Integral Multiset: A Novel Framework for Integral Attacks over Finite Fields

Applications to MiMC and Chaghri

Weizhe Wang and Deng Tang

Shanghai Jiao Tong University, Shanghai, China, {SJTUwwz,dengtang}@sjtu.edu.cn

Abstract. In recent years, symmetric primitives that focus on arithmetic metrics over large finite fields, characterized as arithmetization-oriented (AO) ciphers, are widely used in advanced protocols such as secure multi-party computations (MPC), fully homomorphic encryption (FHE) and zero-knowledge proof systems (ZK). To ensure good performance in protocols, these AO ciphers are commonly designed with a small number of multiplications over finite fields and low multiplicative depths. This feature makes AO ciphers vulnerable to algebraic attacks, especially integral attacks. While a far-developed analysis for integral attacks on traditional block ciphers defined over $\mathbb{F}_2$ exists, there is still a lack of research on this kind of attacks over large finite fields. Previous integral attacks over large finite fields are primarily higher-order differential attacks, which construct distinguishers by simply utilizing algebraic degrees without fully exploiting other algebraic properties of finite fields.

In this paper, we propose a new concept called *integral multiset*, which provides a clear characterization of the integral property of multiset over the finite field $\mathbb{F}_{p^n}$. Based on multiplicative subgroups of finite fields, we present a new class of integral multisets that exhibits completely different integral property compared to the previously studied multisets based on vector subspaces over the finite field $\mathbb{F}_2$. In addition, we also present a method for merging existing integral multisets to create a new one with better integral property. Furthermore, combining with monomial detection techniques, we propose a framework for searching for integral distinguishers based on integral multisets.

We apply our new framework to some competitive AO ciphers, including MiMC and Chaghri. For all these ciphers, we successfully find integral distinguishers with lower time and data complexity. Especially for MiMC, the complexity of some distinguishers we find is only a half or a quarter of the previous best one. Due to the specific algebraic structure, all of our results could not be obtained by higher-order differential attacks. Furthermore, our framework perfectly adapts to various monomial detection techniques like general monomial prediction proposed by Cui et al. at ASIACRYPT 2022 and coefficient grouping invented by Liu et al. at EUROCRYPT 2023. We believe that our work will provide new insight into integral attacks over large finite fields.

Keywords: integral attack · integral multiset · multiplicative subgroup · monomial detection

1 Introduction

In recent years, symmetric-key cryptographic primitives, such as stream ciphers, block ciphers, hash functions, pseudorandom generators, message authentication codes, etc., also serve as fundamental building blocks for many advanced protocols like MPC, FHE and ZK. In the traditional design of symmetric primitives, linear operations and non-linear operations typically have similar costs. However, this is not the case in many MPC/FHE/ZK protocols, where non-linear operations, such as multiplications, often become the bottleneck. As a result, traditional symmetric-key primitives like AES are not efficient and the design of symmetric-key primitives specifically tailored for advanced cryptographic protocols becomes imperative. Since the invention of block cipher LowMC [ARS+15], a large number of MPC/FHE/ZK-friendly symmetric-key primitives have been developed. These include MPC-friendly designs such as MiMC [AGR+16], Ciminion [DGGK21] and AIM [DKR+22]; FHE-friendly designs such as Rasta [DEG+18], Pasta [DGH+21] and Chaghri [AMT22]; and ZK-friendly designs such as MARVELlous [AD18], Poseidon [GKR+21] and Griffin [GHR+23].

Some innovative primitives that prioritize arithmetic metrics are known as AO ciphers. The designs of AO ciphers differ significantly from traditional ones. In contrast to using small S-boxes as non-linear layers, AO ciphers are preferred to employ non-linear functions with explicit and concise algebraic representations over large finite fields. MiMC and Chaghri are two classical AO ciphers. MiMC, a block cipher with extremely low multiplicative complexity, was proposed at ASIACRYPT 2016. Since its birth, MiMC has gained widespread attention, and subsequently, several variants of MiMC have emerged, such as GMiMC [AGP+19] and HadesMiMC [GLR+20]. Due to the algebraic simplicity of round function, MiMC has become the preferred choice for many use cases [GRR+16, RSS17]. It also served as a candidate for Zcash[1] and a baseline in “STARK-Friendly Hash Challenge” competition[2]. Recently, MiMC has been applied in zkBridge [XZC+22], which is an efficient cross-chain bridge proposed at ACM CCS 2022. Chaghri, proposed at ACM CCS 2022 [AMT22], is an FHE-friendly block cipher defined over the finite field $\mathbb{F}_{2^{63}}$ with efficient circuit implementation. Its implementation with HElib can outperform AES by about 65%.

The structure of AO ciphers has a great influence on the type of attacks that can be mounted. Statistical attacks (including linear [Mat94] and differential [BS91] ones), which are considered to be two of the most powerful classical cryptanalytic tools, do not seem to pose a significant threat to the security of AO ciphers. However, AO ciphers are naturally vulnerable to algebraic attacks, especially higher-order differential attacks [EGL+20, BCD+20, BCP23, CHWW22, LAW+23, LGB+23]. Actually, higher-order differential attacks are a specific case of integral attacks over finite fields of characteristic 2. Integral attacks involve the summation over a multiset, while higher-order differential attacks require constraining the multiset to a linear subspace. In higher-order differential attacks, accurately evaluating the algebraic degree is of utmost importance, which can be achieved through direct estimation [CV02, BCD11, BC11, CGG+22] or monomial detection [TIHM17, HSWW20, CHWW22, LAW+23]. There have been numerous research achievements and well-established research methodologies for higher-order differential attacks. However, in the case of integral attacks, particularly over large finite fields, the research is rather insufficient. To the best of our knowledge, most of previously proposed integral distinguishers over large finite fields are constructed solely based on algebraic degree. This greatly limits the integral attacks, thus failing to showcase their significant potentials on symmetric primitives for MPC/FHE/ZK protocols. Hence, proposing a theoretical framework for integral attacks over large finite fields is of great significance. It would enable more systematic research on integral attacks and facilitate a more accurate evaluation of the security of symmetric primitives against integral attacks.

In this paper, we introduce a new concept called *integral multiset* for better characterizing the integral property and propose a novel theoretical framework for finding integral distinguishers targeting MPC/FHE/ZK-friendly symmetric primitives. In order to demonstrate the effectiveness of our framework, we apply it to a selection of competitive AO ciphers, including MiMC and Chaghri. Consequently, we successfully find several

[1] https://github.com/zcash/zcash/issues/2233

[2] https://starkware.co/hash-challenge/

distinguishers with lower complexity compared to previous ones [CHWW22, LAW+23]. In particular, the contributions of this paper are summarized below.

1. We introduce a new concept called integral multiset, which can be seen as a generalization of three-subset division property without unknown set. This concept aims to better characterize the integral property for a given multiset of finite fields. Moreover, we provide a classical construction method for integral multisets based on multiplicative subgroups of finite fields. Additionally, we present an approach for merging existing integral multisets to create a new one with improved integral property. Furthermore, by combining with monomial detection techniques, we propose a framework for searching for integral distinguishers using integral multisets.

2. By combining our new method with two state-of-the-art monomial detection techniques, namely monomial prediction and coefficient grouping, we applied our framework to analyze the symmetric primitives: MiMC and Chaghri. The results are promising, as we successfully find some integral distinguishers with lower time and data complexity compared to the previous best results. In Table 2, 5, 6, we present the results for the standard versions of MiMC, Feistel MiMC, and Chaghri, respectively. Additionally, in Table 3, we show the results for MiMC with different exponents, while Table 4 displays the outcomes of merged multisets. It is worth noting that all the distinguishers we found could not be obtained by using higher-order differential attacks.

Based on the above results, we believe that our framework is useful for carrying out the integral attacks over large finite fields and worth further investigation.

The rest of this paper is organized as follows. In Section 2, we give the notations and introduce some background knowledge about finite field, integral attack and monomial detection. Section 3 is the main section of this paper. In Subsection 3.1, we introduce the concept of integral multiset and reveal its relationship with integral distinguishers. In Subsection 3.2, we show how to construct valuable integral multisets based on multiplicative subgroups of finite fields. In Subsection 3.3, we present a method for generating an integral multiset with better integral property by merging existing ones. In Subsection 3.4, a framework for determining integral distinguishers based on monomial detection technique is introduced. Section 4 shows the details of our experiments on different competitive AO ciphers. We conclude the paper in Section 5.

2 Preliminaries

2.1 Notations

In this paper, we will use the following notations. We denote a finite field with q elements as $\mathbb{F}_q$. We denote $\mathbb{F}_q^* = \mathbb{F}_q \setminus \{0\}$ and it is well known that $\mathbb{F}_q^*$ is a cyclic group. For a finite field $\mathbb{F}_q$, q must be a power of a prime, i.e., $q = p^n$. Let $\mathbb{F}_{p^n}$ denote the finite field with characteristic p and $\mathbb{F}_{p^n}^t$ denote the t-dimensional vector space over the finite field $\mathbb{F}_{p^n}$. We denote the polynomial ring over $\mathbb{F}_{p^n}$ as $\mathbb{F}_{p^n}[x]$ and the set of integers $\{i : a \leq i \leq b\}$ as $[a, b]$. We use bold italic latters to represent vectors, e.g. $\boldsymbol{a} \in \mathbb{F}_q^t$ denotes the vector $\boldsymbol{a} = (a_0, a_1, \ldots, a_{t-1})$, where $a_i \in \mathbb{F}_q$. For any $a \in [0, 2^n - 1]$, we have $a = \sum_{i=0}^{n-1} a_i \cdot 2^i, a_i \in \{0, 1\}$, and its Hamming weight is $wt(a) = |\{i : a_i \neq 0\}|$. Correspongdingly, the Hamming weight of $\boldsymbol{u} \in [0, 2^n - 1]^t$ is $wt(\boldsymbol{u}) = \sum_{i=0}^{t-1} wt(u_i)$. We use $\oplus$ as addition over $\mathbb{F}_2$ or $\mathbb{F}_{2^n}$ and $+$ over $\mathbb{F}_p$ or $\mathbb{F}_{p^n}$. For a multiset $\mathbb{X}$, we use $|\mathbb{X}|$ to represent the number of elements or, in other words, the cardinality of $\mathbb{X}$. For simplicity, we denote a set of all integers $j \in [0, n-1]$ satisfying $d \mid j$ as $\mathcal{D}_d^n$ and a set of all integers $j \in [0, n-1]$ satisfying $d \nmid j$ as $\mathcal{UD}_d^n$, i.e., $\mathcal{D}_d^n = \{j \in [0, n-1] : d \mid j\}, \mathcal{UD}_d^n = \{j \in [0, n-1] : d \nmid j\}$.

