org/10.46586/tosc.v2020.i4.173-194
17. Grassi, L., Rechberger, C., Rønjom, S.: A new structural-differential property of 5-round AES. In: Coron, J., Nielsen, J.B. (eds.) Advances in Cryptology - EUROCRYPT 2017 - 36th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Paris, France, April 30 - May 4, 2017, Proceedings, Part II. Lecture Notes in Computer Science, vol. 10211, pp. 289–317 (2017). https://doi.org/10.1007/978-3-319-56614-6_10, https://doi.org/10.1007/978-3-319-56614-6_10
18. Hadipour, H., Derbez, P., Eichlseder, M.: Revisiting differential-linear attacks via a boomerang perspective with application to aes, ascon, clefia, skinny, present, knot, twine, warp, lblock, simeck, and SERPENT. In: Reyzin, L., Stebila, D. (eds.) Advances in Cryptology - CRYPTO 2024 - 44th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2024, Proceedings, Part IV. Lecture Notes in Computer Science, vol. 14923, pp. 38–72. Springer (2024). https://doi.org/10.1007/978-3-031-68385-5_2, https://doi.org/10.1007/978-3-031-68385-5_2

19. Hu, K., Sun, S., Wang, M., Wang, Q.: An Algebraic Formulation of the Division Property: Revisiting Degree Evaluations, Cube attacks, and key-independent sums. In: Moriai, S., Wang, H. (eds.) Advances in Cryptology - ASIACRYPT 2020. LNCS, vol. 12491, pp. 446–476. Springer (2020). https://doi.org/10.1007/978-3-030-64837-4_15, https://doi.org/10.1007/978-3-030-64837-4_15
20. Jean, J.: TikZ for Cryptographers. https://www.iacr.org/authors/tikz/ (2016)
21. Jean, J., Nikolic, I., Peyrin, T.: Tweaks and keys for block ciphers: The TWEAKEY framework. In: Sarkar, P., Iwata, T. (eds.) Advances in Cryptology - ASIACRYPT 2014 - 20th International Conference on the Theory and Application of Cryptology and Information Security, Kaoshiung, Taiwan, R.O.C., December 7-11, 2014, Proceedings, Part II. Lecture Notes in Computer Science, vol. 8874, pp. 274–288. Springer (2014). https://doi.org/10.1007/978-3-662-45608-8_15, https://doi.org/10.1007/978-3-662-45608-8_15
22. Knudsen, L.R., Wagner, D.A.: Integral Cryptanalysis. In: Daemen, J., Rijmen, V. (eds.) Fast Software Encryption - FSE 2002. LNCS, vol. 2365, pp. 112–127. Springer (2002). https://doi.org/10.1007/3-540-45661-9_9, https://doi.org/10.1007/3-540-45661-9_9
23. Langford, S., Hellman, M.: Differential-Linear Cryptanalysis. In: CRYPTO 1994
24. Leander, G., Tezcan, C., Wiemer, F.: Searching for subspace trails and truncated differentials. IACR Trans. Symmetric Cryptol. **2018**(1), 74–100 (2018). https://doi.org/10.13154/TOSC.V2018.I1.74-100, https://doi.org/10.13154/tosc.v2018.i1.74-100
25. Liu, M., Lu, X., Lin, D.: Differential-linear cryptanalysis from an algebraic perspective. In: Malkin, T., Peikert, C. (eds.) Advances in Cryptology - CRYPTO 2021 - 41st Annual International Cryptology Conference, CRYPTO 2021, Virtual Event, August 16-20, 2021, Proceedings, Part III. Lecture Notes in Computer Science, vol. 12827, pp. 247–277. Springer (2021). https://doi.org/10.1007/978-3-030-84252-9_9, https://doi.org/10.1007/978-3-030-84252-9_9
26. Matsui, M.: Linear Cryptanalysis Method for DES Cipher. In: EUROCRYPT 1993
27. Mouha, N., Wang, Q., Gu, D., Preneel, B.: Differential and linear cryptanalysis using mixed-integer linear programming. In: Wu, C., Yung, M., Lin, D. (eds.) Information Security and Cryptology - 7th International Conference, Inscrypt 2011, Beijing, China, November 30 - December 3, 2011. Revised Selected Papers. Lecture Notes in Computer Science, vol. 7537, pp. 57–76. Springer (2011). https://doi.org/10.1007/978-3-642-34704-7_5, https://doi.org/10.1007/978-3-642-34704-7_5
28. Peng, T., Zhang, W., Weng, J., Ding, T.: New approaches for estimating the bias of differential-linear distinguishers. In: Reyzin, L., Stebila, D. (eds.) Advances in Cryptology - CRYPTO 2024 - 44th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2024, Proceedings, Part IV. Lecture Notes in Computer Science, vol. 14923, pp. 174–205. Springer (2024). https://doi.org/10.1007/978-3-031-68385-5_6, https://doi.org/10.1007/978-3-031-68385-5_6
29. Serre, J.: Linear Representations of Finite Groups, Graduate texts in mathematics, vol. 42. Springer (1977)
30. Sun, S., Hu, L., Wang, P., Qiao, K., Ma, X., Song, L.: Automatic security evaluation and (related-key) differential characteristic search: Application to simon, present, lblock, DES(L) and other bit-oriented block ciphers. In: Sarkar, P., Iwata, T. (eds.) Advances in Cryptology - ASIACRYPT 2014 - 20th International Conference on the Theory and Application of Cryptology and Information Security, Kaoshiung,

Taiwan, R.O.C., December 7-11, 2014. Proceedings, Part I. Lecture Notes in Computer Science, vol. 8873, pp. 158–178. Springer (2014). https://doi.org/10.1007/978-3-662-45611-8_9, https://doi.org/10.1007/978-3-662-45611-8_9

31. Tiessen, T.: Polytopic Cryptanalysis. In: Fischlin, M., Coron, J. (eds.) Advances in Cryptology - EUROCRYPT 2016. LNCS, vol. 9665, pp. 214–239. Springer (2016). https://doi.org/10.1007/978-3-662-49890-3_9, https://doi.org/10.1007/978-3-662-49890-3_9
32. Todo, Y.: Structural evaluation by generalized integral property. In: Oswald, E., Fischlin, M. (eds.) Advances in Cryptology - EUROCRYPT 2015 - 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Sofia, Bulgaria, April 26-30, 2015, Proceedings, Part I. Lecture Notes in Computer Science, vol. 9056, pp. 287–314. Springer (2015). https://doi.org/10.1007/978-3-662-46800-5_12, https://doi.org/10.1007/978-3-662-46800-5_12
33. Todo, Y., Morii, M.: Bit-based division property and application to simon family. In: Peyrin, T. (ed.) Fast Software Encryption - 23rd International Conference, FSE 2016, Bochum, Germany, March 20-23, 2016, Revised Selected Papers. Lecture Notes in Computer Science, vol. 9783, pp. 357–377. Springer (2016). https://doi.org/10.1007/978-3-662-52993-5_18, https://doi.org/10.1007/978-3-662-52993-5_18
34. Wagner, D.A.: The boomerang attack. In: Knudsen, L.R. (ed.) Fast Software Encryption, 6th International Workshop, FSE '99, Rome, Italy, March 24-26, 1999, Proceedings. Lecture Notes in Computer Science, vol. 1636, pp. 156–170. Springer (1999). https://doi.org/10.1007/3-540-48519-8_12, https://doi.org/10.1007/3-540-48519-8_12
35. Wang, L., Song, L., Wu, B., Rahman, M., Isobe, T.: Revisiting the boomerang attack from a perspective of 3-differential. IEEE Trans. Inf. Theory **70**(7), 5343–5357 (2024). https://doi.org/10.1109/TIT.2023.3324738, https://doi.org/10.1109/TIT.2023.3324738
36. Yang, G., Zhu, B., Suder, V., Aagaard, M.D., Gong, G.: The simeck family of lightweight block ciphers. In: Güneysu, T., Handschuh, H. (eds.) Cryptographic Hardware and Embedded Systems - CHES 2015 - 17th International Workshop, Saint-Malo, France, September 13-16, 2015, Proceedings. Lecture Notes in Computer Science, vol. 9293, pp. 307–329. Springer (2015). https://doi.org/10.1007/978-3-662-48324-4_16, https://doi.org/10.1007/978-3-662-48324-4_16

