Mathematical Aspects of Division Property

Phil Hebborn[1], Gregor Leander[1] and Aleksei Udovenko[2,3]

[1]Ruhr-Universität Bochum.
[2]SnT, University of Luxembourg.
[3]CryptoExperts.

Abstract

This work surveys mathematical aspects of *division property*, which is a state-of-the-art technique in cryptanalysis of symmetric-key algorithms, such as authenticated encryption, block ciphers and stream ciphers. It aims to find integral distinguishers and cube attacks, which exploit weaknesses in the algebraic normal forms of the output coordinates of the involved vectorial Boolean functions. Division property can also be used to provide arguments for security of primitives against these attacks. The focus of this work is a formal presentation of the theory behind the division property, including rigorous proofs, which were often omitted in the existing literature. This survey covers the two major variants of division property, namely *conventional* and *perfect* division property. In addition, we explore relationships of the technique with classic degree bounds.

Keywords: symmetric cryptography, Boolean functions, algebraic degree, integral cryptanalysis, division property

Acknowledgements

The work was supported by the Luxembourg National Research Fund's (FNR) and the German Research Foundation's (DFG) joint project APLICA (C19/IS/13641232).

1 Introduction

In this paper we discuss the mathematical aspects of a modern technique in symmetric cryptography. This method allows to both find better attacks as well as to give stronger arguments for the security of given schemes. A large part of the research in the area of division property is devoted to making the actual computation more efficient and applicable to a larger set of primitives. In a nutshell, this part involves setting up a suitable set of equations and inequalities and ask a modern SAT or MILP solver to find solutions. While very important for the field, this aspect is not in the scope of this survey. Instead we focus on the mathematical aspects of division property. But before doing so, we start by giving a bit of context and motivation.

1.1 Symmetric Cryptography

Symmetric primitives play an important role in our daily communication and protect almost all sensitive communication. The security of a symmetric primitive, be it a block cipher, a stream cipher or a permutation-based construction, can not be proven as such by current techniques. Instead, arguments for its security are always arguments why a specific attack or a class of attacks is not applicable to the given scheme. One important class of attacks are these which exploit properties of the algebraic normal form of the cipher in question. In its general form, a symmetric primitive can always be thought of as a function

$$F\colon \mathbb{F}_2^n \times \mathbb{F}_2^\ell \to \mathbb{F}_2^m$$

that takes two bit-strings x and k of lengths n and ℓ as inputs and produces an m-bit output. The input x can be thought of as public information. It could be the message in case of a block cipher, a message and a tweak in case of a tweakable block cipher, or the initial value (IV) in case of a stream cipher. The input k is the secret key which is, as well as (an implementation of) the function F, shared between sender and receiver. The output $F(x,k)$ can again be thought of as public and would correspond to the cipher-text in case of a block cipher, the key-stream in case of a stream cipher. It could also be an authentication tag in case of a message authentication code (MAC), or serve as both the cipher-text together with an authentication tag in case of authenticated encryption.

The fact that both the input x and the output $F(x,k)$ are thought of as public information might seem unnatural as making both the message and the cipher text public seems to nullify the whole point of encryption. However, it is common to take a very powerful attacker into consideration that actually might have access to many, even chosen, plain-text and cipher-text pairs.

1.2 Attacks Based on the ANF

As any Boolean function, each coordinate $F_t : \mathbb{F}_2^n \times \mathbb{F}_2^\ell \to \mathbb{F}_2$ of F, $t \in [1,m]$, can be expressed by its algebraic normal form (ANF). That is, we can represent

F_t uniquely as a multivariate polynomial.

$$F_t(x,k) = \sum_{u\in\mathbb{F}_2^n, v\in\mathbb{F}_2^\ell} \lambda_{u,v} x^u k^v = \sum_{u\in\mathbb{F}_2^n, v\in\mathbb{F}_2^\ell} \lambda_{u,v} \prod_{i:\ u_i=i} x_i \prod_{j:\ v_j=1} k_j, \qquad \lambda_{u,v} \in \mathbb{F}_2.$$

There are many attacks based on specific properties of the ANF of F_t (for any coordinate t), some of them differ in details or targets, but are all based on exploitable knowledge about this ANF. One of the first attacks is based on F_t having small degree in the variables x. Indeed, if F_t has degree d, any its $(d+1)$-th derivative is the constant zero function. Those attacks were coined high-order differential attacks in [1].

Integral attacks [2], make use of a similar but more fine-grained property. Initially, for integral attacks one first fixes part of the input x and then exploits a small degree of the resulting function. In terms of higher-order differentials, those can be seen as a special case where only certain $(d+1)$-th derivatives vanish. However, in their most general form, integral attacks would not correspond to derivatives anymore but to a key-independent sum of cipher-texts for a well chosen subset of all plain-texts.

Cube attacks [3], instead, focus on derivatives that are simple but have not vanished yet. Here, simplicity means low algebraic degree and/or dependence only on few secret variables. Such functions can be used to recover partial information about the secret key. Cube attacks are thus always key recovery attacks, while integral properties by themselves only define a distinguisher of the primitive from a random one. While integral distinguishers can be used to mount key recovery attacks for some primitives such as block ciphers, it is not possible, for example, for stream ciphers.

1.3 Division Property

One major obstacle of the attacks above is to actually find out that, e.g., the degree of a cipher is small, or, more generally, to find derivatives that vanish or have sparse ANF. The function F is naturally very complex as a whole. Almost all symmetric ciphers are composed of simple and almost identical functions $F^{(i)}$. Those functions $F^{(i)}$, called *round functions*, are in particular very efficient to implement and often have very simple ANFs and very low degree. Division property, in all its variants described below, is now a tool that allows to deduce information about the ANFs of F by analyzing its structure as a composition of those simple functions $F^{(i)}$. In its initial form, it would allow to derive better upper bounds on the degree of F (defined as the maximum degree of all its coordinates), i.e., potentially find better attacks. Later, variants were developed that allow the efficient computation of single entries in the ANF. We will discuss the mathematical aspects of the most important variants in this paper.

1.4 Brief Overview of the Division Property Variants

Yosuke Todo [4] introduced the division property as a new technique to find better integral attacks. This initial version is referred to as the *conventional division property*. Here, the state of a cipher is grouped into words (e.g. bytes) and the division property allows to make certain statements about the ANF of the cipher. More precisely, the conventional division property allows to determine upper bounds on the degree of the cipher by keeping track of bounds on the degree *of a certain set indicator* in the separate words and how those change for the basic operations that are applied in the round functions. This led to new attacks, most prominently against the block cipher MISTY1 [5, 6].

While the grouping into words allows rather efficient calculations, its information was limited and therefore Todo and Morii [7] later introduced the *bit based division property* as a refinement.

Both variants allow to compute a set of certain monomials (and their multiples) that do not occur in the ANF of the function $F_t(\cdot, k)$ for any key k. However, for monomials which are not contained in the set, no information is known. One of the main advantages of such a representation is that it allows to cope with the often seen key-addition of an unknown round key to the state.

In this sense the bit based division property splits the space of monomials into two distinct parts, one for which the coefficients are zero, and one for which the coefficients remain unknown.

In a next step, Todo and Morii, in the same paper extended the framework to be able to make statements about monomials of $F_t(\cdot, k)$ which are always present independently of k (i.e., the respective coefficient is the constant 1). This is referred to as the *three-subset division property* to highlight that now all monomials are split into three sets: one for which the coefficients in the ANF of f are known to be zero, one for which they are known to be one, and the rest for which nothing can be concluded.

In order to allow the computation of even more elements of the ANF, Hao et al. [8, 9] introduced the *three-subset division property without unknown subset*. This paper shifted the view from $F_t(\cdot, k)$ to $F_t(\cdot, \cdot)$, i.e., the key is treated as a usual variable. They give algorithms, based on division trails introduced in [10], that allow to compute (a limited number of) ANF coefficients of F_t exactly. At least theoretically, this removes the set of monomials for which no information can be computed, hence the name.

It should be noted that the three-subset division property without unknown subset actually separates all monomials only in two sets, more precisely in one set and its complement. This concept, without the important aspect of computational hardness, was already treated by Boura and Canteaut [11] using the term *parity set*.

1.5 Brief Comparison of Division Property Variants and Classic Degree Bounds

Division property can be seen as evolution and specification of classic methods of bounding the algebraic degree of iterated functions. The most powerful of the classic ones are the naive bound

$$\deg g \circ F \leq \deg g \cdot \deg F, \qquad F : \mathbb{F}_2^n \to \mathbb{F}_2^m,\ g : \mathbb{F}_2^m \to \mathbb{F}_2,$$

the Boura-Canteaut bound [12] (extensively used in cryptanalysis, see e.g. [13])

$$\deg g \circ F \leq n - \left\lceil \frac{n - \deg g}{\deg F^{-1}} \right\rceil, \quad F : \mathbb{F}_2^n \to \mathbb{F}_2^m \text{ a bijection},\ g : \mathbb{F}_2^m \to \mathbb{F}_2,$$

and a more recent family of bounds based on the degrees of the involved *graph indicators* by Carlet [14], in particular,

$$\deg g \circ F \leq \deg g + \deg \mathbb{1}_{\Gamma_F} - m, \qquad F : \mathbb{F}_2^n \to \mathbb{F}_2^m,\ g : \mathbb{F}_2^m \to \mathbb{F}_2,$$

where $\mathbb{1}_{\Gamma_F} : \mathbb{F}_2^{n+m} \to \mathbb{F}_2$ is such that $\mathbb{1}_{\Gamma_F}(x, y) = 1$ if and only if $F(x) = y$.