Let $F : \mathbb{F}_{p^n}^t \to \mathbb{F}_{p^n}$ be a function over $\mathbb{F}_{p^n}[x_0, x_1, \ldots, x_{t-1}]/(x_0^{p^n} - x_0, x_1^{p^n} - x_1, \ldots, x_{n-1}^{p^n} - x_{n-1})$. Function F can be uniquely expressed as

$$F(\boldsymbol{x}) = F(x_0, x_1, \ldots, x_{t-1}) = \sum_{\boldsymbol{u} \in \{0,1,\ldots,p^n-1\}^t} a_{\boldsymbol{u}} \boldsymbol{x}^{\boldsymbol{u}} = \sum_{i_0=0}^{p^n-1} \sum_{i_1=0}^{p^n-1} \cdots \sum_{i_t=0}^{p^n-1} a_{i_0,i_1,\ldots,i_t} x_0^{i_0} x_1^{i_1} \cdots x_t^{i_t},$$

where every coefficient $a_{\boldsymbol{u}} \in \mathbb{F}_{p^n}$ and $\boldsymbol{x}^{\boldsymbol{u}} = \pi_{\boldsymbol{u}}(\boldsymbol{x}) = \prod_{i=0}^{n-1} x_i^{u_i}$ is called a monomial. In particular, the function $G : \mathbb{F}_{2^n}^t \to \mathbb{F}_{2^n}$ can be represented as

$$G(\boldsymbol{x}) = G(x_0, x_1, \ldots, x_{t-1}) = \bigoplus_{\boldsymbol{u} \in \{0,1,\ldots,2^n-1\}^t} a_{\boldsymbol{u}} \boldsymbol{x}^{\boldsymbol{u}} = \bigoplus_{i_0=0}^{2^n-1} \bigoplus_{i_1=0}^{2^n-1} \cdots \bigoplus_{i_t=0}^{2^n-1} a_{i_0,i_1,\ldots,i_t} x_0^{i_0} x_1^{i_1} \cdots x_t^{i_t},$$

If the coefficient of $\boldsymbol{x}^{\boldsymbol{u}}$ is zero, we say $\boldsymbol{x}^{\boldsymbol{u}}$ is not contained by F, denoted by $\boldsymbol{x}^{\boldsymbol{u}} \nrightarrow F$. Otherwise, $\boldsymbol{x}^{\boldsymbol{u}}$ is contained by F, denoted by $\boldsymbol{x}^{\boldsymbol{u}} \to F$. The univariate degree of F on variable x_j is denoted by $d_j(F)$ and defined as:

$$d_j(F) = \max\{i_j : 0 \leq i_j \leq 2^n - 1, a_{i_0,i_1,\ldots,i_t} \neq 0\}.$$

Moreover, the algebraic degree $d(F)$ of F can be computed as

$$\delta(F) = \max\{\sum_{j=1}^{t} wt(i_j) : 0 \leq i_j \leq 2^n - 1, a_{i_0,i_1,\ldots,i_t} \neq 0\}.$$

2.2 Finite field

Finite field is an important algebraic structure that consists of a finite set of elements along with two operations, addition and multiplication. $\mathbb{F}_p$ with a large prime number p and $\mathbb{F}_{2^n}$ are widely used in cryptography. A subfield is a subset of a field that is itself a field, inheriting the same set of operations and satisfying the same properties as the original field. Multiplicative group of a finite field is the group under multiplication operation with respect to this finite field. For finite fields and multiplicative groups, we list some useful theorems here.

Theorem 1. *Let $\mathbb{F}_{p^n}$ be the finite field with p^n elements. Every subfield of $\mathbb{F}_{p^m}$ has p^m elements for some positive integer m dividing n. Conversely, for any positive integer m dividing n there is a unique subfield of $\mathbb{F}_{p^m}$ of order p^m.*

Theorem 2. *The multiplicative group $\mathbb{F}_q^*$ of a finite field $\mathbb{F}_q$ is cyclic.*

Theorem 3. *Let H be a cyclic group generated by x, if $|H| = n$ is finite, then for each positive integer $a \mid n$ there is a unique subgroup of H of order a. This subgroup is the cyclic group $\langle x^d \rangle$, where $d = \frac{n}{a}$.*

2.3 Integral attack

The integral attack was firstly introduced by Daemen et al. [DKR97] and formalized by Knudsen and Wagner [KW02]. It is a chosen plaintext attack and a theoretical generalization higher-order differential cryptanalysis [Lai94]. In an integral attack, the crucial aspect is to find an input multiset for which the sum of their output is a key-independent constant value. Such an input multiset is called an *integral distinguisher*. A concrete definition of integral distinguisher is given as below.

Definition 1 (Integral Distinguisher)**.** Let $E^r : \mathbb{F}_{p^n}^t \to \mathbb{F}_{p^n}^t$ be an r-round cipher and $\mathbb{X}$ be a multiset whose elements belong to $\mathbb{F}_{p^n}^t$. If the sum of ciphertext over $\mathbb{X}$ satisfies

$$\sum_{\boldsymbol{x}\in\mathbb{X}} E^r(\boldsymbol{x}) = \mathbf{0},$$

then $\mathbb{X}$ is an integral distinguisher of cipher E^r.

An integral distinguisher should be able to distinguish a cipher from a random permutation. In previous bit-based integral attacks, the input multiset usually consists of plaintexts taking all possible combinations in d input positions, whereas the remaining bits take a fixed value. According to the theory of higher-order differential cryptanalysis [Lai94], the sum of any Boolean function with algebraic degree smaller than d over such a multiset would become zero and naturally form an integral distinguisher.

With an integral distinguisher, the attacker is able to recover partial information of secret key by guessing and then the remain part of key would be recovered by a brute-force attack. Particularly, this process can be accelerated with the partial-sum technique [FKL+01], the meet-in-the-middle technique [SW13] and the fast fourier transform technique [TA14]. However, the trivial key-guessing approach is inefficient on AO ciphers like MiMC. In [EGL+20], Eichlseder et al. proposed a new strategy of key recovery by constructing a low-degree polynomial.

2.4 Monomial detection

Monomial detection is an important method used to identify the presence of monomials in a given function. The ability to quickly and accurately determine whether a certain monomial is contained by a complicated function is crucial for estimating the (algebraic) degree of function and constructing algebraic attacks. Monomial prediction [HSWW20], based on the division property [Tod15, TM16, HLM+20], is a monomial detection technique that is widely employed in constructing cube attacks and integral attacks. Subsequently, the generalized monomial prediction [CHWW22] has been proposed, showcasing its formidable capabilities in attacking AO ciphers like MiMC. In addition to monomial prediction, coefficient grouping [LAW+23, LGB+23] serves as another monomial detection technique. The main idea of coefficient grouping is to give a compact expression for all possible monomials and encode it as some constraints. Compared to monomial prediction, coefficient grouping is more efficient and lightweight. However, its applicability is somewhat limited and currently confined to field-based ciphers with quadratic round function.

3 Integral distinguishers based on integral multisets

3.1 Integral multiset

Let $y = F(x, k)$ be a function from $\mathbb{F}_{p^n} \times \mathbb{F}_{p^n}$ to $\mathbb{F}_{p^n}$, we can express it as

$$F(x,k) = \sum_{i=0}^{p^n-1} a_i(k)x^i,$$

where the coefficients $a_i(k)$ are related to k. Given a multiset $\mathbb{X}$ whose elements belong to $\mathbb{F}_{p^n}$, then the sum of F on $\mathbb{X}$ is

$$\sum_{x\in\mathbb{X}} F(x,k) = \sum_{x\in\mathbb{X}}\sum_{i=0}^{p^n-1} a_i(k)x^i = \sum_{i=0}^{p^n-1} a_i(k)\sum_{x\in\mathbb{X}} x^i = \sum_{a_i\neq 0} a_i(k)\sum_{x\in\mathbb{X}} x^i. \tag{1}$$

According to Equation (1), we know that if $\sum_{x\in\mathbb{X}} x^i = 0$ for all i such that $a_i \neq 0$, the sum equals to 0; otherwise, the value of sum would depend on k. Then, a natural question raises, that is, how to construct a proper multiset such that the sum of a specific function $F(x,k)$ becomes 0. As is known, the basic component of function is the monomial. For any monomial, the value obtained by summing it over a set is deterministic. Therefore, approaching the problem from the perspective of monomials will make the entire issue clearer and more comprehensive. In order to conduct a more in-depth investigation, we proposed a new concept called the *integral multiset*.

Definition 2 (Integral Multiset)**.** Given a multiset $\mathbb{X}$ whose elements belong to $\mathbb{F}_{p^n}$ and a set $E \subseteq [0, p^n-1]$, if the multiset $\mathbb{X}$ satisfies $\sum_{x\in\mathbb{X}} x^j = 0$ for all $j \in E$ and $\sum_{x\in\mathbb{X}} x^j$ is always non-zero for $j \in [0, p^n-1] \setminus E$, then we say that $\mathbb{X}$ belongs to an E-integral multiset, or $\mathbb{X}$ is an E-integral multiset, where E is the set indicating integral property.

According to Definition 2, it is obvious that for any multiset $\mathbb{X}$, there must exist a set E such that $\mathbb{X}$ belongs to an E-integral multiset. The direct approach to compute an E is to iterate over all monomials. To enhance readers' comprehension of the concept of the integral multiset, a simple example is presented below.

Example 1. Let α be a primitive element of $\mathbb{F}_{2^4}$ with the irreducible polynomial $f(x) = x^4 + x + 1$. Let $\mathbb{X}$ be a multiset whose element belong to $\mathbb{F}_{2^4}$. As an example, we prepare the multiset as

$$\mathbb{X} := \{0, \alpha, \alpha, \alpha, \alpha^6, \alpha^8, \alpha^8, \alpha^{11}\}.$$

Table 1 calculates the sum of x^j for different j.

Table 1: Summation table of $\mathbb{X} := \{0, \alpha, \alpha, \alpha, \alpha^6, \alpha^8, \alpha^8, \alpha^{11}\}$.

j \ x	0	α	α	α	α^6	α^8	α^8	α^{11}	$\bigoplus x^j$
0	1	1	1	1	1	1	1	1	0
1	0	α	α	α	α^6	α^8	α^8	α^{11}	0
2	0	α^2	α^2	α^2	α^{12}	α	α	α^7	0
3	0	α^3	α^3	α^3	α^3	α^9	α^9	α^3	α^3
4	0	α^4	α^4	α^4	α^9	α^2	α^2	α^{14}	0
5	0	α^5	α^5	α^5	1	α^{10}	α^{10}	α^{10}	0
6	0	α^6	α^6	α^6	α^6	α^3	α^3	α^6	α^6
7	0	α^7	α^7	α^7	α^{12}	α^{11}	α^{11}	α^2	0
8	0	α^8	α^8	α^8	α^3	α^4	α^4	α^{13}	0
9	0	α^9	α^9	α^9	α^9	α^{12}	α^{12}	α^9	α^9
10	0	α^{10}	α^{10}	α^{10}	1	α^5	α^5	α^5	0
11	0	α^{11}	α^{11}	α^{11}	α^6	α^{13}	α^{13}	α	0
12	0	α^{12}	α^{12}	α^{12}	α^{12}	α^6	α^6	α^{12}	α^{12}
13	0	α^{13}	α^{13}	α^{13}	α^3	α^{14}	α^{14}	α^8	0
14	0	α^{14}	α^{14}	α^{14}	α^9	α^7	α^7	α^4	0

For all $j \in H = \mathcal{UD}_3^{15} \cup \{0\}$, the sum $\bigoplus_{x\in\mathbb{X}} x^j$ becomes 0. Therefore, $\mathbb{X}$ is an H-integral multiset.