A Specifications of SKINNY-64, PRESENT and SIMECK

A.1 Specifications of PRESENT

PRESENT is a 64-bit block cipher supporting 80-bit and 128-bit keys designed by Bogdanov *et al.* in 2007 [12]. The design is a SPN construction consisting of a round key addition, a 4-bit Sbox layer, and a bit permutation layer. The S-box is specified as follows:

x	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
S[x]	c	5	6	b	9	0	a	d	3	e	f	8	4	7	1	2

The bit permutation and the entire round function are both illustrated in Figure 4.

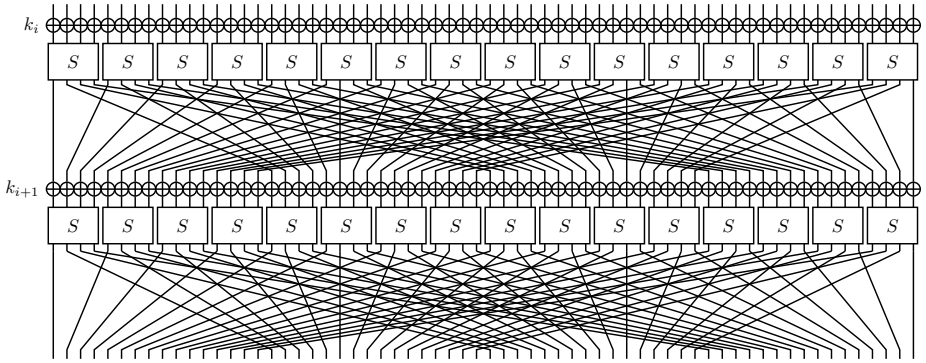

Fig. 4: Round function of PRESENT. The figure is taken from [20].

A.2 Specifications of SKINNY-64

The block cipher family SKINNY was presented at CRYPTO 2016 [3] designed under the TWEAKEY framework [21], whose goal is to compete with the NSA design SIMON [2] in terms of hardware/software performance. According to the length of block and tweakey, the SKINNY family consists of 6 different members represented as SKINNY-n-t, where $n \in \{64, 128\}$ and $t \in \{n, 2n, 3n\}$, which respectively represent the sizes of block and tweakey. Here we introduce the 64-bit version of SKINNY, *i.e.*, SKINNY-64, under the single tweakey model. SKINNY-64 is chosen as its Sbox is 4-bit. Since the multiple-of-n property is described as a 2nd order attack, it is equivalent to describe the propagation for an 8-bit Sbox ciphers in the classical automatic search.

The round function of SKINNY-64 comprises five operations as SubCells (SC), AddConstants (AC), AddRoundTweakey (ART), ShiftRows (SR) and MixColumns (MC), see Figure 5. So a round of SKINNY-64 can be written as

$$R = \mathsf{MC} \circ \mathsf{SR} \circ \mathsf{ART} \circ \mathsf{AC} \circ \mathsf{SC}.$$

1. SC: SC is the only non-linear layer of SKINNY-64, using a 4-bit Sbox S as follows,

x	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
S[x]	c	6	9	0	1	a	2	b	3	8	5	d	4	e	7	f

2. AC and ART: In the AC operation, a 6-bit round-based constant is XORed with the top two cells of the first column, and a constant 2 is XORed with the third cell of the first column. In the ART operation, a 8-cell round key is XORed with the first two rows of the state.
3. SR: SR circularly shifts the i-th row of the internal state to right with i nibbles, where $i = 0, 1, 2, 3$.
4. MC: MC multiplies four nibbles of each state column with the binary matrix M. The details of M are listed below,

$$M = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

Since the non-zero elements in this matrix are only 1, the MC operation can be decomposed into 4 parallel small operations called Lbox, denoted by LBox. Let the input and output of M is x and y, $(y_i, y_{i+4}, y_{i+8}, y_{i+12} = \mathsf{L}(x_i, x_{i+4}, x_{i+8}, x_{i+12}) = (x_0 \oplus x_2 \oplus x_3, x_0, x_1 \oplus x_2, x_0 \oplus x_2)$, for $i \in \{0, 1, 2, 3\}$.

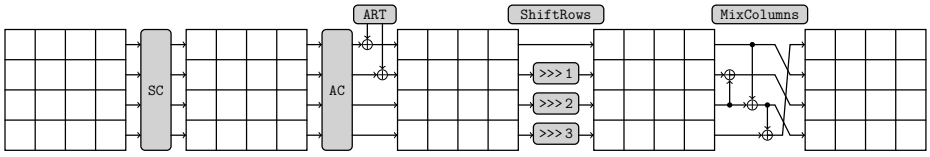

Fig. 5: Round function of SKINNY-64. The figure is taken from [20].

A.3 Specifications of SIMECK

SIMECK is a family of lightweight block ciphers proposed at CHES 2015 [36]. The design is similar to SIMON. The SIMECK family consists of several family members SIMECK-$2n/4n$ operating on n-bit words with a state size of $2n$ bits and a key size of $4n$ bits for $n \in \{16, 24, 32\}$. In round i, the $2n$-bit input state of round i is split into two n-bit words (L_i, R_i) and updated with a Feistel-based round function F to produce (L_{i+1}, R_{i+1}) using the n-bit round key K_i. The round function is a quadratic Feistel function using bitwise XOR $(x \oplus y)$, bitwise AND $x \wedge y$, and cyclic left-shifts by c bits $(x \lll c)$ (see Figure 6):

$$R_{i+1} = L_i$$
$$L_{i+1} = R_i \oplus K_i \oplus (L_i \wedge (L_i \lll 5)) \oplus (L_i \lll 1).$$

The round key K_i is produced using a similar nonlinear update function. The total number of rounds is 32 rounds for SIMECK 32/64 (referred to as SIMECK-32 for short), 36 rounds for SIMECK 48/96 (referred to as SIMECK-48).