These bounds are not always tight and thus clearly can not provide better distinguishers than the *perfect* division property (of course, the downside of the latter is that it is not always feasible to compute). The relationships of these bounds to the conventional division property are described in Section 5. Notably, the state-based conventional division property with ideal propagation can be viewed as a slight generalization of the Carlet's graph indicator-based method (requiring more information about the graph indicators and yielding stronger bounds).

Table 1 summarizes the hierarchy of these classic methods and division property in terms of precision and required information.

In the table, it is assumed that the same decomposition $F = F^{(r)} \circ \ldots \circ F^{(1)}$ of the target function F is used, and the "required information" used by the technique is given precisely for those functions $F^{(i)}$. The table does not include all classic degree bounds (for example, bounds based on the divisibility of the Walsh spectra of the involved functions, or requiring more specific information). The described classic degree bounds are not comparable (i.e., for each bound, there exist instances of functions where this bound is strictly better than the others) and so are grouped together. A relevant detailed study of degree bounds can be found in [14]. Furthermore, relationship of division property variants and the Boura-Canteaut bound was studied in [15].

The argumentation for the provided hierarchy is as follows:

(A)⇒(B): Follows from the relation of conventional division property to maximal monomials in the graph indicators of composed functions (proven in [16], see Section 5).
(B)⇒(C)⇒(D): Follows from inclusion of the conventional division property variants by the partial weights computations (see Section 4).

Table 1 A high-level hierarchy of techniques for searching integral distinguishers. In the top-to-bottom order, the precision increases (each next technique can find more distinguishers when feasible) at the cost of more information required about the involved functions.

Technique	Required information
(A) classic degree bounds (naive, Boura-Canteaut, Carlet), state-based conventional DP with degree-based propagations	degree of the functions, degree of their inverses, degree of their graph indicators
(B) state-based conventional DP	state-based conventional division property propagations, or, equivalently, maximal pairs of degrees of the involved graph indicators
(C) word-based conventional DP	word-based conventional division property propagations
(D) bit-based conventional DP	bit-based conventional division property propagations
(E) bit-based three-subset DP	(a mix of) bit-based conventional/perfect division property propagations
(F) perfect DP (also known as bit-based three-subset DP without the unknown subset, monomial prediction)	perfect division property propagations, equivalent to full specification of functions

(D)⇒(E)⇒(F): Follows from the inclusion of bit-based division property variants: three-subset division property is simply a mix of the conventional and the perfect division property (see [15, 17]).

1.6 Outline

The parity sets and the perfect division property is described in Section 3. It is foundational for a formal description of the original (conventional) division property, which we provide in Section 4. As mentioned above, historically, conventional division property was described and applied earlier than the perfect division property. In Section 5.2, we provide a comparison of (conventional) division property with known generic degree bounds. Section 6 describes the application of perfect division property for providing security arguments of symmetric-key primitives. Finally, Section 7 concludes the work.

2 Preliminaries and Notation

2.1 General Notations

The integer segment $(i, i+1, \ldots, j)$ is denoted by $[i, j]$. The all-one and all-zero vectors are denoted by $\underline{1}$ and $\underline{0}$ respectively. The i-th unit vector is the vector having 1 at the i-th position and 0 otherwise, it is denoted by e_i.

2.2 Boolean Vectors and Functions

An n-bit Boolean function is a function $f\colon \mathbb{F}_2^n \to \mathbb{F}_2$, where $\mathbb{F}_2 = \{0,1\}$ is the binary field and $\mathbb{F}_2^n$ is the n-dimensional vector space over $\mathbb{F}_2$. The *support* of a Boolean function f is the set of all preimages of 1 under f. An $n \times m$ vectorial Boolean function is a function $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$. The *graph* of F is denoted by Γ_F and is given by

$$\Gamma_F \coloneqq \{(x, F(x)) \mid x \in \mathbb{F}_2^n\} \subseteq \mathbb{F}_2^n \times \mathbb{F}_2^m.$$

The bitwise logical AND and OR are denoted by $\wedge\colon \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{F}_2^n$ and $\vee\colon \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{F}_2^n$ respectively. The logical negation $\neg$ is defined by $\neg\colon \mathbb{F}_2^n \to \mathbb{F}_2^n\colon x \mapsto \underline{1} - x$. The inner product of $u, v \in \mathbb{F}_2^n$ is defined as $\langle u, v\rangle \coloneqq \sum_{i=1}^n u_i v_i \in \mathbb{F}_2$. For $x, u \in \mathbb{F}_2^n$, the notation x^u is a shorthand for $\prod_{i=1}^n x_i^{u_i} \coloneqq \prod_{i\in[1,n],u_i=1} x_i$.

The indicator of a set $X \subseteq \mathbb{F}_2^n$ is defined as $\mathbb{1}_X\colon \mathbb{F}_2^n \to \mathbb{F}_2$ such that $\mathbb{1}_X(x) = 1$ if and only if $x \in X$. By an abuse of notation, we will identify a set $X \subseteq \mathbb{F}_2^n$ with its indicator $\mathbb{1}_X : \mathbb{F}_2^n \to \mathbb{F}_2$. In particular, the *degree* of a set X is defined as the algebraic degree of its indicator, i.e., $\deg X \coloneqq \deg \mathbb{1}_X$. Conversely, we will identify a Boolean function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ with its support $\operatorname{supp}(f) \coloneqq \{x \in \mathbb{F}_2^n \mid f(x) = 1\} \subseteq \mathbb{F}_2^n$, The indicator of the graph of a vectorial Boolean function F will be called shortly *the graph indicator of* F.

The *symmetric difference*, denoted by $\triangle$, is a binary operation on sets equal to the set of elements each present in exactly one of the two input sets. It is commutative and associative, so that a symmetric difference of a collection of sets is well defined, and is equal to the set of elements each present in an odd number of the input sets.

We adapt the set-builder notation and applications of functions by defining the resulting set to be constructed from elements having *odd multiplicity* (i.e., ignoring elements with even multiplicity), for example:

- for a function $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$, the notation $\{F(x) \mid x \in \mathbb{F}_2^n\}$ defines the set of all vectors y in the image of F having an odd number of preimages;
- for $F : \mathbb{F}_2^n \to \mathbb{F}_2^m, X \subseteq \mathbb{F}_2^n$, the notation $F(X)$ defines the set of all $y \in \mathbb{F}_2^m$ such that there exists an odd number of $x \in X$ such that $y = F(x)$.

An exception is the partial weights vector function

$$\operatorname*{wt}_{k_1,\ldots,k_r} : \mathbb{F}_2^n \to [0, k_1] \times \ldots \times [0, k_r]$$

(used in Section 4 and Section 5), which is explicitly defined to act on sets as the union of its applications to the individual elements.

Every Boolean function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ admits a unique representation as a multivariate polynomial over $\mathbb{F}_2$ and n variables:

$$f \in \mathbb{F}_2[x_1, \ldots, x_n]/(x_1^2 - x_1, \ldots, x_n^2 - x_n), f(x) = \sum_{u\in\mathbb{F}_2^n} \lambda_u x^u,$$

where $\lambda_u \in \mathbb{F}_2$ are constant coefficients independent of x. This representation is called *the algebraic normal form (ANF)*. We will say that a monomial x^u belongs to the ANF of f (or simply "belongs to $f(x)$") if $\lambda_u = 1$.

The ANF support (the set of all $u \in \mathbb{F}_2^n$ such that x^u is present in the ANF) of a Boolean function $f\colon \mathbb{F}_2^n \to \mathbb{F}_2$ or of a subset X of $\mathbb{F}_2^n$ is denoted by $\mathcal{A}(f)$ and $\mathcal{A}(X) \coloneqq \mathcal{A}(\mathbb{1}_X) \subseteq \mathbb{F}_2^n$ respectively.

Example 1 Let $f : \mathbb{F}_2^4 \to \mathbb{F}_2$ be given by $f(x) = x_1x_2x_4 + x_3x_4 + x_4$ and let $X = \operatorname{supp}(F)$. Then,

$$\mathcal{A}(X) \coloneqq \mathcal{A}(f) = \{(1,1,0,1),(0,0,1,1),(0,0,0,1)\} \subseteq \mathbb{F}_2^n.$$

Fact 1 (See, e.g., [18]) *For any $f : \mathbb{F}_2^n \to \mathbb{F}_2$ and any $u \in \mathbb{F}_2^n$, the corresponding ANF coefficient λ_u is given by*

$$\lambda_u = \sum_{x \preceq u} f(x).$$

2.3 Partial Order

We use the product order ($\preceq$) on $\mathbb{Z}^n$ (including $\mathbb{F}_2^n$ by $\mathbb{F}_2 \simeq \{0,1\} \subseteq \mathbb{Z}$), where $x \preceq y$ if and only if $x_i \leq y_i$ for all $i \in [1,n]$, and $x \prec y$ if and only if $x \preceq y$ and $x \neq y$.

Fact 2 *For $x, u \in \mathbb{F}_2^n$, the relation $x \succeq u$ is equivalent to the expression $x^u = 1$; similarly, $x \preceq u$ is equivalent to $(\neg x)^{\neg u} = 1$.*

Proof The expression $x \succeq u$ means $\bigwedge_{i=1}^n (x_i \geq u_i)$, equivalent to $\bigwedge_{i=1}^n (x_i^{u_i} = 1) = \prod_{i=1}^n x_i^{u_i} = x^u$. The case of $x \preceq u$ is analogous. □

In the following, we assume that a *universe* set $\Omega \subseteq \mathbb{Z}^n$ is given by the context and it is of the shape

$$\Omega = [0,k_1] \times [0,k_2] \times \ldots \times [0,k_n],$$

where $k_1, \ldots, k_n \in \mathbb{Z}_{>0}$.