Specifically, some multisets exhibit a constant integer value in the non-zero part, and this property can be advantageous when constructing an integral distinguisher. To effectively represent such multisets, we present the definition of *integral multiset with two levels*.

Definition 3 (Integral Multiset with Two Levels)**.** Let $\mathbb{X}$ be a multiset whose elements belong to $\mathbb{F}_{p^n}$ and $E \subseteq [0, p^n - 1]$. When the multiset $\mathbb{X}$ satisfies:

$$\sum_{x \in \mathbb{X}} x^j = \begin{cases} 0, & j \in E \\ c, & \text{otherwise} \end{cases},$$

where $c \in \mathbb{F}_{p^n}$ is a constant, we call $\mathbb{X}$ belongs to an E-integral multiset with two levels, or $\mathbb{X}$ is an E-integral multiset with two levels.

For an integral multiset with two levels, consider the following example: $\mathbb{Y} := \{0, 1, \alpha^5, \alpha^{10}\}$, where α is a primitive element of $\mathbb{F}_{2^4}$ with the irreducible polynomial $f(x) = x^4 + x + 1$. In the case of $\mathbb{Y}$, its indicated set is the same as $\mathbb{X}$ in Example 1, while $\bigoplus_{y \in \mathbb{Y}} y^j$ always evaluates to 1 in the non-zero part. It is not crucial for the values in the non-zero part to be fixed constants when directly constructing an integral distinguisher, while combining different two-level integral multisets is more likely to yield a superior integral multiset.

3.1.1 Comparsion with division property

Compared to the word-based division property [Tod15], the integral multiset is more refined. Firstly, the word-based division property treats the non-zero sum value as "unknown", while the integral multiset calculates the exact value of sum. Secondly, the word-based division property partitions the exponents solely based on Hamming weights, whereas the integral multiset partitions them using an indicated set E. In summary, the integral multiset can characterize the summation property more accurately.

To compare it with the bit-based division property [TM16, HLM+20], we firstly provide the definition of *bit-based integral multiset.*

Definition 4 (Bit-based Integral Multiset)**.** Given a multiset $\mathbb{X}$ whose elements belong to $\mathbb{F}_2^n$ and a set $E \subseteq \{0,1\}^n$, if the multiset $\mathbb{X}$ satisfies $\bigoplus_{\boldsymbol{x} \in \mathbb{X}} \boldsymbol{x}^{\boldsymbol{j}} = 0$ for all $\boldsymbol{j} \in E$ and $\bigoplus_{\boldsymbol{x} \in \mathbb{X}} \boldsymbol{x}^{\boldsymbol{j}}$ is always 1 for $\boldsymbol{j} \in \{0,1\}^n \setminus E$, then we say that $\mathbb{X}$ belongs to an E-bit-based integral multiset, or $\mathbb{X}$ is an E-bit-based integral multiset.

Since the only non-zero element in $\mathbb{F}_2$ is 1, there is no difference between bit-based integral multiset and bit-based integral multiset with two levels. According to Definition 4, it is evident that the bit-based integral multiset is essentially equivalent to the three-subset division property without unknown set, where E corresponds to the set of elements that appear an odd number of times in $\mathbb{L}$ in the latter case. Therefore, the core idea behind the integral multiset can be regarded as a generalization of three-subset division property without unknown set over $\mathbb{F}_{p^n}$.

3.1.2 The relation between integral multisets and integral distinguishers

Consider a field-based cipher $Enc : \mathbb{F}_{p^n} \times \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$, it can be represented by a function:

$$F(x, k) = \sum_{i=0}^{p^n-1} \sum_{j=0}^{p^n-1} a_{i,j} x^i k^j = \sum_{i=0}^{p^n-1} g_i(k) x^i = \sum_{i \in I} g_i(k) x^i,$$

where x is the plaintext, k is the secret key and $I = \{j : g_j(k) \neq 0\} = \{j : x^j \to F(x,k)\}$ represents the set of indices corresponding to non-zero coefficients. Given that an input multiset $\mathbb{X}$ is an E-integral multiset, the sum of $F(x,k)$ over $\mathbb{X}$ can be computed as:

$$\sum_{x \in \mathbb{X}} F(x,k) = \sum_{x \in \mathbb{X}} \sum_{i \in I} g_i(k) x^i = \sum_{i \in I \cap E} g_i(k) \sum_{x \in \mathbb{X}} x^i + \sum_{i \in I - E} g_i(k) \sum_{x \in \mathbb{X}} x^i = \sum_{i \in I - E} g_i(k) c_i.$$

Here, we use $I - E$ to represent the set $\{i : i \in I, i \notin E\}$. The sum would become zero when $I - E = \emptyset$, i.e. $I \subseteq E$. Therefore, a proposition is derived.

Proposition 1. *Let $F(x,k) : \mathbb{F}_{p^n} \times \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$ be the function of a field-based cipher and $\mathbb{X}$ be a multiset whose elements belong to $\mathbb{F}_{p^n}$. If $\mathbb{X}$ is an E-integral multiset, then the following statements are equivalent:*

1. *For each monomial $x^j \to F(x,k)$, its expononet satisfies $j \in E$;*
2. *For any secret key $k \in \mathbb{F}_{p^n}$, $\sum_{x\in\mathbb{X}} F(x,k) = 0$;*
3. *$\mathbb{X}$ is an integral distinguisher of $F(x,k)$.*

According to Proposition 1, it appears that a multiset $\mathbb{X}$ would be more useful with a larger indicated set E. However, the cardinality $|\mathbb{X}|$ is also an important role in integral attacks as it is closely related to the complexity of attacks. Additionally, an integral distinguisher should be able to distinguish a cryptographic algorithm from a random permutation. Therefore, some trivial multisets, such as a multiset with p identical elements over $\mathbb{F}_{p^n}$, are useless.

3.2 A new class of integral multisets based on multiplicative subgroups

In this subsection, we will introduce how to construct valuable integral multisets with some algebraic structures such as subfields and multiplicative subgroups. First of all, the sum of certain monomial over a finite field can be summarized as the following lemma:

Lemma 1. *Let p be a prime and x^j be a monomial in $\mathbb{F}_{p^n}[x]$, $p^n > 2$, j is a non-negative integer, then the sum of x^j over the whole finite field is*

$$\sum_{x\in\mathbb{F}_{p^n}} x^j = \begin{cases} 0, & (p^n-1) \nmid j \text{ or } j = 0 \\ p-1, & (p^n-1) \mid j \text{ and } j \neq 0. \end{cases}$$

Proof. When $j = 0$, it is clear that

$$\sum_{x\in\mathbb{F}_{p^n}} x^0 = \sum_{x\in\mathbb{F}_{p^n}} 1 = p = 0.$$

When $(p^n-1) \nmid j$, we have

$$(1-g^j)\sum_{x\in\mathbb{F}_{p^n}} x^j = (1-g^j)\sum_{x\in\mathbb{F}_{p^n}^*} x^j = (1-g^j)\sum_{i=0}^{p^n-2}(g^i)^j = \sum_{i=0}^{p^n-2}(g^i)^j - \sum_{i=0}^{p^n-2}(g^i)^j = 0,$$

where g is a primitive element of $\mathbb{F}_{p^n}^*$. Since $1-g^j \neq 0$, $\sum_{x\in\mathbb{F}_{p^n}} x^j = 0$.
When $(p^n-1) \mid j$ and $j \neq 0$, we have

$$\sum_{x\in\mathbb{F}_{p^n}} x^j = \sum_{x\in\mathbb{F}_{p^n}^*} x^j = p^n - 1 = p-1.$$

The proof is completed. □

Lemma 1 implies that $\mathbb{F}_{p^n}$ is a $[0, p^n-2]$-integral multiset with two levels. Though most of function would yield a zero sum over this multiset, its large cardinality makes it ineffective in attacks. Consequently, a subfield with smaller cardinality will be a better choice when constructing an integral distinguisher.

Theorem 4. *Let m be an integer dividing n, p be a prime and $\mathbb{F}_{p^m}$ be a subfield of $\mathbb{F}_{p^n}$. Consider the multiset $H = \mathbb{F}_{p^m}$, then H belongs to an $\mathcal{UD}_{p^m-1}^{p^n} \cup \{0\}$-integral multiset with two levels.*

The proof of Theorem 4 is similar to Lemma 1 and thus is skipped here. Theorem 4 states that if a function $f(x)$ does not contain a monomials x^j satisfying $j \neq 0$ and $(p^m - 1) \mid j$, then we can find a multiset $\mathbb{X}$ with cardinality p^m such that $\sum_{x \in \mathbb{X}} f(x) = 0$. Instead of solely focusing on the algebraic degree of $f(x)$, this approach pays more attention on the presence of monomial $x^{c \cdot (p^m - 1)}, c > 0$, which presents a new perspective for constructing an integral distinguisher. Furthermore, it is worth noting that Lemma 1 can be extended to multivariate cases, which might also be useful in constructing integral distinguishers.

Corollary 1. *Let p be a prime and $x_0^{j_0} x_1^{j_1} \cdots x_{t-1}^{j_{t-1}}$ be a monomial in $\mathbb{F}_{p^n}[x_0, x_1, \ldots, x_{t-1}], p^n > 2$, $j_0, j_1, \ldots, j_{t-1}$ are non-negative integers, then*

$$\sum_{(x_0, x_1, \ldots, x_{t-1}) \in \mathbb{F}_{p^n}^t} x_0^{j_0} x_1^{j_1} \cdots x_{t-1}^{j_{t-1}} = \begin{cases} (p-1)^t, & (p^n - 1) \mid j_0, (p^n - 1) \mid j_1, \ldots, (p^n - 1) \mid j_{t-1}, \\ 0, & \textit{otherwise}. \end{cases}$$

Although Theorem 4 presents a new method for constructing an integral distinguisher, the requirement of $m \mid n$ in Theorem 1 is too strict, significantly limiting the number of available choices. For example, consider a finite field with size 2^{129}, the size of its subfields is restricted to either 2^3 or 2^{43}. Actually, multiplicative subgroups of finite fields have nice integral property as well.

Lemma 2. *Let p be a prime, x^j be a monomial in $\mathbb{F}_{p^n}[x]$ and G be a multiplicative subgroup of $\mathbb{F}_{p^n}, p^n > 2$, j is a non-negative integer, then the sum of x^j over G is*

$$\sum_{x \in G} x^j = \begin{cases} 0, & |G| \nmid j \\ |G|, & |G| \mid j. \end{cases}$$

It should be noted that $\sum_{x \in G} x^0$ would not be 0 because the cardinality of G is not necessarily a multiple of p. This property is not desirable as, in most cases, the functions we analyze typically include a constant component. To address this issue, one possible approach is to augment the multiset with several 0s, thereby ensuring its cardinality becomes a multiple of p.