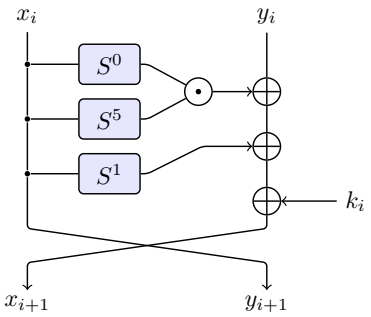

Fig. 6: Round function of SIMECK. The figure is adapted from [20].

B High-Level Viewpoint of Automatic Search for Geometric Approach

In the past decade, automatic search methods have been very popular in cryptanalysis and many classical attack techniques have been modelled. The idea is to express a cryptanalytic problem into a constrained problem, such as Mixed Integer Linear Programming (MILP) or the Satisfiability Problem (SAT), then use off-the-shelf solvers to complete the search. The results are then translated into solutions for the original cryptanalytic problem.

In the case of the geometric approach, the transition matrix is naturally suitable to be modelled in such frameworks.

First, the cipher is divided into many small components, such as Sboxes, bit permutations, and even XORs, ANDs, or COPYs (aka. Branches, where a bit is copied into 2 or multiple bits). Then, each component can be viewed as a "function" (the COPY function is also viewed as a function with one input and two outputs), the statistic derived after choosing two bases for the input and output is then applied to the function. For a function $\mathcal{F} : \mathbb{F}_2^n \longrightarrow \mathbb{F}_2^m$ (note that n and m are usually small as they are the components of the target cipher), we traverse all input and output values. For the d-th order attack, the input and output vectors should be $u_0||u_1||\dots||u_{d-1}$ and $v_0||v_1||\dots||v_{d-1}$, respectively. Then, according to the statistic, a value related to the input/output vectors, denoted by c, is obtained. Values $c \neq 0$ are then made into an entry:

$$(u_0||u_1||\cdots||u_{d-1}, v_0||v_1||\cdots||v_{d-1}, M(c)),$$

where $M(c)$ represents the values after applying some measures to c (usually forcing it to a positive integer number). For example, if one targets a probability, then

$M(c)$ is usually $-\log(c)$. While for the divisibility property, $M(c) = -\log(|c|_2)$ where $|c|_2$ is the 2-adic absolute value of c. Usually, such an entry will be edited as bit strings.

All these entries with $c \neq 0$ will be called *valid propagations*. We generate corresponding variables for the input and output of $\mathcal{F}$, then we can use a set of inequalities, CNF constraints or other methods to make sure that these variables have to be one of these valid propagations.

Finally, we define an objective function that usually sums up all variables from $M(c)$, and we use a solver to search for one trail that makes the summation maximum or minimum. Sometimes, one can also want to search for all valid trails.

We recommend that readers refer to previous research on geometric approach, such as [6] and [8], for a deeper understanding on how automation is applied in this field.

C Step-By-Step Explanation of Equation (4)

$$
\begin{aligned}
&A_{v,u}^{\mathcal{E}^{\otimes d}} = [\beta_y^\star(v), 0 \le y < 2^{dn}]^\top \; T^{\mathcal{E}^{\otimes d}} \; [\alpha_u(x), 0 \le x < 2^{dn}] \\
&= \left([\beta_y^\star(v), 0 \le y < 2^{dn}]^\top \; T^{\mathcal{E}^{\otimes d}}\right) \; [\alpha_u(x), 0 \le x < 2^{dn}] \\
&\text{Note that the } i\text{-th column of } T^{\mathcal{E}^{\otimes d}} \text{ is } [\delta_y(\mathcal{E}^{\otimes d}(i)), y = 0, 1, \ldots, 2^{dn}-1] \\
&= \left[\sum_{y \in (\mathbb{F}_2^n)^{\otimes d}} \beta_y^\star(v)\delta_y(\mathcal{E}(0)), \ldots, \sum_{y \in (\mathbb{F}_2^n)^{\otimes d}} \beta_y^\star(v)\delta_y(\mathcal{E}(2^{dn}-1))\right]^\top \; [\alpha_u(x), 0 \le x < 2^{dn}] \\
&\text{Note that } \textstyle\sum_{y \in (\mathbb{F}_2^n)^{\otimes d}} \beta_y^\star(v)\delta_y(\mathcal{E}(i)) = \beta^\star_{\mathcal{E}^{\otimes d}(i)}(v) \text{ because only when } y = \mathcal{E}^{\otimes d}(i),\ \delta_y(\mathcal{E}^{\otimes d}(i)) = 1 \\
&= \left[\beta^\star_{\mathcal{E}^{\otimes d}(0)}(v), \ldots, \beta^\star_{\mathcal{E}^{\otimes d}(2^{dn}-1)}(v)\right]^\top \; [\alpha_u(x), 0 \le x < 2^{dn}] \\
&= \sum_{x \in (\mathbb{F}_2^n)^{\otimes d}} \beta^\star_{\mathcal{E}^{\otimes d}(x)}(v) \; \alpha_u(x)
\end{aligned}
$$

D Automatic Search for the Simplified Ultrametric Integral Crytanalysis

We replay here the ultrametric integral attack, but in the simplified way described in Section 4. Setting $u =$ `fffffffffffffffe`, we obtain the same results for 9 rounds of `PRESENT` as for [8]. We divide the 9-round `PRESENT` without the last bit permutation into 3 parts: $\mathcal{E}_0$ covers the first 4 rounds, $\mathcal{E}_1$ covers the Sbox layer of the 5th round, and $\mathcal{E}_2$ covers the remaining 4 rounds. The transition matrices of the Sboxes of $\mathcal{E}_0$, $\mathcal{E}_1$ and $\mathcal{E}_2$ can be computed according to Equations (7), (9) and (8). The transition matrix of $\mathcal{E}_2$ is the same as the one in [8], but we still provide it here for a better comparison among the three matrices of the `PRESENT` Sbox.

Now let us consider the propagation. For the Sbox layer, we first construct the transition matrices for a single Sbox, which is not difficult since the statistic

expressions of the three matrices we constructed earlier already exist. Based on the transition matrix of a single Sbox, the propagation rules of the Sbox layer can be conveniently characterized: PRESENT's Sbox layer consists of 16 Sboxes, we have $A_{v,u}^{S_0||\cdots||S_{15}} = \prod_{i=0}^{15} A_{v_i,u_i}^{S_i}$, where $A^{S_0||\cdots||S_{15}}$ is the transition matrix of the Sbox layer and A^{S_i} is the transition matrix of the i-th Sbox.

For the bit permutation layer P, one can easily obtain that for all the three matrices, $M_{v,u} \neq 0$ if and only if $v = P(u)$. For the round key layer $K(x) = x \oplus k$, we can regard it as 64 parallel 2-input-1-output functions, so we have

$$A^K = \bigotimes_{i=0}^{63} A^{K_i} = \bigotimes_{i=0}^{63} \begin{bmatrix} (-1)^{k_i} & 0 \\ k_i & 1 \end{bmatrix} \text{(for } \mathcal{E}_0\text{)}, A^K = \bigotimes_{i=0}^{63} A^{K_i} = \bigotimes_{i=0}^{63} \begin{bmatrix} 1 & 0 \\ k_i & (-1)^{k_i} \end{bmatrix} \text{(for } \mathcal{E}_2\text{)},$$

Our goal is to obtain the 2-adic value of $A_{v,u}^{\mathcal{E}} = \sum_{y \preceq u} \mathcal{E}^v(y)$. Because of the triangle inequality of 2-adic value, we have

$$\left|A_{u_r,u_0}^{\mathcal{E}}\right|_2 \leq \max_{u_{r-1},u_{r-2},\ldots,u_2} \left|\prod_{i=0}^{r-1} A_{u_{i+1},u_i}^{\mathcal{E}^i}\right|_2.$$

Therefore, we only need to utilize an automated search tool to find the path that maximizes the 2-adic value of $\prod_{i=0}^{r-1} A_{u_{i+1},u_i}^{\mathcal{E}^i}$ according to the propagation rules mentioned above. Table 5 presents the search results for 9-round PRESENT.