The logical negation is naturally generalized to operate on $[0,k]$ (where $k \in \mathbb{Z}_{>0}$):

$$\neg\colon [0,k] \to [0,k]\colon x \mapsto k - x.$$

Definition 1 (Lower/upper sets/closures) The *lower closure* of a set $X \subseteq \Omega$, denoted by $\downarrow X$, is the set of all $u \in \Omega$ with $u \preceq x$ for some $x \in X$:

$$\downarrow X \coloneqq \{u \in \Omega \mid \exists x \in X : u \preceq x\} = \bigcup_{x \in X} \{u \in \Omega \mid u \preceq x\}.$$

The *upper closure* of a set $X \subseteq \Omega$, denoted by $\uparrow X$, is the set of all $u \in \Omega$ with $x \preceq u$ for some $x \in X$:

$$\uparrow X \coloneqq \{u \in \Omega \mid \exists x \in X : u \succeq x\} = \bigcup_{x \in X} \{u \in \Omega \mid u \succeq x\}.$$

A set X is an *upper set* if its upper closure is X itself. A set X is a *lower set* if its lower closure is X itself.

Example 2 Let $X = \{(0,0,1),(0,1,1)\} \subseteq \mathbb{F}_2^3$. Then,

$$\downarrow X = \downarrow\{(0,1,1)\} = \{(0,0,0),(0,0,1),(0,1,0),(0,1,1)\},$$
$$\uparrow X = \uparrow\{(0,0,1)\} = \{(0,0,1),(0,1,1),(1,0,1),(1,1,1)\}.$$

Let $Y = \{(0,2),(2,1)\} \subseteq [0,3]^2$. Then,

$$\downarrow Y = \{(0,0),(0,1),(0,2),(1,0),(1,1),(2,0),(2,1)\},$$
$$\uparrow Y = \{(0,2),(0,3),(1,2),(1,3),(2,1),(2,2),(2,3),(3,1),(3,2),(3,3)\}.$$

See Figure 1 for an illustration.

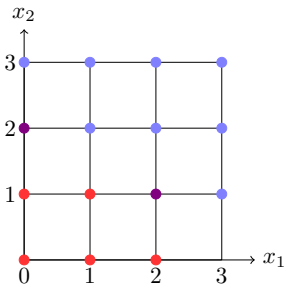


Fig. 1 The lower closure $\downarrow Y$ (red and purple) and the upper closure $\uparrow Y$ (blue and purple) of the set $Y = \{(0,2),(2,1)\}$ (purple).

Definition 2 (Convexity) A subset $X \subseteq \mathbb{Z}^n$ is called *convex*, if for any $a, b, c \in \mathbb{Z}^n$, $a \preceq b \preceq c$ and $a, c \in X$ imply $b \in X$.

Definition 3 (Min-/max-set) The *max-set* of a set $X \subseteq \mathbb{Z}^n$, denoted by $\mathrm{Max}(X)$, is the set of all maximal elements in X:

$$\mathrm{Max}(X) \coloneqq \{u \in X \mid \nexists x \in X : x \succ u\}.$$

The *min-set* of a set $X \subseteq \mathbb{Z}^n$, denoted by $\mathrm{Min}(X)$, is the set of all minimal elements in X:

$$\mathrm{Min}(X) \coloneqq \{u \in X \mid \nexists x \in X : x \prec u\}.$$

Fact 3 *The operator* $\neg$ *anti-commutes with* Min, Max, $\downarrow$, $\uparrow$*: for any set* X*,*

$$\neg\mathrm{Min}(X) = \mathrm{Max}(\neg X), \qquad \neg\downarrow X = \uparrow\neg X,$$
$$\neg\mathrm{Max}(X) = \mathrm{Min}(\neg X), \qquad \neg\uparrow X = \downarrow\neg X.$$

Proof Follows from the definitions and the fact, that, for all u, v, it holds $u \preceq v$ if and only if $\neg u \succeq \neg v$. □

Definition 4 (Antichain) A set $X \subseteq \mathbb{Z}^n$ is called an *antichain* if all pairs of distinct elements from X are incomparable.

Fact 4 *A set $X \subseteq \mathbb{Z}^n$ is a max-set (or a min-set) of some set if and only if it is an antichain.*

Proof If X is a max-set (or a min-set), then it can not contain two elements u, v such that $u \preceq v$ (otherwise, one of them must be removed from X), and so is an antichain. Conversely, if X is an antichain, then it is the max-set of $\downarrow X$ and the min-set of $\uparrow X$. □

2.4 Symmetric-key Primitives

Symmetric cryptography plays an important role in securing our daily sensitive data. Due to their great performance advantages compared to their public-key counterparts, symmetric-key primitives are responsible for the encryption and authentication for any data that is protected.

The most important basic building blocks of symmetric cryptography are block ciphers, stream ciphers and cryptographic permutations.

A *block cipher* is a mapping

$$E\colon \mathbb{F}_2^n \times \mathbb{F}_2^\ell \to \mathbb{F}_2^n$$

mapping an n-bit message x and an ℓ bit key k to an n bit cipher-text

$$c = E(m, k)$$

such that for each fixed key $k \in \mathbb{F}_2^\ell$ the restriction

$$x \mapsto E(x, k)$$

is a permutation on $\mathbb{F}_2^n$. Common sizes for n, called the block size, are 64 and 128 and ℓ, the key-size, range from 80 to 256.

A *stream cipher* most commonly is based on a key stream generator

$$F\colon \mathbb{F}_2^n \times \mathbb{F}_2^\ell \to \mathbb{F}_2^t$$

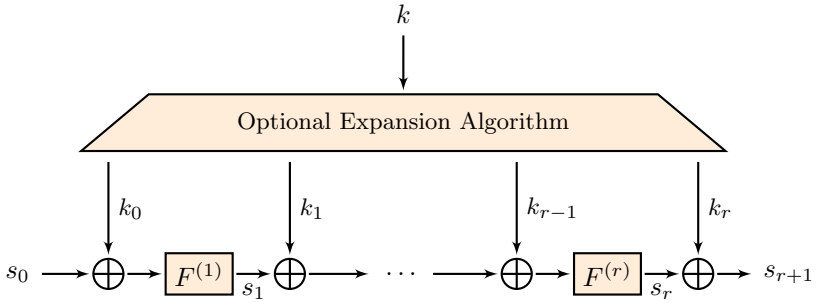
that takes as inputs an initial value IV, a key k, and produces a (very long) stream of bits

$$z = F(\mathrm{IV}, k).$$

The actual encryption of a message is then performed by simply adding (in $\mathbb{F}_2$) the message with the output of the key-stream generator.

Those primitives are used in modes of operation to ensure secure encryption, authentication or a combined mode for authenticated encryption. The security of those modes can usually be proven under some assumption on the underlying primitive. One standard assumption is that it should be practically impossible to distinguish the primitive using an unknown key from a random function, or, in case of a block cipher, from a random permutation. The security of the primitive itself cannot be proven but is rather checked against known classes of attacks. The most well known statistical attacks are differential and linear cryptanalysis and their derivatives. But other large classes exist and are applicable to (reduced) versions of many ciphers. Some of the most important attacks are those that investigate the ANF of a cipher. They are described in the following parts.

Virtually all primitives deployed in practice are iterative designs. For the case of a block cipher, this means that a simple to analyze and easy and efficient to implement function is iterated a fixed, well chosen, number of times. Often, the key, or rather a round-specific part of the key, is simply added to the current state between the rounds. This constitutes what is known as a key-alternating cipher and is depicted[1] below.

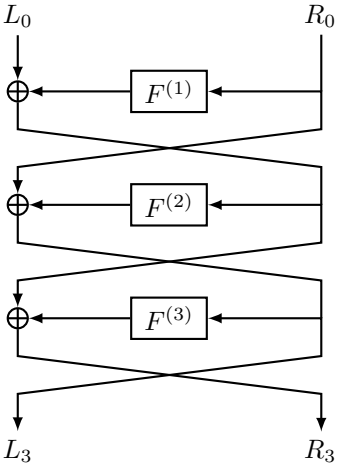


Most of the symmetric primitives we are using today are not only iterative designs but also use what is known as an SP-Network as round function. In the substitution part, a non-linear layer of parallel small-sized permutations (the so-called S-Boxes) is used. The permutation part on the other hand consists of a (binary) linear layer applied to the full state. Using those ingredients when designing efficient and secure ciphers or cryptographic permutations has a long-standing history. It can be seen as having its roots already in Shannon's seminal ideas on confusion and diffusion [20]. While certainly many alternative design strategies exist, the use of S-boxes and linear layers is arguably dominating today's designs and include AES, SHA-3, and many of the primitives for the final round of the NIST lightweight cryptography competition.

Another design strategy we want to briefly mention are Feistel ciphers, with DES being the most prominent but by now out-dated example. Here the message is first split into two halves. The right half is input to a round function f_i, that also takes the key as an input. The output of f_i is added to the left half and finally both parts are swapped. The function f_i are often again based

[1] This, and the picture of the Feistel cipher are taken from [19].

on linear-layers and S-boxes, just as described above for block ciphers. Three rounds of a Feistel cipher are shown below.



2.5 Integral Cryptanalysis

Integral attacks [2] and cube attacks [3] are two important attack vectors on symmetric primitives, which exploit certain properties of the involved ANF expressions. Integral attacks *distinguishing* the analyzed primitive from an ideal one can be applied both to block ciphers and stream ciphers. However, such distinguishers can be extended into *key recovery* attacks only for block ciphers. Cube attacks, on the other hand, can be used to mount key recovery in both block ciphers and stream ciphers.