Theorem 5. *Let p be a prime, G be a multiplicative subgroup of $\mathbb{F}_{p^n}$, S be a multiset consisting only of* 0*s and $p \mid (|S| + |G|)$. Consider the multiset $H = G \cup S$, then H belongs to an $\mathcal{UD}_{|G|}^{p^n} \cup \{0\}$-integral multiset with two levels.*

Proof. When $j = 0$, the sum is

$$\sum_{x \in H} x^0 = \sum_{x \in H} 1 = |S| + |G| = 0.$$

When $|G| \nmid j$, we have

$$(1 - g^j) \sum_{x \in H} x^j = (1 - g^j) \sum_{x \in G} x^j = (1 - g^j) \sum_{i=0}^{|G|-1} (g^i)^j = \sum_{i=0}^{|G|-1} (g^i)^j - \sum_{i=0}^{|G|-1} (g^i)^j = 0,$$

where g is a generator of multiplicative subgroup G. Since $1 - g^j \neq 0$, $\sum_{x \in H} x^j = 0$. When $|G| \mid j$ and $j \neq 0$, we have

$$\sum_{x \in H} x^j = \sum_{x \in G} x^j = |G|.$$

The proof is completed. □

In the case of binary extension field $\mathbb{F}_{2^n}$, only one zero element needs to be added. Therefore, the size of subgroup $|G|$ will play a dominant role in determining the cardinality of multiset H. According to Theorem 2 and 3, the size of multiplicative subgroup G is a factor of $p^n - 1$, providing a wider range of choices. For instance, consider a multiplicative group with size $p = p_1 \cdot p_2 \cdot p_3$, it would yield a total of 6 meaningful subgroups with different sizes.

3.3 Creating integral multisets with better integral property

Having a wider ranger of options for integral multiset will increase the probability of discovering superior integral distinguishers. In Subsection 3.2, we present several integral multisets based on subfields and multiplicative subgroups, which provides a method for utilizing the specific structure of finite fields to construct integral multisets with desirable integral properties. In this subsection, we will introduce a method that merges small multisets to create a multiset with improved integral property.

Let $\mathbb{X}$ belongs to an E-integral multiset and $F(x,k) : \mathbb{F}_{p^n} \times \mathbb{F}_{p^n} \to \mathbb{F}_{p^n}$ be the function of field-based cipher. According to Proposition 1, the sum of $F(x,k)$ over $\mathbb{X}$ would not be zero if there exists a monomial $x^j \to F(x,k)$ and $j \notin E$, and thus this does not fulfill the requirements of an integral distinguisher. If the set of monomials that do not belong to E satisfies specific conditions, it is possible to construct an integral distinguisher by merging $\mathbb{X}$ and its cosets.

Proposition 2. *Let p be a prime, $\mathbb{X}$ be an E-integral multiset whose elements belong to $\mathbb{F}_{p^n}$ and $a\mathbb{X} := \{ax : x \in \mathbb{X}\}$, $b\mathbb{X} := \{bx : x \in \mathbb{X}\}$ be its cosets, where $a, b \in \mathbb{F}_{p^n}$. Consider a new multiset $\mathbb{Y} = \mathbb{X} \cup a\mathbb{X} \cup b\mathbb{X}$, if $a^i + b^i + 1 = 0$, then $\mathbb{Y}$ belongs to an $E \cup \{i, p \cdot i, p^2 \cdot i, \ldots\}$-integral multiset.*

Proof. Because $\mathbb{Y} = \mathbb{X} \cup a\mathbb{X} \cup b\mathbb{X}$, we have

$$\sum_{x \in \mathbb{Y}} x^j = \sum_{x \in \mathbb{X}} x^j + \sum_{x \in a\mathbb{X}} x^j + \sum_{x \in b\mathbb{X}} x^j = \sum_{x \in \mathbb{X}} (x^j + (ax)^j + (bx)^j) = (1 + a^j + b^j) \sum_{x \in \mathbb{X}} x^j.$$

For $j \in E$, it is clear that

$$\sum_{x \in \mathbb{Y}} x^j = (1 + a^j + b^j) \sum_{x \in \mathbb{X}} x^j = (1 + a^j + b^j) \cdot 0 = 0.$$

For $j \in \{i, p \cdot i, p^2 \cdot i, \ldots\}$, there exists l such that $j = p^l \cdot i$ and the sum becomes

$$\sum_{x \in \mathbb{Y}} x^j = (1 + a^j + b^j) \sum_{x \in \mathbb{X}} x^j = (1 + a^{p^l \cdot i} + b^{p^l \cdot i}) \sum_{x \in \mathbb{X}} x^j = (1 + a^i + b^i)^{p^l} \sum_{x \in \mathbb{X}} x^j = 0.$$

Equation $(1 + a + b)^p = 1 + a^p + b^p$ holds as the characteristic of $\mathbb{F}_{p^n}$ is p. □

Finding a pair of elements (a, b) that satisfy $a^i + b^i + 1 = 0$ for a given exponent i in a finite field is not a challenging task. Proposition 2 shows how to construct a better multiset so that its indicated set can cover all the monomials of a certain function. However, merging also leads to a larger cardinality, which means higher complexity in attacks. Naturally, there are many other methods to handle multisets for improved integral properties apart from merging. How to adjust a multiset to construct an integral distinguisher on certain cryptographic algorithms will be a worthwhile research question. Figuring out this problem would require a deep understanding of algebraic structures and cryptographic algorithms.

3.4 Combining with monomial detection

Determining the expression of a cryptographic algorithm is not an easy task. Due to its complexity, computing and directly storing the expression can be challenging. Fortunately, it is unnecessary to do so, and confirming the presence of some monomials is sufficient to determine integral distinguishers with our method. Actually, there have been several studies on monomial detection, such as division property, monomial prediction and coefficient grouping. These techniques exhibit certain differences in terms of accuracy and efficiency. Attackers would prefer applying an efficient no-false-alarm monomial detection algorithm rather than an accurate yet inefficient one. In this context, a no-false-alarm algorithm implies that it may mistakenly identify some non-existent monomials as existing, but it will always correctly identify monomials that do exist. Given an E-integral multiset and a cipher Enc, the algorithm to determine whether it forms an integral distinguisher in certain round is as shown in Algorithm 1.

Algorithm 1 A framework for determining integral distinguishers

```
 1: procedure $IDJudicator$( Indicator set $E$, cipher $Enc$, the number of round $R$)
 2:     $flag \leftarrow 0$
 3:     $\bar{E} \leftarrow [0, p^n - 1] \setminus E$
 4:     for $j \in \bar{E}$ do
 5:         $flag \leftarrow MonomialDetect(Enc, R, j)$
 6:         if $flag == 1$ then
 7:             return False
 8:         end if
 9:     end for
10:     return True
11: end procedure
```

The function $MonomialDetect()$ in line 5 is responsible for monomial detection. When the monomial does not exist, it returns 0; otherwise, it returns 1. The monomial detection process is treated as a black box. Therefore, Algorithm 1 is an universal algorithm that can accommodate all monomial detection methods. Cryptanalysts can choose any monomial detection technique they desire based on the target cipher, computational capabilities, and other requirements.

4 Experiments

In Section 3, we introduced the concept of integral multiset and proposed Algorithm 1 for determining an integral distinguisher. These components naturally form a framework for exploring integral distinguishers. In order to demonstrate the effectiveness of integral multiset, we applied our framework to two specific ciphers: MiMC and Chaghri. As a result, we successfully found some better integral distinguishers, exactly, the currently optimal integral distinguishers for the same number of rounds. In this section, we will provide a detailed description of our experiments and corresponding results.

4.1 Application to MiMC

MiMC [AGR+16] is a key-alternating block cipher over a large binary extension field $\mathbb{F}_{2^n}$ or a prime field $\mathbb{F}_p$. In this paper, we will only focus on the binary field version of MiMC. In detail, the encryption function of MiMC is

$$E_k(x) = (F_{r-1} \circ F_{r-2} \circ \cdots \circ F_0)(x) + k,$$

where $x \in \mathbb{F}_{2^n}$ is the plaintext and r is the number of rounds. The round function F_i of MiMC is quite simple, consisting of a cube function $f(x) = x^3$ and an addition. More specifically, each F_i is defined as

$$F_i(x) = (x + k + c_i)^3,$$

where $c_i \in \mathbb{F}_{2^n}$ is the round constant and $c_0 = 0$. All the round constants are chosen randomly at the instantiation of MiMC and then fixed. Figure 1 presents a high-level overview of MiMC. We denote MiMC with exponent d and block size n as $\mathsf{MiMC}_d(n, r)$. To ensure the round function be a permutation, the block size n has to be odd. In the specification of MiMC, the designers allow exponents of the form $2^l + 1$ and $2^l - 1$. The recommended value for block size is 129. According to the requirement of security, the number of round is $r = \lceil n \cdot \log_d 2 \rceil$. In [AGR$^+$16], designers totally give three types of MiMC, namely MiMC-n/n, MiMC-$2n/n$ (Feistel) and MiMC-p/p. We have considered all the three types and provided the details in the following text.

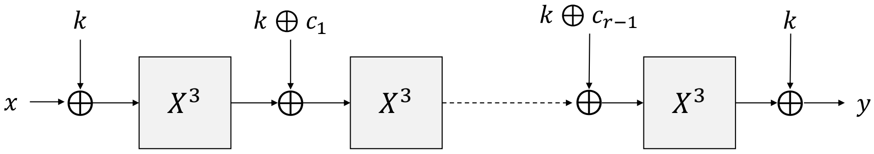

Figure 1: r-round MiMC encryption function.

4.1.1 Application to MiMC-n/n

MiMC-n/n is the most basic block cipher of MiMC, and it is also the most extensively analyzed MiMC block cipher in the literature. Designers of MiMC claim that the large number of rounds ensures that the algebraic degree of MiMC in its native field will be maximum or almost maximum. This naturally thwarts higher-order differential attacks. However, Eichlseder et al. [EGL$^+$20] stated that the algebraic degree of MiMC grow linearly in the number of rounds and not exponentially, and thus the security margin against is only 1 or 2 rounds. Let $\delta^{(d,r)}_{[EGL+20]}$ denote the theoretical upper bound of algebraic degree in [EGL$^+$20]. Then the bound is computed as $\delta^{(d,r)}_{[EGL+20]} = \lfloor \log_2(d^r + 1) \rfloor$. Besides the bound $\delta^{(d,r)}_{[EGL+20]}$, Bouvier et al. [BCP23] proposed another theoretical bound of algebraic degree for the case $d = 2^l + 1$, denoted by $\delta^{(d,r)}_{[BCP23]}$. $\delta^{(d,r)}_{[BCP23]}$ is tighter than $\delta^{(d,r)}_{[EGL+20]}$ and computed as

$$\delta^{(d,r)}_{[BCP23]} = \begin{cases} 2 \cdot \lceil k_r/2 - 1 \rceil, & l = 1, \\ \lfloor r \log_2 d \rfloor - l + 1, & l >= 2, \end{cases}$$

where $k_r = \lfloor r \log_2 d \rfloor$.

The state-of-the-art work on estimating the algebraic degree of MiMC is based on field-based division property [CHWW22]. Cui et al. proposed a method for finding integral properties over binary extension fields $\mathbb{F}_{2^n}$, called general monomial prediction (GMP). They decomposed the cipher into a sequences of simple functions and proposed the corresponding propagation rules of monomials. With the aid of the bit-vector theory of Satisfiability Modulo Theories (SMT), they successfully modeled the propagation of monomials using CVC [BBB$^+$22] input language and took STP [GD07, CGP$^+$08] and

Cryptominisat5 [SNC09] as solvers. In particular, an initial constraint $wt(u^{(0)}) = l$ needs to be added into the SMT model for detecting the monomials with l Hamming weight, where $u^{(0)}$ is the exponent of input variable $x^{(0)}$. A simple framework of their degree estimation method is illustrated in Algorithm 2. The upper bound found by GMP, denoted as $\delta_{GMP}^{(d,r)}$, is the best upper bound so far. Moreover, $\delta_{GMP}^{(d,r)}$ exactly equals to the observed algebraic degree in small-scale instances [CHWW22], which highlights the accuracy of GMP.