Table 5: Divisibility property for the integral distinguisher on 9-round PRESENT searched by our simplified method, with input set $\{x : x \preceq \texttt{fffffffffffffffe}\}$. The number of times the i-th output bit equals one is divisible by 2^{b_i}. These results are completely the same as [8].

bit i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
theoretical b_i	2	1	1	1	2	0	0	0	2	0	0	0	2	0	0	0
	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
	1	1	1	1	1	0	0	0	1	0	0	0	1	0	0	0
	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
	1	1	1	1	1	0	0	0	1	0	0	0	1	0	0	0
	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
	1	1	1	1	1	0	0	0	1	0	0	0	1	0	0	0

E Obtain Blondeau-Leander-Nyberg Formula from Geometric Approach

Given Equation (15), and taking $\mathcal{E}_1 \circ \mathcal{E}_0$ as a whole part, we get

$$\varepsilon[\delta \xrightarrow{\mathcal{E}} \lambda] = A^{\mathcal{E}}_{0||\lambda,0||\delta} = \sum_{\gamma_0||\gamma_1} A^{\mathcal{E}_2}_{0||\lambda,\gamma_0||\gamma_1} A^{\mathcal{E}_1\circ\mathcal{E}_0}_{\gamma_0||\gamma_1,0||\delta}$$

we can set $\gamma = 0$ to force $\mathcal{E}_1 \circ \mathcal{E}_0$ and $\mathcal{E}_2$ to be independent. Indeed, $A^{\mathcal{E}_1\circ\mathcal{E}_0}_{\gamma_0=0||\gamma_1,0||\delta}$ represents the correlation of a DL approximation of $\mathcal{E}_1 \circ \mathcal{E}_0$ with the input difference δ and the output mask γ_1. The independence of $\mathcal{E}_1 \circ \mathcal{E}_0$ and $\mathcal{E}_2$ is equivalent to say that the intermediate values at the connection point can be any values, which is equivalent to $\gamma_0 = 0$.

Note that

$$\begin{aligned}
&A^{\mathcal{E}_2}_{0||\lambda,0||\gamma_1} = 2^{-2n} \sum_{x_0\in\mathbb{F}_2^n, x_1\in\mathbb{F}_2^n} (-1)^{\gamma_1^\top x_1 \oplus \lambda^\top \mathcal{E}(x_1)} \\
&= 2^{-2n} \sum_{x_0\in\mathbb{F}_2^n, x_0\oplus x_1\in\mathbb{F}_2^n} (-1)^{\gamma_1^\top x_0 \oplus \gamma_1^\top (x_0\oplus x_1) \oplus \lambda^\top \mathcal{E}(x_0) \oplus \lambda^\top \mathcal{E}(x_0) \oplus \lambda^\top \mathcal{E}(x_0\oplus x_1)} \\
&= \left(2^{-n} \sum_{x_0\in\mathbb{F}_2^n} (-1)^{\gamma_1^\top x_0 \oplus \lambda^\top \mathcal{E}(x_0)}\right) \left(2^{-n} \sum_{x_0\oplus x_1\in\mathbb{F}_2^n} (-1)^{\gamma_1^\top (x_0\oplus x_1) \oplus \lambda^\top (\mathcal{E}(x_0)\oplus\mathcal{E}(x_1))}\right) \\
&= c^2[\gamma_1 \xrightarrow{\mathcal{E}_2} \lambda]
\end{aligned}$$

Thus, Equation (15) becomes to

$$\begin{aligned}
\varepsilon[\delta \xrightarrow{\mathcal{E}} \lambda] = A^{\mathcal{E}}_{0||\lambda,0||\delta} &= \sum_{0||\gamma_1} A^{\mathcal{E}_2}_{0||\lambda,0||\gamma_1} A^{\mathcal{E}_1\circ\mathcal{E}_0}_{0||\gamma_1,0||\delta} = \sum_{0||\gamma_1} A^{\mathcal{E}_2}_{0||\lambda,0||\gamma_1} A^{\mathcal{E}_1\circ\mathcal{E}_0}_{0||\gamma_1,0||\delta} \\
&= \sum_{\gamma_1} \varepsilon[\delta \xrightarrow{\mathcal{E}_1\circ\mathcal{E}_0} \gamma_1] A^{\mathcal{E}_2}_{0||\lambda,0||\gamma_1} = \sum_{\gamma_1} \varepsilon[\delta \xrightarrow{\mathcal{E}_1\circ\mathcal{E}_0} \gamma_1] c^2[\gamma_1 \xrightarrow{\mathcal{E}_2} \lambda]
\end{aligned}$$

which is exactly Equation (13).

Table 6: First order attacks (first part)