To mount a distinguishing integral attack on a block cipher $E\colon \mathbb{F}_2^n \times \mathbb{F}_2^\ell \to \mathbb{F}_2^n$, an adversary chooses a proper non-empty subset $S \subset \mathbb{F}_2^n$ and a non-zero output mask $\beta \in \mathbb{F}_2^n$, so that the sum

$$\sum_{x \in S} \langle \beta, E_k(x) \rangle$$

is constant zero or one, i.e., is independent of the key k. This property can be used as a distinguisher since for a random permutation $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^n$, the sum $\sum_{x \in S} \langle \beta, F(x) \rangle$ is zero or one with a probability 1/2 each (excluding the trivial cases $S = \emptyset$ and $S = \mathbb{F}_2^n$).

An integral distinguisher can be extended into a key recovery attack in the following way. Let $E_k = G_k \circ F_k$, where $G_k, F_k : \mathbb{F}_2^n \to \mathbb{F}_2^n$ are bijective key-dependent parts of the cipher, and let F_k have an integral distinguisher

$$\sum_{x \in S} \langle \beta, F_k(x) \rangle = 0$$

for all keys k. Then, it is also true that

$$\sum_{x \in S} \langle \beta, G_k^{-1}(E_k(x)) \rangle = 0.$$

However, under certain assumptions, for a key $k' \neq k$, the equation

$$\sum_{y \in S'} \langle \beta, G_{k'}^{-1}(y) \rangle = 0 \tag{1}$$

would hold only with probability $1/2$, providing a way to distinguish a wrong key k' from the right key k. In particular, if $\langle \beta, G_k^{-1} \rangle$ depends only on part of the key, the set of candidate values for this part can be reduced by checking the equation (1). This method can be improved by advanced techniques such as partial sums [21] or FFT-based key recovery [22].

2.5.1 Finding Integral Distinguishers with Division Property

Division property is a technique for finding integral distinguishers based on degeneracies of ANF expressions of the analyzed ciphers.

Conventional division property (Section 4) may only exhibit a monomial such that *all its monomial multiples* are missing in the ANF. In addition, the technique is imperfect, meaning that it does not guarantee finding a distinguisher even if it exists; nonexistence of a distinguisher can never be proven either. Roughly speaking, conventional division property does not detect *cancellations* of monomials in intermediate operations, which may or may not lead to distinguishers; the technique only detects cases when the given monomial can not be computed in principle, due to nonexistence of multiplication paths (the so-called "division trails", see Section 3.2). However, application of conventional division property is feasible for nearly all used ciphers in the literature and often yields powerful distinguishers.

Perfect variants of division property (Section 3), on the other hand, can be used to compute a chosen ANF coefficient exactly. Simply speaking, the technique counts the number of multiplication paths (division trails) for the target monomial, and the *parity* of the count determines the respective ANF coefficient. In practice, however, it is feasible only in some use-cases (e.g., 64-bit block ciphers) and often requires specific optimizations and fine-tuning of utilized solvers or optimization software.

2.6 Cube-based Cryptanalysis

For cube attacks, we define a Boolean function $F\colon \mathbb{F}_2^n \times \mathbb{F}_2^\ell \to \mathbb{F}_2$, where the first input is public and the second input is the secret key. In the context of stream ciphers, the public input is the initialization vector and F computes a bit of the key-stream. Now, the public input is split into two parts and F decomposed as

$$F(x, y, k) = x^{\underline{1}} \cdot p(y, k) + r(x, y, k)$$

where p, r are polynomials, $(x, y) \in \mathbb{F}_2^n$, and $x^{\underline{1}}$ does not occur in the ANF of r. Recall that $x^{\underline{1}}$ corresponds to the product of all variables in x. The polynomial p is called the *superpoly*. It holds that

$$\sum_x F(x, y, k) = p(y, k),$$

since r vanishes. When now for a fixed y the polynomial $p(y, k)$ has a low complexity in k, the adversary obtains information about the key or can use p to filter keys.

2.6.1 Finding Cube Attacks with Division Property

Division property is a state-of-the-art tool for finding cube attacks. Its feasibility range is much larger than classic methods based on empirical evaluation of cubes (as in [3]).

Conventional division property may be used to search for potential cube attacks. More precisely, it allows to upper-bound the shape of monomials in the superpoly $p(y, k)$. However, it can not show that $p(y, \cdot)$ is non-trivial, i.e., that it may actually be used to recover information about the key k (otherwise, the cube key recovery attack degrades to an integral distinguisher; this actually happened in the literature [23]).

Perfect division property can recover the exact coefficients of the superpoly and guarantee a successful key recovery attack. However, its application is much more computationally intensive and is often not feasible.

3 Parity Sets and Perfect Division Property

Parity sets were introduced by Boura and Canteaut [11] to formalize rigorously Todo's original (conventional) division property [4], which can be viewed as a union of vectorial lower bounds on the parity set of a given set. Considering propagation of parity sets directly leads to the *perfect* division property. While an intermediate variant called 3-subset division property was proposed by Todo and Moriai already in [7], fully perfect variants were proposed and shown feasible in practice only a few years later [8, 17, 24].

In this section, we introduce parity sets and study their properties, focusing on propagation of parity sets through functions. These constitute the key concepts behind the perfect division property and also prepare formalisation of the conventional division property (Section 4).

3.1 Parity Sets

Definition 5 (Parity set [11]) Let $X \subseteq \mathbb{F}_2^n$. The parity set of X, denoted by $\mathcal{U}(X)$, is the subset of $\mathbb{F}_2^n$ defined by

$$\mathcal{U}(X) = \left\{ u \in \mathbb{F}_2^n \mid \sum_{x \in X} x^u = 1 \right\}.$$

By an abuse of notation, the parity set of a Boolean function is defined as the parity set of its support.

Remark 1 In the division property literature, parity sets correspond to the "3-subset division property without the unknown subset". More precisely, the 3-subset division property defined in [7] included a set $\mathbb{L} \setminus \mathbb{K}$ being a subset of $\mathcal{U}(X)$ feasible to maintain. Later, [8] defined a set $\tilde{\mathbb{L}}$ exactly as the parity set of X and called the variant "3-subset division property without the unknown subset".

The parity set is equal up to negations to the support of the ANF of the set's indicator.

Proposition 1 ([16]) *The following statements are equivalent characterizations of the parity set:*

$$u \in \mathcal{U}(X) \quad \Leftrightarrow \quad \sum_{x \in \mathbb{F}_2^n} x^u \cdot \mathbb{1}_X(x) = 1 \quad \Leftrightarrow \quad \sum_{x \succeq u} \mathbb{1}_X(x) = 1 \quad \Leftrightarrow \quad (\neg u) \in \mathcal{A}(\neg X).$$

Proof The first equivalence is trivial; the second equivalence follows from Fact 2; the third equivalence is proven by

$$\sum_{x \succeq u} \mathbb{1}_X(x) = 1 \;\Leftrightarrow\; \sum_{\neg z \succeq u} \mathbb{1}_{\neg X}(z) = 1 \;\Leftrightarrow\; \sum_{z \preceq \neg u} \mathbb{1}_{\neg X}(z) = 1,$$

where the latter condition clearly defines the coefficient of $x^{\neg u}$ in the ANF of $\mathbb{1}_{\neg X}(x)$ (Fact 1). □

The second characterization, although trivial, shows an interesting formulation of the parity set as describing monomials x^u that, when multiplied with the (indicator) function, produce the monomial $x^{\underline{1}}$.

The third characterization shows a close similarity to the Möbius transform-based formulation of the ANF coefficients. The only difference is the direction of the summation: the parity set coefficients are sums of the function over *supermasks*, while the ANF coefficients are sums of the function over *submasks*.

The fourth characterization uncovers an equivalence up to negations of the ANF support and of the parity set.

Corollary 1 ([11]) *The mapping $\mathcal{U}$ is involutive.*

Proof Follows from the fact the $\mathcal{A}$ is involutive. □

Proposition 2 *The parity set operator acts linearly on the set indicator. That is, for an integer $k \geq 1$ and for all sets $X, X_1, \ldots, X_k \in \mathbb{F}_2^n$, it holds*

$$\mathbb{1}_X = \sum_{i=1}^{k} \mathbb{1}_{X_i} \quad \text{if and only if} \quad \mathbb{1}_{\mathcal{U}(X)} = \sum_{i=1}^{k} \mathbb{1}_{\mathcal{U}(X_i)}\,.$$

This condition is also equivalent to $X = \triangle_{i=1}^{k} X_i$.

Proof Let $X, X_1, \ldots, X_k \in \mathbb{F}_2^n$ be such that $\mathbb{1}_X = \sum_{i=1}^{k} \mathbb{1}_{X_i}$. Then, for any $u \in \mathbb{F}_2^n$,

$$\mathbb{1}_{\mathcal{U}(X)}(u) = \sum_{x \in X} x^u = \sum_{x \in X_1} x^u + \ldots + \sum_{x \in X_k} x^u = \mathbb{1}_{\mathcal{U}(X_1)}(u) + \ldots + \mathbb{1}_{\mathcal{U}(X_k)}(u).$$

The converse follows from the fact that $\mathcal{U}$ is involutive. The equivalence to the symmetric difference follows from the definition and the fact that the field has characteristic 2. □

Finally, we list a few common examples of parity sets.

Corollary 2 ([11, Cor.1]) *Let X be a subset of $\mathbb{F}_2^n$, $u \in \mathbb{F}_2^n$. Then,*

1. $\mathcal{U}(\emptyset) = \emptyset$.
2. $\mathcal{U}(\{u\}) = {\downarrow}\{u\}$.
3. $\mathcal{U}({\downarrow}\{u\}) = \{u\}$.
4. $\mathcal{U}(\mathbb{F}_2^n) = \{\underline{1}\}$.

Proof 1. Since $\mathbb{1}_\emptyset = 0$ (as a function), its ANF is empty and so must be the parity set (by Proposition 1).
2. Follows from the definition of a parity set, the product $x^{u'}$ equals to 1 if and only if $u' \preceq u$ (the necessary and sufficient condition is to have $u_i = 0$ imply $u'_i = 0$ for all i).
3. Follows from point 2 and the fact that $\mathcal{U}$ is an involution.
4. Is a special case of point 3 with $u = \underline{1}$ and ${\downarrow}u = \mathbb{F}_2^n$.