Algorithm 2 Degree estimation method in [CHWW22]

```
procedure $DegEst$(The $r$-round SMT model $M_r$, the maximum algebraic degree $\Delta$)
    $i \leftarrow \lfloor \log_2(d^r + 1) \rfloor$
    while $i > 0$ do
        $M \leftarrow M_r$
        $M.addConstr(wt(u^{(0)}) = i)$
        solve $M$
        if $M$ is satisfiable then
            return $i$
        end if
        $i \leftarrow i - 1$
    end while
    return 0
end procedure
```

In this experiment, we employ GMP as our monomial detection method and find the highest round integral distinguishers for each candidate integral multiset. As a result, we successfully find some effective integral distinguishers. First of all, to determine whether a monomial x^j is contained by r-round MiMC, we need to modify the initial constraint of SMT model from $wt(u^{(0)}) = l$ to $u^{(0)} = j$. Given a multiset $\mathbb{X}$ belongs to an E-integral multiset, according to Proposition 1, it is necessary to ensure that none of the monomials x^j, satisfying $j \notin E$, are contained by r-round MiMC in order to form an integral distinguisher. Obviously, it is inefficient to evaluate monomials individually. By observing the integral multiset obtained from Theorem 5, it is evident that we only need to focus on non-zero exponents that are divisible by the size of the multiplicative subgroup, i.e. the set $\mathcal{D}_{|G|}^{2^n} \setminus \{0\}$. Fortunately, SMT supports modular arithmetic constraints. Therefore, by modifying the initial constraints to $u^{(0)} \% |G| = 0$ and $u^{(0)} > 0$, we can determine whether the multiset is an r-round integral distinguisher by solving only one SMT model. The detail of searching algorithm is shown in Algorithm 3.

The function $ModelGen(r, d)$ in line 5 would return the SMT model for MiMC$_d(129, r)$[3]. In Algorithm 3, we start the process of searching from round $br = \lfloor \log_d |G| \rfloor$ instead of round 0. This is because the highest degree monomial that can appear in MiMC$_d(129, r)$ is x^{d^r}, which means that the monomial $x^{c \cdot |G|}, c > 0$, will not appear in the first br rounds. As a result, the input multiset will always form an integral distinguisher within the initial br rounds.

By factoring $2^{129} - 1$, we get an expression $2^{129} - 1 = 7 \cdot 431 \cdot 9719 \cdot 2099863 \cdot 11053036065049294753459639$, which means there are totally 30 meaningful subgroups with different sizes. We firstly applied Algorithm 3 to MiMC$_3(129, r)$, the standard version of MiMC, and found some effective integral distinguishers of different rounds. The result on MiMC$_3(129, r)$ is listed in Table 2. In the subsequent results presentation, we will use **bold font** to highlight significant improvements.

[3] As the process of generating model is not the main focus, it will not be further discussed in this paper. Interested readers can refer to [CHWW22].

Algorithm 3 Search for integral distinguishers using $\mathcal{UD}_{|G|}^{2^n} \cup \{0\}$-integral multiset

```
procedure $IDSearch$(The size of group $|G|$, block size $n$, the exponent of round function $d$)
    $R \leftarrow \lceil n \cdot \log_d 2 \rceil$
    $br \leftarrow \lfloor \log_d |G| \rfloor$
    while $br \leq R$ do
        $M \leftarrow ModelGen(br, d)$
        $M.addConstr(u^{(0)} \% |G| = 0)$
        $M.addConstr(u^{(0)} > 0)$
        solve $M$
        if $M$ is satisfiable then
            return $br - 1$
        end if
        $br \leftarrow br + 1$
    end while
    return $R$
end procedure
```

Table 2: Effective integral distinguishers found on $\mathsf{MiMC}_3(129, r)$.

Round	$\delta_{GMP}^{(d,r)}$	Comp. of HD-ID*	$\|$G$\|^*$	Comp. of IM-ID†
3	**4**	$\mathbf{2^5}$	$\mathbf{p_1}$	$\mathbf{2^3}$
6	8	2^9	p_2	$2^{8.75}$
16	24	2^{25}	$p_1 \cdot p_2 \cdot p_3$	$2^{24.81}$
21	32	2^{33}	$p_1 \cdot p_2 \cdot p_4$	$2^{32.56}$
60	94	2^{95}	$p_1 \cdot p_2 \cdot p_5$	$2^{94.75}$

* The complexity of integral distinguisher based on higher-order derivative is computed as $2^{\delta_{GMP}^{(d,r)}+1}$, where $\delta_{GMP}^{(d,r)}$ is the estimated upper bound of algebraic degree.

* $(p_1, p_2, p_3, p_4, p_5) = (7, 431, 9719, 2099863, 11053036065049294753459639)$.

† According to Theorem 5, a zero element needs to be added. Hence, the complexity of integral distinguisher based on integral multiset is computed as $2^{\log_2(|G|+1)}$.

To verify the results, we implemented the cipher $\mathsf{MiMC}_3(129, r)$ with SageMath [The20] and tested the sum for a large number of randomly generated keys and round constants in small-scale instances. Across different keys and round constants, we consistently obtained a sum of zero, which proves the correctness of our method.

During the process of searching, we totally found 5 integral distinguishers better than the previous ones constructed using higher-order derivative. It is worth noting that GMP is an efficient and accurate method for algebraic degree estimation, and the difference between the bound $\delta_{GMP}^{(d,r)}$ and the actual algebraic degree is small. This statement holds true for at least some verifiable small instances [CHWW22]. Therefore, although the new integral distinguishers found by our method do not exhibit a significant improvement on complexity, such results cannot be obtained solely through the use of algebraic degree. This advantage is well demonstrated in the round 3 result shown in Table 2. By directly computing the expression of output function, it can be observed that the monomial with highest degree is 27 and the set of monomials $\{x^5, x^7, x^{13}, x^{14}, x^{15}, x^{20}, x^{21}, x^{22}, x^{23}\}$ does not appear. A subspace with 2^5 elements is required for a higher-order differential attack because the algebraic degree of output function is 4. However, we can derive a new integral distinguisher based on the multiplicative subgroup of size 7 due to the absence of $\{x^7, x^{14}, x^{21}\}$.

We also applied Algorithm 3 to $\mathsf{MiMC}_d(129, r)$ with exponents $d \in \{5, 7, 9, 15, 17\}$, respectively. Table 3 shows our results on $\mathsf{MiMC}_d(129, r)$ with different exponents. We successfully found some integral distinguishers for all $d \in \{5, 7, 9, 15, 17\}$. Specifically, when $d = 17$, we obtained a 12-round integral distinguisher with a complexity of 2^{43}, which is only a quarter of the complexity of previous one. This result demonstrates that it is possible to enhance the efficiency of integral attacks on complex functions with our method. According to Lucas's Theorem[4], after one round of iteration, a power function with $d = 2^l - 1$ will generate more terms/monomials than the one with $d = 2^l + 1$. This implies that the output of cipher with x^{2^l-1} as round function is more likely to become dense after several rounds, which intuitively helps prevent attacks based on integral multiset. However, Table 3 shows that integral multiset still proves effective when $d = 2^l - 1$. This indicates that the designer cannot resist our attack by simply replacing the round function of the cipher.

Table 3: Effective integral distinguishers found on diffrent versions of $\mathsf{MiMC}_d(129, r)$.

d	Round	$\delta_{GMP}^{(d,r)}$	Comp. of HD-ID	|G|	Comp. of IM-ID
5	**2**	**3**	$\mathbf{2^4}$	$\mathbf{p_1}$	$\mathbf{2^3}$
	41	94	2^{95}	$p_1 \cdot p_2 \cdot p_5$	$2^{94.75}$
	49	112	2^{113}	$p_2 \cdot p_4 \cdot p_5$	$2^{112.95}$
7	9	24	2^{25}	$p_1 \cdot p_2 \cdot p_3$	$2^{24.81}$
	17	45	2^{46}	$p_1 \cdot p_2 \cdot p_3 \cdot p_4$	$2^{45.81}$
9	**2**	**3**	$\mathbf{2^4}$	$\mathbf{p_1}$	$\mathbf{2^3}$
	11	32	2^{33}	$p_1 \cdot p_2 \cdot p_4$	$2^{32.56}$
15	29	112	2^{113}	$p_2 \cdot p_4 \cdot p_5$	$2^{112.95}$
	31	120	2^{121}	$p_1 \cdot p_3 \cdot p_4 \cdot p_5$	$2^{120.25}$
17	**2**	**3**	$\mathbf{2^4}$	$\mathbf{p_1}$	$\mathbf{2^3}$
	12	**44**	$\mathbf{2^{45}}$	$\mathbf{p_2 \cdot p_3 \cdot p_4}$	$\mathbf{2^{43}}$

When searching for integral distinguishers using a given multiplicative subgroup G, we found that some rounds cannot form an integral distinguisher due to the existence of a small number of monomials x^j satisfying $|G| \mid j$. However, according to Proposition 2, if these appearing monomials meet certain conditions, we can obtain a new integral distinguisher by merging the original integral multiset with its cosets. We also applied this method to different version of $\mathsf{MiMC}_d(129, r)$. The results of experiment are shown in Table 4.

4.1.2 Application to MiMC-$2n/n$ (Feistel)

The designer of MiMC also proposed MiMC-$2n/n$, or Feistel MiMC, whose non-linear permutation is the same as MiMC. We denote Feistel MiMC with exponent d and block size n as $\mathsf{FeistelMiMC}_d(n, r)$. The i-th round function of $\mathsf{FeistelMiMC}_d(n, r)$ is shown in Figure 2 and defined as

$$(x_0^{(i+1)}, x_1^{(i+1)}) \leftarrow (x_1^{(i)} \oplus (x_0^{(i)} \oplus k \oplus c_i)^d, x_0^{(i)}).$$

The round constants c_i are again chosen randomly except for the first and last round

[4] By Lucas's Theorem, $\binom{n}{m} = \prod_{i=0}^{k} \binom{n_i}{m_i} (\bmod 2)$, where $n = \sum_{i=0}^{k} n_i \cdot 2^i$, $m = \sum_{i=0}^{k} m_i \cdot 2^i$ and $n_i, m_i \in \{0, 1\}$.