Output/Input	$[\delta_u(x)]_{x,u}$	$[(-1)^{u^\top x}]_{x,u}$	$[2^{-n}(-1)^{u^\top x}]_{x,u}$	$[u^x]_{x,u}$
$[\delta_u(x)]_{x,u}$	$\sum_{x=u,\mathcal{E}(x)=v} 1$	$\sum_{\substack{x\in\mathbb{F}_2^n\\ \mathcal{E}(x)=v}} (-1)^{u^\top x}$	$2^{-n}\sum_{\substack{x\in\mathbb{F}_2^n\\ \mathcal{E}(x)=v}} (-1)^{u^\top x}$	$\sum_{\substack{x\preceq u\\ \mathcal{E}(x)=v}} 1$
$[(-1)^{u^\top x}]_{x,u}$	$2^{-n}\sum_{x=u}(-1)^{v^\top \mathcal{E}(x)}$	$2^{-n}\sum_{x\in\mathbb{F}_2^n}(-1)^{u^\top x\oplus v^\top \mathcal{E}(x)}$	$2^{-2n}\sum_{x\in\mathbb{F}_2^n}(-1)^{u^\top x\oplus v^\top \mathcal{E}(x)}$	$2^{-n}\sum_{x\preceq u}(-1)^{v^\top \mathcal{E}(x)}$
$[2^{-n}(-1)^{u^\top x}]_{x,u}$	$\sum_{x=u}(-1)^{v^\top \mathcal{E}(x)}$	$\sum_{x\in\mathbb{F}_2^n}(-1)^{u^\top x\oplus v^\top \mathcal{E}(x)}$	$2^{-n}\sum_{x\in\mathbb{F}_2^n}(-1)^{u^\top x\oplus v^\top \mathcal{E}(x)}$	$\sum_{x\preceq u}(-1)^{v^\top \mathcal{E}(x)}$
$[u^x]_{x,u}$	$\sum_{x=u}(-1)^{\mathsf{wt}(v\oplus\mathcal{E}(x))}\mathcal{E}^v(x)$	$\sum_{x\in\mathbb{F}_2^n}(-1)^{u^\top x}(-1)^{\mathsf{wt}(v\oplus\mathcal{E}(x))}\mathcal{E}^v(x)$	$2^{-n}\sum_{x\in\mathbb{F}_2^n}(-1)^{u^\top x}(-1)^{\mathsf{wt}(v\oplus\mathcal{E}(x))}\mathcal{E}^v(x)$	$\sum_{x\preceq u}(-1)^{\mathsf{wt}(v\oplus\mathcal{E}(x))}\mathcal{E}^v(x)$
$[(-1)^{\mathsf{wt}(u\oplus x)}u^x]_{x,u}$	$\sum_{x=u}\mathcal{E}^v(x)$	$\sum_{x\in\mathbb{F}_2^n}(-1)^{u^\top x}\mathcal{E}^v(x)$	$2^{-n}\sum_{x\in\mathbb{F}_2^n}\mathcal{E}^v(x)$	$\sum_{x\preceq u}\mathcal{E}^v(x)$
$[x^u]_{x,u}$	$\sum_{x=u}(-1)^{\mathsf{wt}(v\oplus\mathcal{E}(x))}v^{\mathcal{E}(x)}$	$\sum_{x\in\mathbb{F}_2^n}(-1)^{u^\top x}(-1)^{\mathsf{wt}(v\oplus\mathcal{E}(x))}v^{\mathcal{E}(x)}$	$2^{-n}\sum_{x\in\mathbb{F}_2^n}(-1)^{u^\top x}(-1)^{\mathsf{wt}(v\oplus\mathcal{E}(x))}v^{\mathcal{E}(x)}$	$\sum_{x\preceq u}(-1)^{\mathsf{wt}(v\oplus\mathcal{E}(x))}v^{\mathcal{E}(x)}$
$[(-1)^{\mathsf{wt}(u\oplus x)}x^u]_{x,u}$	$\sum_{x=u}v^{\mathcal{E}(x)}$	$\sum_{x\in\mathbb{F}_2^n}(-1)^{u^\top x}v^{\mathcal{E}(x)}$	$2^{-n}\sum_{x\in\mathbb{F}_2^n}(-1)^{u^\top x}v^{\mathcal{E}(x)}$	$\sum_{x\preceq u}v^{\mathcal{E}(x)}$

Table 7: First order attacks (second part)

Output/Input	$[(-1)^{\mathsf{wt}(u\oplus x)}u^x]_{x,u}$	$[x^u]_{x,u}$	$[(-1)^{\mathsf{wt}(u\oplus x)}x^u]_{x,u}$
$[\delta_u(x)]_{x,u}$	$\sum_{\substack{x\preceq u\\ \mathcal{E}(x)=v}}(-1)^{\mathsf{wt}(u\oplus x)}$	$\sum_{\substack{x\succeq u\\ \mathcal{E}(x)=v}}1$	$\sum_{\substack{x\succeq u\\ \mathcal{E}(x)=v}}(-1)^{\mathsf{wt}(u\oplus x)}$
$[(-1)^{u^\top x}]_{x,u}$	$2^{-n}\sum_{x\preceq u}(-1)^{\mathsf{wt}(u\oplus x)}(-1)^{v^\top\mathcal{E}(x)}$	$2^{-n}\sum_{x\preceq u}(-1)^{v^\top\mathcal{E}(x)}$	$2^{-n}\sum_{x\succeq u}(-1)^{\mathsf{wt}(u\oplus x)}(-1)^{v^\top\mathcal{E}(x)}$
$[2^{-n}(-1)^{u^\top x}]_{x,u}$	$\sum_{x\preceq u}(-1)^{\mathsf{wt}(u\oplus x)}(-1)^{v^\top\mathcal{E}(x)}$	$\sum_{x\succeq u}(-1)^{v^\top\mathcal{E}(x)}$	$\sum_{x\succeq u}(-1)^{\mathsf{wt}(u\oplus x)}(-1)^{v^\top\mathcal{E}(x)}$
$[u^x]_{x,u}$	$\sum_{x\preceq u}(-1)^{\mathsf{wt}(u\oplus x)}(-1)^{\mathsf{wt}(v\oplus\mathcal{E}(x))}\mathcal{E}^v(x)$	$\sum_{x\succeq u}(-1)^{\mathsf{wt}(v\oplus\mathcal{E}(x))}\mathcal{E}^v(x)$	$\sum_{x\succeq u}(-1)^{\mathsf{wt}(u\oplus x)}(-1)^{\mathsf{wt}(v\oplus\mathcal{E}(x))}\mathcal{E}^v(x)$
$[(-1)^{\mathsf{wt}(u\oplus x)}u^x]_{x,u}$	$\sum_{x\preceq u}(-1)^{\mathsf{wt}(u\oplus x)}\mathcal{E}^v(x)$	$\sum_{x\succeq u}\mathcal{E}^v(x)$	$\sum_{x\succeq u}(-1)^{\mathsf{wt}(u\oplus x)}\mathcal{E}^v(x)$
$[x^u]_{x,u}$	$\sum_{x\preceq u}(-1)^{\mathsf{wt}(u\oplus x)}(-1)^{\mathsf{wt}(v\oplus\mathcal{E}(x))}v^{\mathcal{E}(x)}$	$\sum_{x\succeq u}(-1)^{\mathsf{wt}(v\oplus\mathcal{E}(x))}v^{\mathcal{E}(x)}$	$\sum_{x\succeq u}(-1)^{\mathsf{wt}(u\oplus x)}(-1)^{\mathsf{wt}(v\oplus\mathcal{E}(x))}v^{\mathcal{E}(x)}$
$[(-1)^{\mathsf{wt}(u\oplus x)}x^u]_{x,u}$	$\sum_{x\preceq u}(-1)^{\mathsf{wt}(u\oplus x)}v^{\mathcal{E}(x)}$	$\sum_{x\succeq u}v^{\mathcal{E}(x)}$	$\sum_{x\succeq u}(-1)^{\mathsf{wt}(u\oplus x)}v^{\mathcal{E}(x)}$

Table 8: 49 Bases for the second order attacks and their effects(first part)