□

The typical use of parity sets and perfect division property in cryptanalysis reduces to computing an ANF coefficient of a given function, as shown in the following proposition.

Proposition 3 ([24, Cor.1]) *Let $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$ be a vectorial Boolean function, so that $F_i(x) = \sum_{u \in \mathbb{F}_2^n} \lambda_u^{(i)} x^u$. For any $u \in \mathbb{F}_2^n$, let*

$$X = \mathcal{U}(\{u\}) = {\downarrow}\{u\} \;\subseteq\; \mathbb{F}_2^n, \quad Y = F(X).$$

Then, $\lambda_u^{(i)} = 1$ if and only if $e_i \in \mathcal{U}(Y)$.

Proof Follows from Fact 1, and Corollary 2. Indeed, it holds $e_i \in \mathcal{U}(Y)$ if and only if $\sum_{y \in Y} y_i = 1$ if and only if $\sum_{x \in X} F_i(x) = 1$ and X by construction is equal to all $x \in \mathbb{F}_2^n$ such that $x \preceq u$, matching the summing condition of Fact 1. □

3.2 Propagation of Parity Sets and Division Trails

Due to the linearity of the parity set transformation $\mathcal{U}$ (on the set of 2^n-bit vectors representing subsets of $\mathbb{F}_2^n$), propagation of parity sets through a function can be decomposed into $\mathbb{F}_2$-sums of propagations of single entries through that function.

Definition 6 (Propagation [24, Def.2]) Given $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$ and $u \in \mathbb{F}_2^n, v \in \mathbb{F}_2^m$, we say that the input division property u propagates to the output division property v, denoted by

$$u \xrightarrow{F} v$$

(also called a *transition*), if

$$v \in \mathcal{U}(F(\mathcal{U}(\{u\}))).$$

Remark 2 Since $\mathcal{U}$ is an involution, an equivalent formulation is that the set X having $\mathcal{U}(X) = \{u\}$ propagates into $Y = F(X)$ with $v \in \mathcal{U}(Y)$.

Proposition 4 ([24, Prop.1]) *It holds that $u \xrightarrow{F} v$ if and only if the ANF of $F^v(x)$ contains x^u.*

Proof Observe that $v \in \mathcal{U}(F(\mathcal{U}(\{u\})))$ holds if and only if $\sum_{x \in \mathcal{U}(\{u\})} F^v(x) = 1$, by Definition 5. The summing condition $x \in \mathcal{U}(\{u\})$ is equivalent to $x \preceq u$ (see Corollary 2), making the expression equivalent to $\sum_{x \preceq u} F^v(x)$ which is exactly the condition that $F^v(x)$ contains the monomial x^u (by Fact 1). □

An interesting observation is that propagations through a bijective function F are closely related to propagations through the compositional inverse of F.

Proposition 5 ([11, Lem.3]) *Let $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^n$ be a bijection, $u \in \mathbb{F}_2^n, v \in \mathbb{F}_2^m$. Then,*

$$u \xrightarrow{F} v \quad \textit{if and only if} \quad \neg v \xrightarrow{(\neg) \circ F^{-1} \circ (\neg)} \neg u.$$

Proof The proof is simple in the ANF formulation from Proposition 4 by using Fact 2.

$$\sum_{x \preceq u} F^v(x) = 1$$
$$\iff \sum_{x \in \mathbb{F}_2^n} F^v(x) \cdot (\neg x)^{\neg u} = 1$$

$$\iff \sum_{y \in \mathbb{F}_2^n} y^v \cdot (\neg F^{-1}(y))^{\neg u} = 1$$
$$\iff \sum_{y \preceq \neg v} (\neg F^{-1}(\neg y))^{\neg u} = 1.$$

□

The proposition below shows that a vector $b \in \mathbb{F}_2^m$ belongs to the output parity set if and only if the number of vectors $a \in \mathbb{F}_2^n$ in the input parity set, such that a propagates to b, is odd. Before the proposition, we state the following lemma required for the proof.

Lemma 1 *Let $X \subseteq \mathbb{F}_2^n$. Then, $X = \triangle_{u \in U(X)} U(\{u\})$.*

Proof By Proposition 1, we have

$$\mathbb{1}_{\mathcal{U}(X)}(u) = \sum_{x \succeq u} \mathbb{1}_X(x) = \sum_{x \in X} \mathbb{1}_{\mathcal{U}(\{x\})}(u).$$

Swapping the roles of X and $U(X)$ (Proposition 2), we obtain

$$\mathbb{1}_X(x) = \sum_{u \in \mathcal{U}(X)} \mathbb{1}_{\mathcal{U}(\{u\})}(x).$$

The lemma follows. □

Proposition 6 *Let $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m, X \subseteq \mathbb{F}_2^n, Y = F(X) \subseteq \mathbb{F}_2^m$. Then,*

$$\mathcal{U}(Y) = \left\{ v \in \mathbb{F}_2^m \;\middle|\; \left| \left\{ u \in \mathcal{U}(X) \mid u \xrightarrow{F} v \right\} \right| \text{ is odd} \right\}.$$

Proof We recall that here $Y = F(X)$ corresponds to the symmetric difference of sets $\{F(x)\}_{x \in X}$. We have

$$v \in \mathcal{U}(Y) \quad \Leftrightarrow \quad \sum_{y \in Y} y^v = 1 \quad \Leftrightarrow \quad \sum_{x \in X} F^v(x) = 1,$$

where the last equivalence relies on the field having characteristic 2. By Lemma 1, the latter condition is further equivalent to

$$\sum_{u \in \mathcal{U}(X)} \sum_{x \in \mathcal{U}(\{u\})} F^v(x) = 1.$$

The inner sum is equal to 1 if and only if $F^v(x)$ contains the monomial x^u (using Corollary 2 and Fact 1). Using Proposition 4, we conclude that the equivalent to $v \in \mathcal{U}(Y)$ condition is that the number of $u \in \mathcal{U}(X)$ such that $u \xrightarrow{F} v$ holds is odd.

□

In 2016, Xiang et al. [10] introduced the notion of division trails for conventional division property, and it also generalizes naturally to trails of perfect division property [17, 24]. Here, the propagation is not evaluated for a function in a single step. Instead, the definition above is generalized to the setting where F is actually given as the composition of many functions:

$$F = F^{(r)} \circ \cdots \circ F^{(2)} \circ F^{(1)}.$$

Definition 7 (Division/monomial Trail [10, 17, 24]) Let $F\colon \mathbb{F}_2^{n_0} \to \mathbb{F}_2^{n_r}$ be given as

$$F = F^{(r)} \circ \cdots \circ F^{(2)} \circ F^{(1)}$$

where

$$F^{(i)}\colon \mathbb{F}_2^{n_{i-1}} \to \mathbb{F}_2^{n_i}.$$

We call $\left(u^{(0)}, \ldots, u^{(r)}\right), u^{(i)} \in \mathbb{F}_2^{n_i}$, a *division trail* over the sequence $(F^{(i)})_{i\in[1,r]}$ if

$$\forall i \in [1, r]\colon u^{(i-1)} \xrightarrow{F^{(i)}} u^{(i)}.$$

We denote such a trail by

$$u^{(0)} \xrightarrow{F^{(1)}} u^{(1)} \xrightarrow{F^{(2)}} \cdots \xrightarrow{F^{(r)}} u^{(r)}.$$

Remark 3 Hu et al. [17] called division trail a *monomial trail*, since each transition $u \xrightarrow{F^{(i)}} v$ in fact defines monomials x^u and y^v such that $y^v(x)$ contains the monomial x^u, where $y = y(x) = F^{(i)}(x)$.

Theorem 1 ([17, 24]) *Let $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$ be such that $F = F^{(r)} \circ \cdots \circ F^{(2)} \circ F^{(1)}$, where $F^{(i)} : \mathbb{F}_2^{n_i} \to \mathbb{F}_2^{n_{i+1}}$. Let $u \in \mathbb{F}_2^{n_0}, v \in \mathbb{F}_2^{n_r}$. Then,*

$$u \xrightarrow{F} v$$

if and only if the number of trails

$$u \xrightarrow{F^{(1)}} \cdots \xrightarrow{F^{(r)}} v$$

is odd.

Proof The proof is by induction on r, with the base case $r = 2$.