Table 4: Effective integral distinguishers found on diffrent versions of $\mathsf{MiMC}_d(129, r)$ by merging integral multiset.

d	Round	$\delta_{GMP}^{(d,r)}$	Comp. of HD-ID	$\lvert G\rvert$	Monmials[⋆]	Comp. of IM-ID[*]
3	7	10	2^{11}	p_2	$\{x^{2\lvert G\rvert}, x^{4\lvert G\rvert}\}$	$2^{10.33}$
	64	100	2^{101}	$p_1 \cdot p_3 \cdot p_5$		$2^{100.83}$
5	20	44	2^{45}	$p_2 \cdot p_3 \cdot p_4$	$\{x^{4\lvert G\rvert}, x^{8\lvert G\rvert}\}$	$2^{44.59}$
15	1	4	2^{5}	p_1	$\{x^{\lvert G\rvert}, x^{2\lvert G\rvert}\}$	$2^{4.46}$
	12	44	2^{33}	$p_2 \cdot p_3 \cdot p_4$	$\{x^{2\lvert G\rvert}, x^{4\lvert G\rvert}, x^{8\lvert G\rvert}\}$	$2^{44.59}$

[⋆] The set of monomials $\{x^j : |G| \mid j\}$ that exists in $\mathsf{MiMC}_d(129, r)$.

[*] According to Proposition 2, the size of new merged multiset is $3 \cdot (|G| + 1)$. Due to the binary extension field, we can reduce the number of 0 from 3 to 1. Thus, the complexity of integral distinguisher based on merged multiset is computed as $3|G| + 1$.

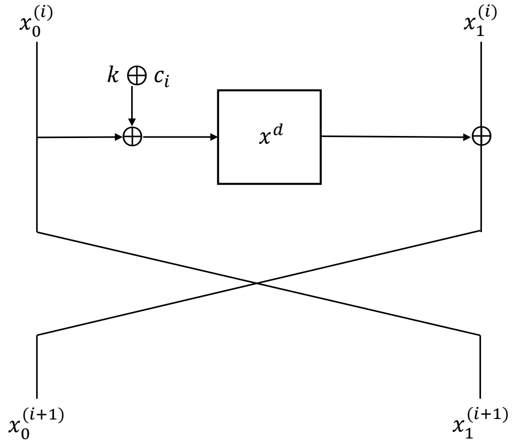

Figure 2: The i-th round function of FeistelMiMC$_d(n, r)$

constants which are equal to 0. The number of rounds for $\mathsf{FeistelMiMC}_d(n, r)$ is $r' = 2 \cdot r = 2 \cdot \lceil n \cdot \log_d 2 \rceil$.

We focus on the univariate case of Feistel MiMC. In this case, we only consider the input and output of left branch. Because of the structure of Feistel network, all integral distinguishers found in this way can be naturally extended with 2 more rounds. Without a doubt, when considering the same nonlinear function and number of rounds, the output of left branch of Feistel MiMC is denser than that of MiMC. New variable x_1 and addition of both branches decrease the probability of two identical monomials canceling each other out. Taking the example of three rounds, the set of missing monomials for MiMC is $\{x^5, x^7, x^{13}, x^{14}, x^{15}, x^{20}, x^{21}, x^{22}, x^{23}\}$, while for Feistel MiMC it is $\{x_0^{14}, x_0^{15}, x_0^{20}, x_0^{21}, x_0^{22}, x_0^{23}\}$. However, when we applied Algorithm 3 to the standard version $\mathsf{FeistelMiMC}_3(n, r)$, we found that the Feistel network did not prevent us from finding better integral distinguishers. The result is shown in Table 5.

Table 5: Effective integral distinguishers found on $\mathsf{FeistelMiMC}_3(129, r)$.

Round	$\delta_{x_0}(x_0^{(r)})^\star$	Comp. of HD-ID	|G|	Comp. of IM-ID
7	11	2^{12}	$p_1 \cdot p_2$	$2^{11.56}$
15	23	2^{24}	$p_1 \cdot p_4$	$2^{23.81}$
58	91	2^{92}	$p_2 \cdot p_5$	$2^{91.94}$
71	112	2^{113}	$p_2 \cdot p_4 \cdot p_5$	$2^{112.95}$
73	115	2^{116}	$p_1 \cdot p_2 \cdot p_4 \cdot p_5$	$2^{115.75}$
74	117	2^{118}	$p_3 \cdot p_4 \cdot p_5$	$2^{117.74}$

* We denote the algebraic degree of left branch over $x_0^{(0)}$ as $\delta_{x_0}(x_0^{(r)})$.

4.1.3 Applications to MiMC-p/p

A variant of MiMC over prime finite fields is denoted as MiMC-p/p. MiMCHash over prime finite field is one of the candidate hash function in STARKWARE hash challenge. To ensure the round function as a permutation, the prime number p should satisfy $\gcd(3, p-1) = 1$. When MiMC-p/p is used to construct hash function, the prime number p should also satisfy $p - 1 = 2^l \cdot q$.

When we attempted to directly apply our framework for seeking integral distinguishers on MiMC-p/p, we encountered two problems. The first problem is the lack of monomial detection over $\mathbb{F}_p$. Though GMP can work perfectly on MiMC-n/n, the propogation of monomial over $\mathbb{F}_p$ is greatly different from that over $\mathbb{F}_{2^n}$, which limits the application of GMP on MiMC-p/p. The second problem is more essential. During our previous process of searching, we make use of the absence of certain monomials, e.g. though $d(F) > |G|$, monomials $x^{c\cdot|G|}$ do not appear. For a ciphers over $\mathbb{F}$, where p is a large prime, the corresponding function would be very dense. This property greatly restricts the existence of lower complexity integral distinguishers based on multiplicative subgroups. However, we can still construct distinguishers based on multiplicative subgroups if $d(F) < |G|$. Beyne et. al. [BCD+20] give the concrete proposition.

Proposition 3. *Let G be a multiplicative subgroup of $\mathbb{F}_p^*$, where p is a prime. For any $F : \mathbb{F}_p \to \mathbb{F}_p$ such that $d(F) < |G|$, we have*

$$\sum_{x\in G} F(x) = F(0) \cdot |G|.$$

It is worth noting that, besides Proposition 3, we are still able to find integral distinguishers with lower complexity by using the framework of integral multiset. This can be achieved by our merging method or a new class of integral multisets. We will provide an example of merging multisets to find better integral distinguishers. As for how to construct a new class of integral multisets, it is a topic that requires further research and investigation.

Example 2. Consider MiMC-p/p with $p = 2^{161} + 23 \cdot 2^{128} + 1$, we can factor $p - 1$ as $2^{128} \cdot 5 \cdot 31 \cdot 37 \cdot 1497809$. For 80-round MiMC-$p/p$, we can use a multiplicative subgroup G_1 with size 2^{128}, since $3^{80} < 2^{128}$. When the round number of cipher comes to 81, according to Proposition 3, the smallest multiplicative subgroup we can use is the group G_2 with size $5 \cdot 2^{128}$.

By investigating the polynomial of 81-round MiMC-p/p, we know that among monomials $x^{c\cdot|G_1|}$ only $x^{|G_1|}$ would appear. According to Proposition 2, we can construct an integral distinguisher based on a multiset with size $3 \cdot |G_1| = 3 \cdot 2^{128}$.

4.2 Applications to Chaghri

Chaghri [AMT22] is a FHE-friendly block cipher defined over a large finite field. The block size of Chaghri is 189. We denote the state by $\boldsymbol{a} = (a_1, a_2, a_3) \in \mathbb{F}_{2^{63}}^3$. The number of round is 8 and each round is composed of two steps. The step function $S(\boldsymbol{a})$ of Chaghri[5] is shown in Figure 3 and defined as

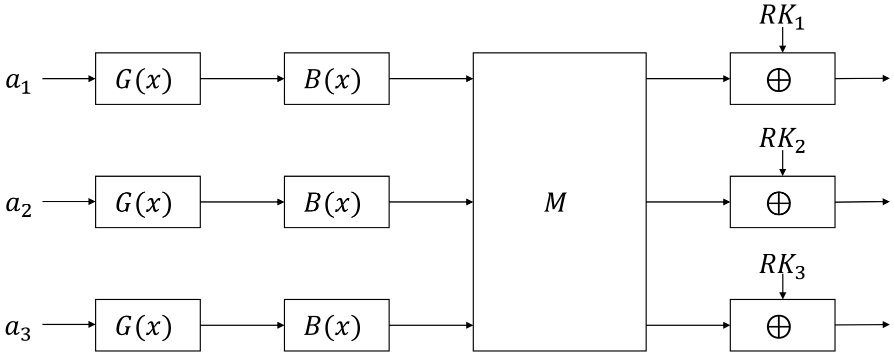

Figure 3: The step function of Chaghri.

$$S(\boldsymbol{a}) = M \cdot B(G(\boldsymbol{a}))^T + \boldsymbol{RK},$$

where $\boldsymbol{RK} = (RK_1, RK_2, RK_3) \in \mathbb{F}_{2^{63}}^3$ is the round key. The components used in step function are the nonlinear function $G(\boldsymbol{a}) = (a_1^{2^{32}+1}, a_2^{2^{32}+1}, a_3^{2^{32}+1})$, the affine transform[6] $B(\boldsymbol{a}) = (c_1 a_1^{2^3} + c_2, c_1 a_2^{2^3} + c_2, c_1 a_3^{2^3} + c_2)$ and a 3×3 Maximum Distance Separable (MDS) matrix M.

Coefficient grouping technique [LAW+23, LGB+23], proposed and refined by Liu et al., is a novel method for monomial detection. It has been successfully employed to construct integral distinguishers targeting Chaghri. The main idea of coefficient grouping is to give an expression of all possible monomials and find the maximum Hamming weight among the exponents. To be specific, by setting the input state as

$$a_1^{(0)} = A_{0,1}X + B_{0,1}, a_2^{(0)} = A_{0,2}X + B_{0,2}, a_3^{(0)} = A_{0,3}X + B_{0,3},$$

where $X \in \mathbb{F}_{2^{63}}$ is the variable and $A_{0,i}, B_{0,i} \in \mathbb{F}_{2^{63}} (1 \leq i \leq 3)$ are randomly chosen constants, the set of possible exponents in r-th round is

$$w_r = \{e : e = \mathcal{M}_n(\sum_{i=0}^{n-1} 2^i \gamma_i), 0 \leq \gamma_i \leq N_i^r\}. \tag{2}$$

$\mathcal{M}_n(x)$ is a function defined as follows:

$$\mathcal{M}_n(x) = \begin{cases} 2^n - 1, & (2^n - 1) \nmid x \text{ and } x \neq 0, \\ x\%(2^n - 1), & \text{otherwise.} \end{cases}$$

In Equation (2), the vector of integers $(N_{n-1}^r, N_{n-2}^r, \ldots, N_0^r)$ can be computed by the recursive relation

$$N_i^{j+1} = N_{(i-35)\%n}^j + N_{(i-32)\%n}^j, 0 \leq i \leq n-1, j \geq 0$$

[5] In this paper, we consider the decryption of Chaghri, since the secure number of rounds mainly depends on the security of decryption and low multiplicative depth in decryption desired in FHE schemes.

[6] $c_1, c_2 \in \mathbb{F}_{2^{63}}$ are constants. In fact, after the proposal of coefficient grouping [LAW+23], the affine transform has been replaced by a denser one. However, we focus on the original version in order to make a better comparison with the result in [LAW+23].

and the initial value is

$$N_0^0 = 1, N_i^0 = 0, 1 \leq i \leq n-1.$$

With the value of $(N_{n-1}^r, N_{n-2}^r, \ldots, N_0^r)$, the algebraic degree can be obtain by finding the maximum Hamming weight. In [LAW$^+$23], the authors encoded the problem as an MILP (Mixed Integer Linear Programming) problem and presented the methodology for modeling arithmetic addition and comparison as MILP constraints.