Index	**Basis**	**Effect of input** $\alpha_u(x)$	**Effect of output** $\beta^{\star}_{\mathcal{E}(x)}(v)$
0	$[\delta_{u_0}(x_0)]_{x_0,u_0} \otimes [\delta_{u_1}(x_1)]_{x_1,u_1}$	$\delta_{u_0}(x_0)\delta_{u_1}(x_1)$	$\delta_{v_0}(\mathcal{E}(x_0))\delta_{v_1}(\mathcal{E}(x_1))$
1	$[\delta_{u_0}(x_0)]_{x_0,u_0} \otimes [2^{-n}(-1)^{{u_1}^\top x_1}]_{x_1,u_1}$	$\delta_{u_0}(x_0)2^{-n}(-1)^{u_1^\top x_1}$	$\delta_{v_0}(\mathcal{E}(x_0))(-1)^{v_1^\top \mathcal{E}(x_1)}$
2	$[\delta_{u_0}(x_0)]_{x_0,u_0} \otimes [(-1)^{{u_1}^\top x_1}]_{x_1,u_1}$	$\delta_{u_0}(x_0)(-1)^{u_1^\top x_1}$	$\delta_{v_0}(\mathcal{E}(x_0))2^{-n}(-1)^{v_1^\top \mathcal{E}(x_1)}$
3	$[\delta_{u_0}(x_0)]_{x_0,u_0} \otimes [{u_1}^{x_1}]_{x_1,u_1}$	$\delta_{u_0}(x_0){u_1}^{x_1}$	$\delta_{v_0}(\mathcal{E}(x_0))(-1)^{wt(v_1\oplus\mathcal{E}(x_1))}\mathcal{E}^{v_1}(x_1)$
4	$[\delta_{u_0}(x_0)]_{x_0,u_0} \otimes [(-1)^{wt(u_1\oplus x_1)}{u_1}^{x_1}]_{x_1,u_1}$	$\delta_{u_0}(x_0)(-1)^{wt(u_1\oplus x_1)}{u_1}^{x_1}$	$\delta_{v_0}(\mathcal{E}(x_0))\mathcal{E}^{v_1}(x_1)$
5	$[\delta_{u_0}(x_0)]_{x_0,u_0} \otimes [{x_1}^{u_1}]_{x_1,u_1}$	$\delta_{u_0}(x_0){x_1}^{u_1}$	$\delta_{v_0}(\mathcal{E}(x_0))(-1)^{wt(v_1\oplus\mathcal{E}(x_1))}v_1^{\mathcal{E}(x_1)}$
6	$[\delta_{u_0}(x_0)]_{x_0,u_0} \otimes [(-1)^{wt(u_1\oplus x_1)}{x_1}^{u_1}]_{x_1,u_1}$	$\delta_{u_0}(x_0)(-1)^{wt(u_1\oplus x_1)}{x_1}^{u_1}$	$\delta_{v_0}(\mathcal{E}(x_0))v_1^{\mathcal{E}(x_1)}$
7	$[2^{-n}(-1)^{{u_0}^\top x_0}]_{x_0,u_0} \otimes [\delta_{u_1}(x_1)]_{x_1,u_1}$	$2^{-n}(-1)^{u_0^\top x_0}\delta_{u_1}(x_1)$	$(-1)^{v_0^\top \mathcal{E}(x_0)}\delta_{v_1}(\mathcal{E}(x_1))$
8	$[2^{-n}(-1)^{{u_0}^\top x_0}]_{x_0,u_0} \otimes [2^{-n}(-1)^{{u_1}^\top x_1}]_{x_1,u_1}$	$2^{-n}(-1)^{u_0^\top x_0}2^{-n}(-1)^{u_1^\top x_1}$	$(-1)^{v_0^\top \mathcal{E}(x_0)}(-1)^{v_1^\top \mathcal{E}(x_1)}$
9	$[2^{-n}(-1)^{{u_0}^\top x_0}]_{x_0,u_0} \otimes [(-1)^{{u_1}^\top x_1}]_{x_1,u_1}$	$2^{-n}(-1)^{u_0^\top x_0}(-1)^{u_1^\top x_1}$	$(-1)^{v_0^\top \mathcal{E}(x_0)}2^{-n}(-1)^{v_1^\top \mathcal{E}(x_1)}$
10	$[2^{-n}(-1)^{{u_0}^\top x_0}]_{x_0,u_0} \otimes [{u_1}^{x_1}]_{x_1,u_1}$	$2^{-n}(-1)^{u_0^\top x_0}{u_1}^{x_1}$	$(-1)^{v_0^\top \mathcal{E}(x_0)}(-1)^{wt(v_1\oplus\mathcal{E}(x_1))}\mathcal{E}^{v_1}(x_1)$
11	$[2^{-n}(-1)^{{u_0}^\top x_0}]_{x_0,u_0} \otimes [(-1)^{wt(u_1\oplus x_1)}{u_1}^{x_1}]_{x_1,u_1}$	$2^{-n}(-1)^{u_0^\top x_0}(-1)^{wt(u_1\oplus x_1)}{u_1}^{x_1}$	$(-1)^{v_0^\top \mathcal{E}(x_0)}\mathcal{E}^{v_1}(x_1)$
12	$[2^{-n}(-1)^{{u_0}^\top x_0}]_{x_0,u_0} \otimes [{x_1}^{u_1}]_{x_1,u_1}$	$2^{-n}(-1)^{u_0^\top x_0}{x_1}^{u_1}$	$(-1)^{v_0^\top \mathcal{E}(x_0)}(-1)^{wt(v_1\oplus\mathcal{E}(x_1))}v_1^{\mathcal{E}(x_1)}$
13	$[2^{-n}(-1)^{{u_0}^\top x_0}]_{x_0,u_0} \otimes [(-1)^{wt(u_1\oplus x_1)}{x_1}^{u_1}]_{x_1,u_1}$	$2^{-n}(-1)^{u_0^\top x_0}(-1)^{wt(u_1\oplus x_1)}{x_1}^{u_1}$	$(-1)^{v_0^\top \mathcal{E}(x_0)}v_1^{\mathcal{E}(x_1)}$
14	$[(-1)^{{u_0}^\top x_0}]_{x_0,u_0} \otimes [\delta_{u_1}(x_1)]_{x_1,u_1}$	$(-1)^{u_0^\top x_0}\delta_{u_1}(x_1)$	$2^{-n}(-1)^{v_0^\top \mathcal{E}(x_0)}\delta_{v_1}(\mathcal{E}(x_1))$
15	$[(-1)^{{u_0}^\top x_0}]_{x_0,u_0} \otimes [2^{-n}(-1)^{{u_1}^\top x_1}]_{x_1,u_1}$	$(-1)^{u_0^\top x_0}2^{-n}(-1)^{u_1^\top x_1}$	$2^{-n}(-1)^{v_0^\top \mathcal{E}(x_0)}(-1)^{v_1^\top \mathcal{E}(x_1)}$
16	$[(-1)^{{u_0}^\top x_0}]_{x_0,u_0} \otimes [(-1)^{{u_1}^\top x_1}]_{x_1,u_1}$	$(-1)^{u_0^\top x_0}(-1)^{u_1^\top x_1}$	$2^{-n}(-1)^{v_0^\top \mathcal{E}(x_0)}2^{-n}(-1)^{v_1^\top \mathcal{E}(x_1)}$
17	$[(-1)^{{u_0}^\top x_0}]_{x_0,u_0} \otimes [{u_1}^{x_1}]_{x_1,u_1}$	$(-1)^{u_0^\top x_0}{u_1}^{x_1}$	$2^{-n}(-1)^{v_0^\top \mathcal{E}(x_0)}(-1)^{wt(v_1\oplus\mathcal{E}(x_1))}\mathcal{E}^{v_1}(x_1)$
18	$[(-1)^{{u_0}^\top x_0}]_{x_0,u_0} \otimes [(-1)^{wt(u_1\oplus x_1)}{u_1}^{x_1}]_{x_1,u_1}$	$(-1)^{u_0^\top x_0}(-1)^{wt(u_1\oplus x_1)}{u_1}^{x_1}$	$2^{-n}(-1)^{v_0^\top \mathcal{E}(x_0)}\mathcal{E}^{v_1}(x_1)$
19	$[(-1)^{{u_0}^\top x_0}]_{x_0,u_0} \otimes [{x_1}^{u_1}]_{x_1,u_1}$	$(-1)^{u_0^\top x_0}{x_1}^{u_1}$	$2^{-n}(-1)^{v_0^\top \mathcal{E}(x_0)}(-1)^{wt(v_1\oplus\mathcal{E}(x_1))}v_1^{\mathcal{E}(x_1)}$
20	$[(-1)^{{u_0}^\top x_0}]_{x_0,u_0} \otimes [(-1)^{wt(u_1\oplus x_1)}{x_1}^{u_1}]_{x_1,u_1}$	$(-1)^{u_0^\top x_0}(-1)^{wt(u_1\oplus x_1)}{x_1}^{u_1}$	$2^{-n}(-1)^{v_0^\top \mathcal{E}(x_0)}v_1^{\mathcal{E}(x_1)}$
21	$[{u_0}^{x_0}]_{x_0,u_0} \otimes [\delta_{u_1}(x_1)]_{x_1,u_1}$	${u_0}^{x_0}\delta_{u_1}(x_1)$	$(-1)^{wt(v_0\oplus\mathcal{E}(x_0))}\mathcal{E}^{v_0}(x_0)\delta_{v_1}(\mathcal{E}(x_1))$
22	$[{u_0}^{x_0}]_{x_0,u_0} \otimes [2^{-n}(-1)^{{u_1}^\top x_1}]_{x_1,u_1}$	${u_0}^{x_0}2^{-n}(-1)^{u_1^\top x_1}$	$(-1)^{wt(v_0\oplus\mathcal{E}(x_0))}\mathcal{E}^{v_0}(x_0)(-1)^{v_1^\top \mathcal{E}(x_1)}$
23	$[{u_0}^{x_0}]_{x_0,u_0} \otimes [(-1)^{{u_1}^\top x_1}]_{x_1,u_1}$	${u_0}^{x_0}(-1)^{u_1^\top x_1}$	$(-1)^{wt(v_0\oplus\mathcal{E}(x_0))}\mathcal{E}^{v_0}(x_0)2^{-n}(-1)^{v_1^\top \mathcal{E}(x_1)}$