Case $r = 2$:
We have $F = F^{(2)} \circ F^{(1)}$ and then

$$\begin{aligned}
& u \xrightarrow{F} v \\
\Leftrightarrow\ & v \in (\mathcal{U} \circ F \circ \mathcal{U})(\{u\}) && \text{(by Definition 6)} \\
\Leftrightarrow\ & v \in (\mathcal{U} \circ F^{(2)} \circ F^{(1)} \circ \mathcal{U})(\{u\}) && \text{(by } F = F^{(2)} \circ F^{(1)}\text{)} \\
\Leftrightarrow\ & v \in (\mathcal{U} \circ F^{(2)} \circ \mathcal{U})(\mathcal{U} \circ F^{(1)} \circ \mathcal{U})(\{u\}) && (\mathcal{U} \text{ is an involution}) \\
\Leftrightarrow\ & v \in (\mathcal{U} \circ F^{(2)} \circ \mathcal{U})(\{w \in \mathbb{F}_2^{n_1} \mid u \xrightarrow{F^{(1)}} w\}) && \text{(by Definition 6)}
\end{aligned}$$

$$\Leftrightarrow \quad |\{w \in \mathbb{F}_2^{n_1} \mid u \xrightarrow{F^{(1)}} w \xrightarrow{F^{(2)}} v\}| \text{ is odd} \qquad \text{(by linearity of } \mathcal{U}\text{)}.$$

Case $r \geq 3$**:**
Now, we prove the induction step. Let $r' \geq 2$ be such that the theorem is proven up to $r = r'$. Consider the case $r = r' + 1 \geq 3$. Let $F^{(2)\prime} = F^{(r)} \circ \ldots \circ F^{(2)}$ so that $F = F^{(2)\prime} \circ F^{(1)}$. By the base case, $u \xrightarrow{F} v$ if and only if the number of $w \in \mathbb{F}_2^{r_1}$ such that $u \xrightarrow{F^{(1)}} w \xrightarrow{F^{(2)\prime}} v$ holds is odd. For each such w, by the induction hypothesis, we have $w \xrightarrow{F^{(2)\prime}} v$ if and only the number of trails $w \xrightarrow{F^{(2)}} \ldots \xrightarrow{F^{(r)}} v$ is odd. Therefore, the parity of the number of w such that $u \xrightarrow{F^{(1)}} w \xrightarrow{F^{(2)\prime}} v$ holds matches the parity of the number of full trails $u \xrightarrow{F^{(1)}} \ldots \xrightarrow{F^{(2)}} \ldots \xrightarrow{F^{(r)}} v$, and at the same time determines whether $u \xrightarrow{F} v$ holds. □

Example 3 [24, Cor.2] In order to determine whether the i-th output coordinate of F contains a given monomial x^u, it is sufficient to compute the parity (by counting) of the number of trails

$$u \xrightarrow{F^{(1)}} \ldots \xrightarrow{F^{(2)}} \ldots \xrightarrow{F^{(r)}} e_i.$$

3.3 Propagation Rules for Basic Operations

In the previous sections we have already seen that we can describe the division property of a function step by step by splitting the function in multiple parts. Especially, when it comes to real block ciphers with a block size $n = 64$ or $n = 128$ bits, we cannot compute a propagation table of size $2^n \times 2^n$.

In the following we show the propagation rules for five functional complete *basic operations*: the XOR operation, the AND operation, the NEGATION, the COPY operation, and a bit permutation.

Proposition 7 (XOR Propagation) *Let* $\oplus\colon \mathbb{F}_2^{n+1} \to \mathbb{F}_2^n$ *be defined as*

$$\oplus(x_1, x_2, \ldots, x_{n+1}) \coloneqq (x_1 + x_2, x_3, \ldots, x_{n+1}).$$

For $u \in \mathbb{F}_2^{n+1}$ *and* $v \in \mathbb{F}_2^n$*, it holds that* $u \xrightarrow{\oplus} v$ *if and only if*

$$v_1 = u_1 + u_2 \text{ over integers and } \forall i \in \{2, \ldots, n\}\colon v_i = u_{i+1}.$$

That is, either u_1 *or* u_2 *can be equal to* 1*, which leads to* $v_1 = 1$ *in that case, or* $u_1 = u_2 = 0$*, which leads to* $v_1 = 0$*.*

Proof We write $y = \oplus(x)$.

Case $v_1 = 0$**:**
Then $u \xrightarrow{\oplus} v$ if and only if

$$y^v = (y_2, \ldots, y_n)^{(v_2, \ldots, v_n)} = (x_3, \ldots, x_{n+1})^{(v_2, \ldots, v_n)}$$

contains the monomial x^u, thus leading to only one possible trail with $u = (0, 0, v_2, \ldots, v_n)$ and it satisfies the relation above.

Case $v_1 = 1$**:**
Then $u \xrightarrow{\oplus} v$ if and only if

$$\begin{aligned} y^v &= y_1 \cdot (y_2, \ldots, y_n)^{(v_2,\ldots,v_n)} \\ &= (x_1 + x_2) \cdot (x_3, \ldots, x_{n+1})^{(v_2,\ldots,v_n)} \\ &= x_1 \cdot (x_3, \ldots, x_{n+1})^{(v_2,\ldots,v_n)} + x_2 \cdot (x_3, \ldots, x_{n+1})^{(v_2,\ldots,v_n)} \end{aligned}$$

contains the monomial x^u, which holds exactly for $u = (1, 0, v_2, \ldots, v_n)$ and $u = (0, 1, v_2, \ldots, v_n)$. □

Proposition 8 (NEGATION Propagation) *Let* $\neg_1 \colon \mathbb{F}_2^n \to \mathbb{F}_2^n$ *be defined as*

$$\neg_1(x_1, \ldots, x_n) \coloneqq (x_1 + 1, x_2, \ldots, x_n).$$

For $u, v \in \mathbb{F}_2^n$*, it holds that* $u \xrightarrow{\neg_1} v$ *if and only if*

$$\forall i \in \{2, \ldots, n\} \colon v_i = u_i \text{ and } \begin{cases} u_1 = 0 & \text{if } v_1 = 0 \\ u_1 \in \{0, 1\} & \text{if } v_1 = 1 \end{cases}.$$

Proof We denote $y = \neg_1(x)$.

Case $v_1 = 0$**:**
Then $u \xrightarrow{\neg_1} v$ if and only if

$$y^v = (x_2, \ldots, x_n)^{(v_2,\ldots,v_n)}$$

contains the monomial x^u, which leads to the only possible trail, where $u = (0, v_2, \ldots, v_n)$.

Case $v_1 = 1$**:**
Then $u \xrightarrow{\neg_1} v$ if and only if

$$\begin{aligned} y^v &= (x_1 + 1)(x_2, \ldots, x_n)^{(v_2,\ldots,v_n)} \\ &= x_1 \cdot (x_2, \ldots, x_n)^{(v_2,\ldots,v_n)} + (x_2, \ldots, x_n)^{(v_2,\ldots,v_n)} \end{aligned}$$

contains the monomial x^u, which holds exactly for $u = (1, v_2, \ldots, v_n)$ and $u = (0, v_2, \ldots, v_n)$. □

Proposition 9 (COPY Propagation) *Let* $\vdash \colon \mathbb{F}_2^n \to \mathbb{F}_2^{n+1}$ *be defined as*

$$\vdash (x_1, \ldots, x_n) \coloneqq (x_1, x_1, x_2, \ldots, x_n).$$

For $u \in \mathbb{F}_2^n$ *and* $v \in \mathbb{F}_2^{n+1}$*, it holds that* $u \xrightarrow{\vdash} v$ *if and only if*

$$u_1 = v_1 \vee v_2 \quad \text{and} \quad \forall i \in \{2, \ldots, n\} \colon u_i = v_{i+1},$$

that is, $u_1 = 0$ *leads to* $v_1 = v_2 = 0$*, otherwise there is at least one* $i \in \{1, 2\}$ *such that* $v_i = 1$*.*

Proof We denote $y = \vdash(x)$. Then we have $u \xrightarrow{\vdash} v$ if and only if

$$\begin{aligned} y^v &= (x_1, x_1, x_2, \ldots, x_n)^v \\ &= x_1^{v_1} \cdot x_1^{v_2} \cdot (x_2, \ldots, x_n)^{(v_3, \ldots, v_{n+1})} \end{aligned}$$

contains the monomial x^u.

Case $u_1 = 0$**:**
This leads to the only possible trail where $v = (0, 0, u_2, \ldots, u_n)$.

Case $u_1 = 1$**:**
Here, we have the following three possible trails:

$$\begin{aligned} u &\xrightarrow{\vdash} (1, 0, u_2, \ldots, u_n) \text{ or} \\ u &\xrightarrow{\vdash} (0, 1, u_2, \ldots, u_n) \text{ or} \\ u &\xrightarrow{\vdash} (1, 1, u_2, \ldots, u_n). \end{aligned}$$

□

Proposition 10 (AND Propagation) *Let* $\odot\colon \mathbb{F}_2^{n+1} \to \mathbb{F}_2^n$ *be defined as*

$$\odot(x_1, \ldots, x_{n+1}) \coloneqq (x_1 x_2, x_3, \ldots, x_{n+1}).$$

For $u \in \mathbb{F}_2^{n+1}$ *and* $v \in \mathbb{F}_2^n$*, it holds that* $u \xrightarrow{\odot} v$ *if and only if*

$$v_1 = u_1 = u_2 \quad \textit{and} \quad \forall i \in \{2, \ldots, n\}\colon v_i = u_{i+1}.$$

Proof We denote $y = \odot(x)$.

Case $v_1 = 0$**:**
Then we have $u \xrightarrow{\odot} v$ if and only if

$$\begin{aligned} y^v &= (x_1 x_2, x_3, \ldots, x_{n+1})^v \\ &= (x_3, \ldots, x_{n+1})^{(v_2, \ldots, v_n)} \end{aligned}$$

contains the monomial x^u, which exactly holds for $u = (0, 0, v_2, \ldots, v_n)$.

Case $v_1 = 1$**:**
Here, we have $u \xrightarrow{\odot} v$ if and only if

$$y^v = (x_1 x_2)^{v_1} \cdot (x_3, \ldots, x_{n+1})^{(v_2, \ldots, v_n)}$$

contains the monomial x^u. This leads to the only possible trail where $u = (1, 1, v_2, \ldots, v_n)$. □

Proposition 11 (Bit Permutation Propagation) *Let* $P\colon \mathbb{F}_2^n \to \mathbb{F}_2^n$ *be defined as*

$$P(x_1, \ldots, x_n)_i \coloneqq x_{\pi(i)}$$

where $\pi\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ *is a permutation. For* $u, v \in \mathbb{F}_2^n$*, it holds that* $u \xrightarrow{P} v$ *if and only if* $P(u) = v$.