It should be noted that only when $j \in w_r$ will the monomial x^j appear in the r-round Chaghri. The detection of certain monomial x^j naturally becomes a constraint problem, as shown below:

$$\begin{aligned} &\mathcal{M}_n(\sum_{i=0}^{n-1} 2^i \gamma_i) = j, \\ &0 \leq \gamma_i \leq N_i^r, i = 0, 1, \ldots, n-1. \end{aligned} \tag{3}$$

In order to detect the existence of monomials with non-zero exponents that are divisible by the size of a certain multiplicative subgroup G, we need to replace $\mathcal{M}_n(\sum_{i=0}^{n-1} 2^i \gamma_i) = j$ with $\mathcal{M}_n(\sum_{i=0}^{n-1} 2^i \gamma_i) \% |G| = 0$ and $\mathcal{M}_n(\sum_{i=0}^{n-1} 2^i \gamma_i) > 0$. However, the presence of $\mathcal{M}_n(\sum_{i=0}^{n-1} 2^i \gamma_i) \% |G| = 0$ prevents us from encoding the problem as an MILP problem as [LAW$^+$23]. After carefully examining the operations and variable domains in the constraint problem, we have realized that SMT, supporting bit vectors and arithmetic operations, is a better choice. Therefore, we model the problem with CVC input language and solve it using STP as the solver. In particular, an algorithm for exploring the set of monomials $\{x^j : |G| \mid j, j > 0\}$ is shown in Algorithm 4.

Algorithm 4 Explore the set of monomials $\{x^j : |G| \mid j, j > 0\}$ with coefficient grouping

```
1:  procedure MonSearch(The size of group $|G|$, the $r$-th round coefficient limit $(N_{n-1}^r, N_{n-2}^r, \ldots, N_0^r)$)
2:      $M \leftarrow$ an empty model
3:      $R \leftarrow \emptyset$
4:      for $i$ from 0 to $n-1$ do
5:          $M.addVar(a_i)$
6:          $M.addConstr(a_i \geq 0)$
7:          $M.addConstr(a_i \leq N_i^r)$
8:      end for
9:      $M.addVar(s)$
10:     $M.addConstr(s = \sum_{i=0}^{n-1} 2^i a_i)$
11:     $M.addConstr(s > 0)$
12:     $M.addConstr(s \% |G| = 0)$
13:     solve $M$
14:     while $M$ is satisfiable do
15:         $t \leftarrow M.getVal(s)$
16:         $R \leftarrow R \cup \{t\}$
17:         $M.addConstr(s \neq t)$
18:         solve $M$
19:     end while
20:     return $R$
21: end procedure
```

Algorithm 4 includes two phases: constructing the basic model and finding all solutions. Lines 2-12 depict the process of constructing SMT model with the limit of r-round coefficient $(N_{n-1}^r, N_{n-2}^r, \ldots, N_0^r)$. On the other hand, lines 13-19 iteratively exclude feasible solutions until the model becomes unsatisfiable. If the output of Algorithm 4 is an empty set, an integral distinguisher can be obtained based on Proposition 1 and Theorem 5. Otherwise, the sum over $|G| \cup \{0\}$ would not be equal to 0. However, if the output satisfies certain conditions, it is still possible to form an integral distinguisher by merging integral multiset, as described in Proposition 2.

With the above algorithm, it is easy to search for integral distinguishers for multiplicative subgroup with any size. By factoring $2^{63}-1$, an expression $2^{63}-1 = 7 \cdot 7 \cdot 73 \cdot 127 \cdot 337 \cdot 92737 \cdot 649657$ is obtained. We denote the algebraic degree upper bound of r-step Chaghri as δ_r. Applying Algorithm 4 to Chaghri, some integral distinguishers which are more effective than previous ones in [LAW⁺23]. The result is listed in Table 6.

Table 6: Effective integral distinguishers found on Chaghri

Step	δ_r	Comp. of HD-ID	$\|G\|^\star$	Comp. of IM-ID
8	17	2^{18}	$q_1 \cdot q_2 \cdot q_4$	$2^{17.39}$
15	35	2^{36}	$q_5 \cdot q_6$	$2^{35.81}$
17	40	2^{41}	$q_2 \cdot q_3 \cdot q_4 \cdot q_6$	$2^{40.88}$
17	40	2^{41}	$q_1 \cdot q_2 \cdot q_3 \cdot q_4 \cdot q_5$	$2^{40.88}$
24	58	2^{59}	$q_2 \cdot q_3 \cdot q_4 \cdot q_5 \cdot q_6$	$2^{58.97*}$

* $(q_1, q_2, q_3, q_4, q_5, q_6) = (7, 73, 127, 337, 92737, 649657)$.

* This distinguisher is constructed by merging integral multiset. The monomial with non-zero exponent that is divisible by $|G|$ found in 24-th step is $x^{2|G|}$.

With integral distinguishers in Table 6, we can mount the integral attack on different steps of Chaghri with lower data and time complexity. Furthermore, according to [LAW⁺23], one more step can be easily extended by setting the input carefully.

5 Conclusion

In the past two decades, bit-based block ciphers or ones with efficient bit-based implementation have played a dominant role in symmetric cryptographic primitives. Consequently, when we refer to integral attacks, more often than not, we are referring to higher-order differential attacks. Recently, there has been a surge of MPC/FHE/ZK-friendly symmetric primitives. These primitives are typically defined over large finite fields such $\mathbb{F}_{2^n}$ and $\mathbb{F}_p$. When dealing with these primitives, using higher-order derivative to construct integral distinguishers is no longer the most natural choice. In fact, it would even fail to be effective over finite field with odd characteristic. How to effectively perform integral attacks over large finite fields has become a subject worthy of research. While there have been scattered studies in the past, there is still a lack of research on the framework of integral attacks over large finite fields.

In this paper, we introduce a novel concept called *integral multiset*, which provides a clear and unified characterization of the integral property of multisets over the finite field $\mathbb{F}_{p^n}$. Previous higher-order differential distinguishers can be regarded as a class of integral multiset based on vector subspaces. In particular, we present a new class of integral multisets based on multiplicative subgroups and a method about how to merge existing integral multisets to create one with better integral property. Combining with monomial detection techniques, we propose a framework for searching for integral distinguishers and apply it to MiMC and Chaghri. For the ciphers defined over $\mathbb{F}_{2^n}$, we find some integral distinguishers with lower time and data complexity, which were not achievable by higher-order differential attacks. For the ciphers defined over $\mathbb{F}_p$, we also demonstrated that the merging method indeed enables us to find better distinguishers. Furthermore, our framework can perfectly adapt to various monomial detection techniques. From this fact, we expect that the integral multiset will be of great help in improving integral attacks. In order to resist integral attacks based on multiplicative subgroups, the designer of cipher should choose a field with as few multiplicative subgroups as possible.

Finally, there are still some open questions that require further research. For example, how to construct E-integral multisets for any indicated set E and whether there exists a theoretical lower bound on the cardinality of E-integral multisets for certain E.

References

[AD18] Tomer Ashur and Siemen Dhooghe. MARVELlous: a STARK-friendly family of cryptographic primitives. Cryptology ePrint Archive, Report 2018/1098, 2018. https://eprint.iacr.org/2018/1098.

[AGP+19] Martin R. Albrecht, Lorenzo Grassi, Léo Perrin, Sebastian Ramacher, Christian Rechberger, Dragos Rotaru, Arnab Roy, and Markus Schofnegger. Feistel structures for MPC, and more. In Kazue Sako, Steve Schneider, and Peter Y. A. Ryan, editors, *ESORICS 2019, Part II*, volume 11736 of *LNCS*, pages 151–171. Springer, Heidelberg, September 2019.

[AGR+16] Martin R. Albrecht, Lorenzo Grassi, Christian Rechberger, Arnab Roy, and Tyge Tiessen. MiMC: Efficient encryption and cryptographic hashing with minimal multiplicative complexity. In Jung Hee Cheon and Tsuyoshi Takagi, editors, *ASIACRYPT 2016, Part I*, volume 10031 of *LNCS*, pages 191–219. Springer, Heidelberg, December 2016.

[AMT22] Tomer Ashur, Mohammad Mahzoun, and Dilara Toprakhisar. Chaghri - A fhe-friendly block cipher. In Heng Yin, Angelos Stavrou, Cas Cremers, and Elaine Shi, editors, *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security, CCS 2022, Los Angeles, CA, USA, November 7-11, 2022*, pages 139–150. ACM, 2022.

[ARS+15] Martin R. Albrecht, Christian Rechberger, Thomas Schneider, Tyge Tiessen, and Michael Zohner. Ciphers for MPC and FHE. In Elisabeth Oswald and Marc Fischlin, editors, *EUROCRYPT 2015, Part I*, volume 9056 of *LNCS*, pages 430–454. Springer, Heidelberg, April 2015.

[BBB+22] Haniel Barbosa, Clark W. Barrett, Martin Brain, Gereon Kremer, Hanna Lachnitt, Makai Mann, Abdalrhman Mohamed, Mudathir Mohamed, Aina Niemetz, Andres Nötzli, Alex Ozdemir, Mathias Preiner, Andrew Reynolds, Ying Sheng, Cesare Tinelli, and Yoni Zohar. cvc5: A versatile and industrial-strength SMT solver. In Dana Fisman and Grigore Rosu, editors, *Tools and Algorithms for the Construction and Analysis of Systems - 28th International Conference, TACAS 2022, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2022, Munich, Germany, April 2-7, 2022, Proceedings, Part I*, volume 13243 of *Lecture Notes in Computer Science*, pages 415–442. Springer, 2022.

[BC11] Christina Boura and Anne Canteaut. On the influence of the algebraic degree of f^{-1} on the algebraic degree of G circ F. *IACR Cryptol. ePrint Arch.*, page 503, 2011.

[BCD11] Christina Boura, Anne Canteaut, and Christophe De Cannière. Higher-order differential properties of Keccak and Luffa. In Antoine Joux, editor, *FSE 2011*, volume 6733 of *LNCS*, pages 252–269. Springer, Heidelberg, February 2011.

[BCD$^+$20] Tim Beyne, Anne Canteaut, Itai Dinur, Maria Eichlseder, Gregor Leander, Gaëtan Leurent, María Naya-Plasencia, Léo Perrin, Yu Sasaki, Yosuke Todo, and Friedrich Wiemer. Out of oddity - new cryptanalytic techniques against symmetric primitives optimized for integrity proof systems. In Daniele Micciancio and Thomas Ristenpart, editors, *CRYPTO 2020, Part III*, volume 12172 of *LNCS*, pages 299–328. Springer, Heidelberg, August 2020.

[BCP23] Clémence Bouvier, Anne Canteaut, and Léo Perrin. On the algebraic degree of iterated power functions. *Des. Codes Cryptogr.*, 91(3):997–1033, 2023.

[BS91] Eli Biham and Adi Shamir. Differential cryptanalysis of DES-like cryptosystems. In Alfred J. Menezes and Scott A. Vanstone, editors, *CRYPTO'90*, volume 537 of *LNCS*, pages 2–21. Springer, Heidelberg, August 1991.