Table 9: 49 Bases for the second order attacks and their effects (second part)

Index	Basis	Effect of input $\alpha_u(x)$	Effect of output $\beta^\star_{\mathcal{E}(x)}(v)$
24	$[{u_0}^{x_0}]_{x_0,u_0} \otimes [{u_1}^{x_1}]_{x_1,u_1}$	${u_0}^{x_0}{u_1}^{x_1}$	$(-1)^{wt(v_0 \oplus \mathcal{E}(x_0))}\mathcal{E}^{v_0}(x_0)(-1)^{wt(v_1 \oplus \mathcal{E}(x_1))}\mathcal{E}^{v_1}(x_1)$
25	$[{u_0}^{x_0}]_{x_0,u_0} \otimes [(-1)^{wt(u_1 \oplus x_1)}{u_1}^{x_1}]_{x_1,u_1}$	${u_0}^{x_0}(-1)^{wt(u_1 \oplus x_1)}{u_1}^{x_1}$	$(-1)^{wt(v_0 \oplus \mathcal{E}(x_0))}\mathcal{E}^{v_0}(x_0)\mathcal{E}^{v_1}(x_1)$
26	$[{u_0}^{x_0}]_{x_0,u_0} \otimes [{x_1}^{u_1}]_{x_1,u_1}$	${u_0}^{x_0}{x_1}^{u_1}$	$(-1)^{wt(v_0 \oplus \mathcal{E}(x_0))}\mathcal{E}^{v_0}(x_0)(-1)^{wt(v_1 \oplus \mathcal{E}(x_1))}v_1^{\mathcal{E}(x_1)}$
27	$[{u_0}^{x_0}]_{x_0,u_0} \otimes [(-1)^{wt(u_1 \oplus x_1)}{x_1}^{u_1}]_{x_1,u_1}$	${u_0}^{x_0}(-1)^{wt(u_1 \oplus x_1)}{x_1}^{u_1}$	$(-1)^{wt(v_0 \oplus \mathcal{E}(x_0))}\mathcal{E}^{v_0}(x_0)v_1^{\mathcal{E}(x_1)}$
28	$[(-1)^{wt(u_0 \oplus x_0)}{u_0}^{x_0}]_{x_0,u_0} \otimes [\delta_{u_1}(x_1)]_{x_1,u_1}$	$(-1)^{wt(u_0 \oplus x_0)}{u_0}^{x_0}\delta_{u_1}(x_1)$	$\mathcal{E}^{v_0}(x_0)\delta_{v_1}(\mathcal{E}(x_1))$
29	$[(-1)^{wt(u_0 \oplus x_0)}{u_0}^{x_0}]_{x_0,u_0} \otimes [2^{-n}(-1)^{{u_1}^\top x_1}]_{x_1,u_1}$	$(-1)^{wt(u_0 \oplus x_0)}{u_0}^{x_0}2^{-n}(-1)^{u_1^\top x_1}$	$\mathcal{E}^{v_0}(x_0)(-1)^{v_1^\top \mathcal{E}(x_1)}$
30	$[(-1)^{wt(u_0 \oplus x_0)}{u_0}^{x_0}]_{x_0,u_0} \otimes [(-1)^{{u_1}^\top x_1}]_{x_1,u_1}$	$(-1)^{wt(u_0 \oplus x_0)}{u_0}^{x_0}(-1)^{u_1^\top x_1}$	$\mathcal{E}^{v_0}(x_0)2^{-n}(-1)^{v_1^\top \mathcal{E}(x_1)}$
31	$[(-1)^{wt(u_0 \oplus x_0)}{u_0}^{x_0}]_{x_0,u_0} \otimes [{u_1}^{x_1}]_{x_1,u_1}$	$(-1)^{wt(u_0 \oplus x_0)}{u_0}^{x_0}{u_1}^{x_1}$	$\mathcal{E}^{v_0}(x_0)(-1)^{wt(v_1 \oplus \mathcal{E}(x_1))}\mathcal{E}^{v_1}(x_1)$
32	$[(-1)^{wt(u_0 \oplus x_0)}{u_0}^{x_0}]_{x_0,u_0} \otimes [(-1)^{wt(u_1 \oplus x_1)}{u_1}^{x_1}]_{x_1,u_1}$	$(-1)^{wt(u_0 \oplus x_0)}{u_0}^{x_0}(-1)^{wt(u_1 \oplus x_1)}{u_1}^{x_1}$	$\mathcal{E}^{v_0}(x_0)\mathcal{E}^{v_1}(x_1)$
33	$[(-1)^{wt(u_0 \oplus x_0)}{u_0}^{x_0}]_{x_0,u_0} \otimes [{x_1}^{u_1}]_{x_1,u_1}$	$(-1)^{wt(u_0 \oplus x_0)}{u_0}^{x_0}{x_1}^{u_1}$	$\mathcal{E}^{v_0}(x_0)(-1)^{wt(v_1 \oplus \mathcal{E}(x_1))}v_1^{\mathcal{E}(x_1)}$
34	$[(-1)^{wt(u_0 \oplus x_0)}{u_0}^{x_0}]_{x_0,u_0} \otimes [(-1)^{wt(u_1 \oplus x_1)}{x_1}^{u_1}]_{x_1,u_1}$	$(-1)^{wt(u_0 \oplus x_0)}{u_0}^{x_0}(-1)^{wt(u_1 \oplus x_1)}{x_1}^{u_1}$	$\mathcal{E}^{v_0}(x_0)v_1^{\mathcal{E}(x_1)}$
35	$[{x_0}^{u_0}]_{x_0,u_0} \otimes [\delta_{u_1}(x_1)]_{x_1,u_1}$	${x_0}^{u_0}\delta_{u_1}(x_1)$	$(-1)^{wt(v_0 \oplus \mathcal{E}(x_0))}v_0^{\mathcal{E}(x_0)}\delta_{v_1}(\mathcal{E}(x_1))$