Proof Let $y = P(x)$, then we have

$$\begin{aligned} y^v &= P(x)^v \\ &= (x_{\pi(1)}, \ldots, x_{\pi(n)})^v \\ &= x^{\left(v_{\pi^{-1}(1)}, \ldots, v_{\pi^{-1}(n)}\right)} \\ &= x^{P^{-1}(v)} \end{aligned}$$

which should be equal to x^u, so that $u \xrightarrow{P} v$ if and only if $P(u) = v$. □

Now, we take a look at an illustrating example of the trail propagation for a simple function.

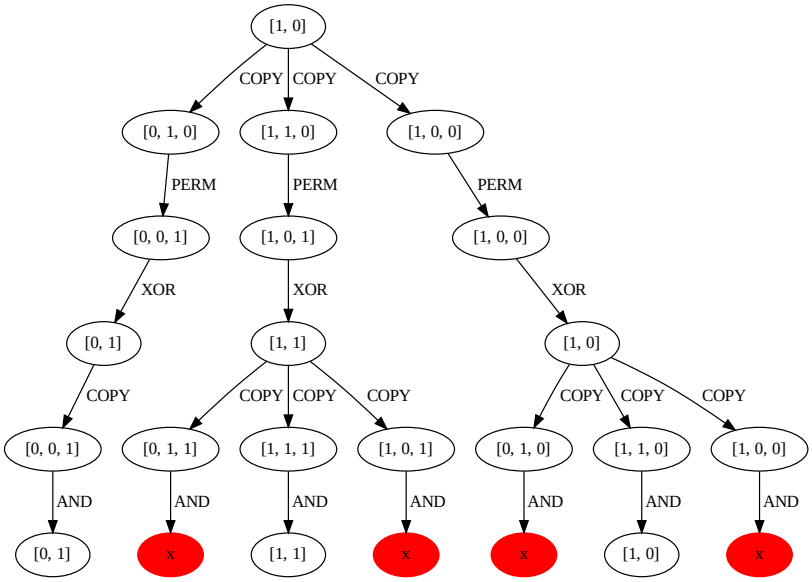


Fig. 2 Trail propagation for $F = \odot \circ \vdash \circ \oplus \circ P \circ \vdash$.

Example 4 (Trail Propagation [25]) Let $F\colon \mathbb{F}_2^2 \to \mathbb{F}_2^2$ be defined as

$$F := \odot \circ \vdash \circ \oplus \circ P \circ \vdash$$

where P interchanges the second and third bit, so that

$$F(x_1, x_2) = \begin{pmatrix} x_1 + x_2 \\ x_1 \end{pmatrix}.$$

Figure 2 shows all trails based on the propagation rules for the input division property $(1, 0)$, i.e. x_1 is contained in y_1, y_2 and $y_1 y_2$.

3.4 Propagation through a Linear Map

Zhang and Rijmen [26] proved a necessary and sufficient characterization of division property transitions through a linear map. While in [26] the statement was given only for invertible maps, Hu, Wang and Wang [27] later proved that it applies to arbitrary linear maps. Although this result was proven in the framework of conventional division property (see Section 4), it can be directly translated to the perfect division property when the weights of the input and the output vectors are equal.

Theorem 2 ([26, 27]) *Let M be a linear map given by a linear matrix $M \in \mathbb{F}_2^{m \times n}$, $u \in \mathbb{F}_2^n, v \in \mathbb{F}_2^m$ such that $\mathrm{wt}(u) = \mathrm{wt}(v)$. Denote by $M_{v,u}$ the submatrix of M with rows selected by ones of the vector v and columns selected by ones of the vector u. Then, $u \xrightarrow{M} v$ if and only if $M_{v,u}$ is invertible.*

Proof Let $M' = M_{v,u}$, $\ell \coloneqq \mathrm{wt}(u) = \mathrm{wt}(v)$. Consider the product $M^v(x)$:

$$M^v(x) = (M'_{1,1}x_1 + \ldots + M'_{1,\ell}x_\ell) \cdot \ldots \cdot (M'_{\ell,1}x_1 + \ldots + M'_{\ell,\ell}x_\ell).$$

The monomial x^u is present in the ANF of $M^v(x)$ if and only if

$$\sum_{\sigma \in S_\ell} \prod_{i=1}^{\ell} M'_{i,\sigma(i)} = 1,$$

where S_ℓ is the set of all permutations of $[1, \ell]$. Note that the left-hand side of the expression is exactly the expression for the permanent of the matrix $M_{v,u}$, which, for the binary field, coincides with the determinant. The theorem follows. □

It is trivial to show that $\mathrm{wt}(u) \leq \mathrm{wt}(v)$ in general, but there is no known simple characterization of transitions $u \xrightarrow{M} v$ when $\mathrm{wt}(u) < \mathrm{wt}(v)$. One possibility is to decompose the map into a sequence of elementary operations and apply the XOR and COPY propagation rules (Proposition 7 and Proposition 9).

Proposition 12 *Let M be a linear map given by a linear matrix $M \in \mathbb{F}_2^{m \times n}$, $u \in \mathbb{F}_2^n, v \in \mathbb{F}_2^m$. Then, $u \xrightarrow{M} v$ implies $\mathrm{wt}(u) \leq \mathrm{wt}(v)$.*

Proof The transition $u \xrightarrow{M} v$ means that the product $M^v(x)$, which is equal to a sum of monomials of degree at most $\mathrm{wt}(v)$, contains the monomial x^u. It follows that $\mathrm{wt}(u) \leq \mathrm{wt}(v)$. □

4 Conventional Division Property

4.1 Definitions

Conventional division property, as introduced in the seminal work by Todo [4], can be viewed as a (vectorial) lower bound on (partial) weights of the

parity set elements. In other words, it considers monomials in the division trails *up to monomial multiples* and, in addition, may also group monomials by their degrees on some *words*, rather than on single bit variables. While these relaxations loose precision (on the contrast to perfect division property), it makes the computational analysis easier and more often feasible.

Definition 8 (Partial weights vector) Let $x \in \mathbb{F}_2^n$ and $k_1, \ldots, k_r \in \mathbb{Z}_{>0}$ such that $k_1 + \ldots + k_r = n$. Define the partial weights vector of x with respect to $(k_1, \ldots, k_r)$ as

$$\underset{k_1,\ldots,k_r}{\mathrm{wt}}(x) \coloneqq (\mathrm{wt}(\mathbf{x}_1), \ldots, \mathrm{wt}(\mathbf{x}_r)) \in [0, k_1] \times \ldots \times [0, k_r],$$

where

$$x \in \mathbb{F}_2^n = (\mathbf{x}_1, \ldots, \mathbf{x}_r) \in \mathbb{F}_2^{k_1} \times \ldots \times \mathbb{F}_2^{k_r}.$$

The universe set is then defined as $\Omega \coloneqq [0, k_1] \times \ldots \times [0, k_r]$. The application of wt to a set $X \subseteq \mathbb{F}_2^n$ is defined as the *union* of the outputs of wt on individual elements:

$$\underset{k_1,\ldots,k_r}{\mathrm{wt}}(X) \coloneqq \bigcup_{x \in X} \underset{k_1,\ldots,k_r}{\mathrm{wt}}(x).$$

Example 5 Let $x = (0, 1, 0, 1, 1, 1) \in \mathbb{F}_2^6$. Then,

$$\begin{aligned} \underset{2,2,2}{\mathrm{wt}}(x) &= (1, 1, 2) \in [0, 2] \times [0, 2] \times [0, 2], \\ \underset{3,3}{\mathrm{wt}}(x) &= (1, 3) \in [0, 3] \times [0, 3], \\ \underset{6}{\mathrm{wt}}(x) &= (4) \in [0, 6]. \end{aligned}$$

Example 6 Let $X = \{(1, 0), (0, 1)\}$. Then, $\mathrm{wt}_2(X) = \{1\}$.

Definition 9 (Conventional division property) A set $X \subseteq \mathbb{F}_2^n$ is said to satisfy division property $\mathbb{K} \subseteq \mathbb{F}_2^n$ if

$$\mathcal{U}(X) \subseteq \uparrow\mathbb{K}.$$

More generally, X is said to satisfy r-dimensional division property $\mathbb{K}' \subseteq [0, k_1] \times \ldots \times [0, k_r]$ if

$$\underset{k_1,\ldots,k_r}{\mathrm{wt}}(\mathcal{U}(X)) \subseteq \uparrow\mathbb{K}'.$$

Division property $\mathbb{K}$ is said to be *minimal* if it is an antichain.

Remark 4 Clearly, a set X satisfies division property $\mathbb{K}$ if and only if X satisfies division property $\mathrm{Min}(\mathbb{K})$, which is minimal.

Remark 5 The most "tight" division properties that a set $X \subseteq \mathbb{F}_2^n$ satisfies are those with $\uparrow\mathcal{U}(X) = \uparrow\mathbb{K}$, for example, $\mathbb{K} \coloneqq \mathrm{Min}(\mathcal{U}(X))$.

The division property $\mathbb{K}$ defined over $\mathbb{F}_2^n$ is called *bit-based* division property. Note that it is included in the general definition due to $\mathrm{wt}_{1,\ldots,1}(x) = x$

for all $x \in \mathbb{F}_2^n$ and corresponds to the case $r = n$ (n-dimensional division property). When $r = 1$, i.e., when $\mathbb{K} \subseteq [0, n]$ (1-dimensional division property), the division property is called *state-based*. When $1 < r < n$, the division property may be called *word-based* if it is aligned with words (e.g. S-boxes) used in the analyzed cryptographic primitive.

From Proposition 1 it follows that conventional division property of a set defines an upper bound on the degree vectors of the monomials in the ANF of the set's indicator. In the special case of 1-dimensional division property, this fact was given already in several initial studies [11, 13, 28]. The following proposition shows the bit-based version of this fact.