[CGG$^+$22] Carlos Cid, Lorenzo Grassi, Aldo Gunsing, Reinhard Lüftenegger, Christian Rechberger, and Markus Schofnegger. Influence of the linear layer on the algebraic degree in SP-networks. *IACR Trans. Symm. Cryptol.*, 2022(1):110–137, 2022.

[CGP$^+$08] Cristian Cadar, Vijay Ganesh, Peter M. Pawlowski, David L. Dill, and Dawson R. Engler. EXE: automatically generating inputs of death. *ACM Trans. Inf. Syst. Secur.*, 12(2):10:1–10:38, 2008.

[CHWW22] Jiamin Cui, Kai Hu, Meiqin Wang, and Puwen Wei. On the field-based division property: Applications to MiMC, feistel MiMC and GMiMC. In Shweta Agrawal and Dongdai Lin, editors, *ASIACRYPT 2022, Part III*, volume 13793 of *LNCS*, pages 241–270. Springer, Heidelberg, December 2022.

[CV02] Anne Canteaut and Marion Videau. Degree of composition of highly nonlinear functions and applications to higher order differential cryptanalysis. In Lars R. Knudsen, editor, *EUROCRYPT 2002*, volume 2332 of *LNCS*, pages 518–533. Springer, Heidelberg, April / May 2002.

[DEG$^+$18] Christoph Dobraunig, Maria Eichlseder, Lorenzo Grassi, Virginie Lallemand, Gregor Leander, Eik List, Florian Mendel, and Christian Rechberger. Rasta: A cipher with low ANDdepth and few ANDs per bit. In Hovav Shacham and Alexandra Boldyreva, editors, *CRYPTO 2018, Part I*, volume 10991 of *LNCS*, pages 662–692. Springer, Heidelberg, August 2018.

[DGGK21] Christoph Dobraunig, Lorenzo Grassi, Anna Guinet, and Daniël Kuijsters. Ciminion: Symmetric encryption based on Toffoli-gates over large finite fields. In Anne Canteaut and François-Xavier Standaert, editors, *EUROCRYPT 2021, Part II*, volume 12697 of *LNCS*, pages 3–34. Springer, Heidelberg, October 2021.

[DGH$^+$21] Christoph Dobraunig, Lorenzo Grassi, Lukas Helminger, Christian Rechberger, Markus Schofnegger, and Roman Walch. Pasta: A case for hybrid homomorphic encryption. Cryptology ePrint Archive, Report 2021/731, 2021. https://eprint.iacr.org/2021/731.

[DKR97] Joan Daemen, Lars R. Knudsen, and Vincent Rijmen. The block cipher Square. In Eli Biham, editor, *FSE'97*, volume 1267 of *LNCS*, pages 149–165. Springer, Heidelberg, January 1997.

[DKR$^+$22] Christoph Dobraunig, Daniel Kales, Christian Rechberger, Markus Schofnegger, and Greg Zaverucha. Shorter signatures based on tailor-made minimalist symmetric-key crypto. In Heng Yin, Angelos Stavrou, Cas Cremers, and

Elaine Shi, editors, *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security, CCS 2022, Los Angeles, CA, USA, November 7-11, 2022*, pages 843–857. ACM, 2022.

[EGL+20] Maria Eichlseder, Lorenzo Grassi, Reinhard Lüftenegger, Morten Øygarden, Christian Rechberger, Markus Schofnegger, and Qingju Wang. An algebraic attack on ciphers with low-degree round functions: Application to full MiMC. In Shiho Moriai and Huaxiong Wang, editors, *ASIACRYPT 2020, Part I*, volume 12491 of *LNCS*, pages 477–506. Springer, Heidelberg, December 2020.

[FKL+01] Niels Ferguson, John Kelsey, Stefan Lucks, Bruce Schneier, Michael Stay, David Wagner, and Doug Whiting. Improved cryptanalysis of Rijndael. In Bruce Schneier, editor, *FSE 2000*, volume 1978 of *LNCS*, pages 213–230. Springer, Heidelberg, April 2001.

[GD07] Vijay Ganesh and David L. Dill. A decision procedure for bit-vectors and arrays. In Werner Damm and Holger Hermanns, editors, *Computer Aided Verification, 19th International Conference, CAV 2007, Berlin, Germany, July 3-7, 2007, Proceedings*, volume 4590 of *Lecture Notes in Computer Science*, pages 519–531. Springer, 2007.

[GHR+23] Lorenzo Grassi, Yonglin Hao, Christian Rechberger, Markus Schofnegger, Roman Walch, and Qingju Wang. Horst meets fluid-spn: Griffin for zero-knowledge applications. In Helena Handschuh and Anna Lysyanskaya, editors, *Advances in Cryptology - CRYPTO 2023 - 43rd Annual International Cryptology Conference, CRYPTO 2023, Santa Barbara, CA, USA, August 20-24, 2023, Proceedings, Part III*, volume 14083 of *Lecture Notes in Computer Science*, pages 573–606. Springer, 2023.

[GKR+21] Lorenzo Grassi, Dmitry Khovratovich, Christian Rechberger, Arnab Roy, and Markus Schofnegger. Poseidon: A new hash function for zero-knowledge proof systems. In Michael Bailey and Rachel Greenstadt, editors, *30th USENIX Security Symposium, USENIX Security 2021, August 11-13, 2021*, pages 519–535. USENIX Association, 2021.

[GLR+20] Lorenzo Grassi, Reinhard Lüftenegger, Christian Rechberger, Dragos Rotaru, and Markus Schofnegger. On a generalization of substitution-permutation networks: The HADES design strategy. In Anne Canteaut and Yuval Ishai, editors, *EUROCRYPT 2020, Part II*, volume 12106 of *LNCS*, pages 674–704. Springer, Heidelberg, May 2020.

[GRR+16] Lorenzo Grassi, Christian Rechberger, Dragos Rotaru, Peter Scholl, and Nigel P. Smart. Mpc-friendly symmetric key primitives. In Edgar R. Weippl, Stefan Katzenbeisser, Christopher Kruegel, Andrew C. Myers, and Shai Halevi, editors, *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, October 24-28, 2016*, pages 430–443. ACM, 2016.

[HLM+20] Yonglin Hao, Gregor Leander, Willi Meier, Yosuke Todo, and Qingju Wang. Modeling for three-subset division property without unknown subset - improved cube attacks against Trivium and Grain-128AEAD. In Anne Canteaut and Yuval Ishai, editors, *EUROCRYPT 2020, Part I*, volume 12105 of *LNCS*, pages 466–495. Springer, Heidelberg, May 2020.

[HSWW20] Kai Hu, Siwei Sun, Meiqin Wang, and Qingju Wang. An algebraic formulation of the division property: Revisiting degree evaluations, cube attacks, and key-independent sums. In Shiho Moriai and Huaxiong Wang, editors, *ASIACRYPT 2020, Part I*, volume 12491 of *LNCS*, pages 446–476. Springer, Heidelberg, December 2020.

[KW02] Lars R. Knudsen and David Wagner. Integral cryptanalysis. In Joan Daemen and Vincent Rijmen, editors, *FSE 2002*, volume 2365 of *LNCS*, pages 112–127. Springer, Heidelberg, February 2002.

[Lai94] Xuejia Lai. Higher order derivatives and differential cryptanalysis. *Communications and Cryptography: Two Sides of One Tapestry*, pages 227–233, 1994.

[LAW$^+$23] Fukang Liu, Ravi Anand, Libo Wang, Willi Meier, and Takanori Isobe. Coefficient grouping: Breaking chaghri and more. In Carmit Hazay and Martijn Stam, editors, *EUROCRYPT 2023, Part IV*, volume 14007 of *LNCS*, pages 287–317. Springer, Heidelberg, April 2023.

[LGB$^+$23] Fukang Liu, Lorenzo Grassi, Clémence Bouvier, Willi Meier, and Takanori Isobe. Coefficient grouping for complex affine layers. In Helena Handschuh and Anna Lysyanskaya, editors, *Advances in Cryptology - CRYPTO 2023 - 43rd Annual International Cryptology Conference, CRYPTO 2023, Santa Barbara, CA, USA, August 20-24, 2023, Proceedings, Part III*, volume 14083 of *Lecture Notes in Computer Science*, pages 540–572. Springer, 2023.

[Mat94] Mitsuru Matsui. Linear cryptanalysis method for DES cipher. In Tor Helleseth, editor, *EUROCRYPT'93*, volume 765 of *LNCS*, pages 386–397. Springer, Heidelberg, May 1994.

[RSS17] Dragos Rotaru, Nigel P. Smart, and Martijn Stam. Modes of operation suitable for computing on encrypted data. *IACR Trans. Symm. Cryptol.*, 2017(3):294–324, 2017.

[SNC09] Mate Soos, Karsten Nohl, and Claude Castelluccia. Extending SAT solvers to cryptographic problems. In Oliver Kullmann, editor, *Theory and Applications of Satisfiability Testing - SAT 2009, 12th International Conference, SAT 2009, Swansea, UK, June 30 - July 3, 2009. Proceedings*, volume 5584 of *Lecture Notes in Computer Science*, pages 244–257. Springer, 2009.

[SW13] Yu Sasaki and Lei Wang. Meet-in-the-middle technique for integral attacks against Feistel ciphers. In Lars R. Knudsen and Huapeng Wu, editors, *SAC 2012*, volume 7707 of *LNCS*, pages 234–251. Springer, Heidelberg, August 2013.

[TA14] Yosuke Todo and Kazumaro Aoki. FFT key recovery for integral attack. In Dimitris Gritzalis, Aggelos Kiayias, and Ioannis G. Askoxylakis, editors, *CANS 14*, volume 8813 of *LNCS*, pages 64–81. Springer, Heidelberg, October 2014.

[The20] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 9.1)*, 2020. `https://www.sagemath.org`.

[TIHM17] Yosuke Todo, Takanori Isobe, Yonglin Hao, and Willi Meier. Cube attacks on non-blackbox polynomials based on division property. In Jonathan Katz and Hovav Shacham, editors, *CRYPTO 2017, Part III*, volume 10403 of *LNCS*, pages 250–279. Springer, Heidelberg, August 2017.

[TM16] Yosuke Todo and Masakatu Morii. Bit-based division property and application to simon family. In Thomas Peyrin, editor, *FSE 2016*, volume 9783 of *LNCS*, pages 357–377. Springer, Heidelberg, March 2016.

[Tod15] Yosuke Todo. Structural evaluation by generalized integral property. In Elisabeth Oswald and Marc Fischlin, editors, *EUROCRYPT 2015, Part I*, volume 9056 of *LNCS*, pages 287–314. Springer, Heidelberg, April 2015.

[XZC$^+$22] Tiancheng Xie, Jiaheng Zhang, Zerui Cheng, Fan Zhang, Yupeng Zhang, Yongzheng Jia, Dan Boneh, and Dawn Song. zkbridge: Trustless cross-chain bridges made practical. In Heng Yin, Angelos Stavrou, Cas Cremers, and Elaine Shi, editors, *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security, CCS 2022, Los Angeles, CA, USA, November 7-11, 2022*, pages 3003–3017. ACM, 2022.