36	$[{x_0}^{u_0}]_{x_0,u_0} \otimes [2^{-n}(-1)^{{u_1}^\top x_1}]_{x_1,u_1}$	${x_0}^{u_0}2^{-n}(-1)^{u_1^\top x_1}$	$(-1)^{wt(v_0 \oplus \mathcal{E}(x_0))}v_0^{\mathcal{E}(x_0)}(-1)^{v_1^\top \mathcal{E}(x_1)}$
37	$[{x_0}^{u_0}]_{x_0,u_0} \otimes [(-1)^{{u_1}^\top x_1}]_{x_1,u_1}$	${x_0}^{u_0}(-1)^{u_1^\top x_1}$	$(-1)^{wt(v_0 \oplus \mathcal{E}(x_0))}v_0^{\mathcal{E}(x_0)}2^{-n}(-1)^{v_1^\top \mathcal{E}(x_1)}$
38	$[{x_0}^{u_0}]_{x_0,u_0} \otimes [{u_1}^{x_1}]_{x_1,u_1}$	${x_0}^{u_0}{u_1}^{x_1}$	$(-1)^{wt(v_0 \oplus \mathcal{E}(x_0))}v_0^{\mathcal{E}(x_0)}(-1)^{wt(v_1 \oplus \mathcal{E}(x_1))}\mathcal{E}^{v_1}(x_1)$
39	$[{x_0}^{u_0}]_{x_0,u_0} \otimes [(-1)^{wt(u_1 \oplus x_1)}{u_1}^{x_1}]_{x_1,u_1}$	${x_0}^{u_0}(-1)^{wt(u_1 \oplus x_1)}{u_1}^{x_1}$	$(-1)^{wt(v_0 \oplus \mathcal{E}(x_0))}v_0^{\mathcal{E}(x_0)}\mathcal{E}^{v_1}(x_1)$
40	$[{x_0}^{u_0}]_{x_0,u_0} \otimes [{x_1}^{u_1}]_{x_1,u_1}$	${x_0}^{u_0}{x_1}^{u_1}$	$(-1)^{wt(v_0 \oplus \mathcal{E}(x_0))}v_0^{\mathcal{E}(x_0)}(-1)^{wt(v_1 \oplus \mathcal{E}(x_1))}v_1^{\mathcal{E}(x_1)}$
41	$[{x_0}^{u_0}]_{x_0,u_0} \otimes [(-1)^{wt(u_1 \oplus x_1)}{x_1}^{u_1}]_{x_1,u_1}$	${x_0}^{u_0}(-1)^{wt(u_1 \oplus x_1)}{x_1}^{u_1}$	$(-1)^{wt(v_0 \oplus \mathcal{E}(x_0))}v_0^{\mathcal{E}(x_0)}v_1^{\mathcal{E}(x_1)}$
42	$[(-1)^{wt(u_0 \oplus x_0)}{x_0}^{u_0}]_{x_0,u_0} \otimes [\delta_{u_1}(x_1)]_{x_1,u_1}$	$(-1)^{wt(u_0 \oplus x_0)}{x_0}^{u_0}\delta_{u_1}(x_1)$	$v_0^{\mathcal{E}(x_0)}\delta_{v_1}(\mathcal{E}(x_1))$
43	$[(-1)^{wt(u_0 \oplus x_0)}{x_0}^{u_0}]_{x_0,u_0} \otimes [2^{-n}(-1)^{{u_1}^\top x_1}]_{x_1,u_1}$	$(-1)^{wt(u_0 \oplus x_0)}{x_0}^{u_0}2^{-n}(-1)^{u_1^\top x_1}$	$v_0^{\mathcal{E}(x_0)}(-1)^{v_1^\top \mathcal{E}(x_1)}$
44	$[(-1)^{wt(u_0 \oplus x_0)}{x_0}^{u_0}]_{x_0,u_0} \otimes [(-1)^{{u_1}^\top x_1}]_{x_1,u_1}$	$(-1)^{wt(u_0 \oplus x_0)}{x_0}^{u_0}(-1)^{u_1^\top x_1}$	$v_0^{\mathcal{E}(x_0)}2^{-n}(-1)^{v_1^\top \mathcal{E}(x_1)}$
45	$[(-1)^{wt(u_0 \oplus x_0)}{x_0}^{u_0}]_{x_0,u_0} \otimes [{u_1}^{x_1}]_{x_1,u_1}$	$(-1)^{wt(u_0 \oplus x_0)}{x_0}^{u_0}{u_1}^{x_1}$	$v_0^{\mathcal{E}(x_0)}(-1)^{wt(v_1 \oplus \mathcal{E}(x_1))}\mathcal{E}^{v_1}(x_1)$
46	$[(-1)^{wt(u_0 \oplus x_0)}{x_0}^{u_0}]_{x_0,u_0} \otimes [(-1)^{wt(u_1 \oplus x_1)}{u_1}^{x_1}]_{x_1,u_1}$	$(-1)^{wt(u_0 \oplus x_0)}{x_0}^{u_0}(-1)^{wt(u_1 \oplus x_1)}{u_1}^{x_1}$	$v_0^{\mathcal{E}(x_0)}\mathcal{E}^{v_1}(x_1)$
47	$[(-1)^{wt(u_0 \oplus x_0)}{x_0}^{u_0}]_{x_0,u_0} \otimes [{x_1}^{u_1}]_{x_1,u_1}$	$(-1)^{wt(u_0 \oplus x_0)}{x_0}^{u_0}{x_1}^{u_1}$	$v_0^{\mathcal{E}(x_0)}(-1)^{wt(v_1 \oplus \mathcal{E}(x_1))}v_1^{\mathcal{E}(x_1)}$
48	$[(-1)^{wt(u_0 \oplus x_0)}{x_0}^{u_0}]_{x_0,u_0} \otimes [(-1)^{wt(u_1 \oplus x_1)}{x_1}^{u_1}]_{x_1,u_1}$	$(-1)^{wt(u_0 \oplus x_0)}{x_0}^{u_0}(-1)^{wt(u_1 \oplus x_1)}{x_1}^{u_1}$	$v_0^{\mathcal{E}(x_0)}v_1^{\mathcal{E}(x_1)}$