Proposition 13 ([16, Prop.4]) *A set $X \subseteq \mathbb{F}_2^n$ satisfies division property $\mathbb{K} \subseteq \mathbb{F}_2^n$ if and only if*

$$\mathcal{A}(X) \subseteq \downarrow\neg\mathbb{K}.$$

More generally, a set $X \subseteq \mathbb{F}_2^n$ satisfies division property $\mathbb{K}' \subseteq [0, k_1] \times \ldots \times [0, k_r]$ if and only if

$$\underset{k_1,\ldots,k_r}{\mathrm{wt}}(\mathcal{A}(X)) \subseteq \downarrow\neg\mathbb{K}'.$$

Proof The proposition follows from Proposition 1 (stating that $\mathcal{U}(X) = \neg\mathcal{A}(\neg(X))$), and the fact that $\mathrm{Max}(\mathcal{A}(X)) = \mathrm{Max}(\mathcal{A}(\neg X))$ and so $\downarrow\mathcal{A}(X) = \downarrow\mathcal{A}(\neg X)$. For the generalized version, it is left to use that $\neg$ and $\downarrow$ commute with wt. □

Example 7 Consider the function

$$f : \mathbb{F}_2^4 \to \mathbb{F}_2 : (x_1, x_2, x_3, x_4) \mapsto x_1x_2x_3 + x_2x_3 + x_3x_4 + x_1$$

and let $X = \mathrm{supp}\,(f) \subseteq \mathbb{F}_2^4$. Then, X satisfies the minimal division property

$$\mathbb{K} = \{(0, 0, 0, 1), (1, 1, 0, 0)\},$$

derived from the maximal monomials in $f = \mathbb{1}_X$ ($x_1x_2x_3$ and x_3x_4). Note that the full parity set of X is given by

$$\mathcal{U}(X) = \{(0, 0, 0, 1), (0, 0, 1, 1), (0, 1, 0, 1), (1, 1, 0, 0), (1, 1, 0, 1), (1, 1, 1, 0)\} \subseteq \uparrow\mathbb{K},$$

which, by Proposition 1, can be derived by negating the exponents in the ANF of

$$f \circ (\neg) : (x_1, x_2, x_3, x_4) \mapsto x_1x_2x_3 + x_1x_2 + x_1x_3 + x_3x_4 + x_3 + x_4.$$

4.2 1-dimensional Conventional Division Property

In this section, we briefly describe main properties and some special cases of 1-dimensional division property, outlined in [11] and partly in [28]. Proposition 13 states that 1-dimensional division property simply corresponds to an upper bound on the degree of a set (i.e., on the algebraic degree of its indicator function). Therefore, this case is covered by the theory of Reed-Muller codes.

For the case of 1-dimensional division property $\mathbb{K} \in \mathbb{F}_2^n$, the following notation is used.

Notation 1 Let $n \in \mathbb{Z}_{>0}, k \in \mathbb{Z}, 0 \leq k \leq n$. Denote by D_k^n the set of all vectors from $\mathbb{F}_2^n$ with weight at least k:

$$D_k^n \coloneqq \{u \in \mathbb{F}_2^n \mid \mathrm{wt}(u) \geq k\} \subseteq \mathbb{F}_2^n.$$

Clearly, if a set X satisfies D_k^n, then X satisfies $D_{k'}^n$ for all k' with $0 \leq k' \leq k$. Therefore, we are typically interested in the highest such value of k. In fact, this is more clear when it is concluded from Proposition 13 that a set X satisfies D_k^n if and only if $\deg X \leq n - k$. We restate this result as in [11] to show the direct connection with Reed-Muller codes.

Proposition 14 ([11, Prop.1]) *Let $X \subseteq \mathbb{F}_2^n$, $k \in [0, n]$. The following statements are equivalent:*

1. *X satisfies division property D_k^n;*
2. *the incidence vector of X belongs to the Reed-Muller code of length 2^n and order $n - k$;*
3. *the incidence vector of X belongs to the dual of the Reed-Muller code of length 2^n and order $k - 1$.*

In particular, a tight lower bound on the size of a set satisfying given D_k^n follows.

Proposition 15 ([11, Prop.2]) *If a non-empty subset $X \subseteq \mathbb{F}_2^n$ satisfies division property D_k^n, then*

$$|X| \geq 2^k.$$

Moreover, the equality $|X| = 2^k$ holds if and only if X is an affine subspace of dimension k.

The sets satisfying division property D_k^n for very small or very large values of k have simple characterizations.

Proposition 16 ([11]) *Let $X \subseteq \mathbb{F}_2^n$. Then,*

1. *X satisfies D_0^n;*
2. *X satisfies D_1^n if and only if its cardinality is even;*
3. *X satisfies D_2^n if and only if its cardinality is even and it is a zero-sum set (i.e., $\sum_{x \in X} x = 0$);*
4. *X satisfies D_3^n if and only if X and all n sets $\{x \in X \mid x_i = 0\}$ $(1 \leq i \leq n)$ are zero-sum sets;*
5. *X satisfies D_{n-1}^n and not D_n^n if and only if X is an affine hyperplane of $\mathbb{F}_2^n$.*
6. *X satisfies D_n^n if and only if X is empty or $X = \mathbb{F}_2^n$.*

Remark 6 One might allow the division property $D_{n+1}^n = \emptyset$ which only the empty set satisfies.

4.3 Propagation of Conventional Division Property

From now on, we focus on bit-based conventional division property. Most definitions and results can be naturally generalized by applying the corresponding partial weights projection map $\mathrm{wt}_{k_1,\ldots,k_r}$.

Conventional bit-based division property may precisely capture the parity set up to the presence of an (unknown) constant addition. However, it is also useful in the analysis of key-less functions such as cryptographic permutations due to its high efficiency.

Proposition 17 ([11, Prop.6]) *Let X be a subset of $\mathbb{F}_2^n$. Then, for any $c \in \mathbb{F}_2^n$,*

$$\mathcal{U}(X+c) \subseteq \uparrow\mathcal{U}(X).$$

Proof By Proposition 13, the set $\neg\mathrm{Min}(\mathcal{U}(X))$ is the set of maximal monomials in the ANF of $\mathbb{1}_X$. Clearly, this set is invariant under shifting X by a constant, therefore, we have $\mathrm{Min}(\mathcal{U}(X)) = \mathrm{Min}(\mathcal{U}(X+c))$ and also

$$\mathcal{U}(X+c) \subseteq \uparrow\mathcal{U}(X+c) = \uparrow\mathcal{U}(X).$$

□

The following proposition shows that this bound is tight and can not be improved without constraining c.

Proposition 18 *For all $X \subseteq \mathbb{F}_2^n$ and for all $u \in \uparrow\mathcal{U}(X)$, there exists $c \in \mathbb{F}_2^n$ such that $u \in \mathcal{U}(X+c)$.*

Equivalently, for all $X \subseteq \mathbb{F}_2^n$ and for all $u \in \downarrow\mathcal{A}(X)$, there exists $c \in \mathbb{F}_2^n$ such that $u \in \mathcal{A}(X+c)$.

Proof The equivalence follows from Proposition 1, stating that $u \in \mathcal{U}(X)$ if and only if $(\neg u) \in \mathcal{A}(\neg X)$. We prove the proposition in the ANF domain.

The proof is by contradiction. Let $X \subseteq \mathbb{F}_2^n$ and $u \in \downarrow\mathcal{A}(X)$ be such that $u \notin \mathcal{A}(X+c)$ for all $c \in \mathbb{F}_2^n$. Let $v \in \mathbb{F}_2^n$ be minimal such that $u \prec v$ and $v \in \mathcal{A}(X)$. In other words, the monomial x^u is not present in the ANF of $\mathbb{1}_{X+c}(x)$ for all c, and the monomial x^v is present in the ANF of $\mathbb{1}_X(x)$. Consider $c = v + u$. An arbitrary monomial x^t in $\mathbb{1}_X(x)$ is replaced in $\mathbb{1}_X(x+c)$ by

$$\prod_{i\,:\,t_i=1,c_i=0} x_i \prod_{i\,:\,t_i=1,c_i=1} (x_i+1).$$

Thus, the monomial x^u appears in $\mathbb{1}_{X+c}(x)$ only from monomials x^t in $\mathbb{1}_X(x)$ that are multiples of x^u (i.e., $u \preceq t$) having $c_i = 1$ in positions where $t_i = 1, u_i = 0$ (i.e., $t_i + u_i = 1$). Since the chosen c is such that $c \wedge u = \underline{0}$, it must be $u \preceq t \preceq u + c = v$. But x^v is the minimal monomial multiple of x^u in the ANF of $\mathbb{1}_X(x)$, so that x^u is added to $\mathbb{1}_{X+c}(x)$ exactly once (from the monomial $x^t = x^v$), and thus x^u must be present in $\mathcal{A}(X+c)$. □

Remark 7 This proposition can be also applied in the integral distinguisher scenario (Section 2.5). That is, in the presence of an unknown constant addition in the input (e.g., a whitening key addition in a block cipher), there exist no monomials always missing in the ANF that are divisors of maximal ANF monomials. In other words, if a given monomial is not present in the ANF (for all values of the constant added in the input), then all its multiples are not present either.

We now switch to the most important component of conventional division property theory - propagation through a public function. From Proposition 6 it is clear that propagation of parity sets through a function can be derived from propagation of each single entry of the input parity set, simply by folding the respective image sets with the symmetric difference. Since conventional division property studies lower bounds of parity sets, a (generally tight) lower bound on the parity set propagation can derived as the union of lower bounds of propagated single-entry parity sets. This is based on the following simple observation: for any sets $X_1, \ldots, X_t$, it holds that

$$\begin{aligned} X_1 \triangle \ldots \triangle X_t &\subseteq X_1 \cup \ldots \cup X_t, \text{ or, equivalently,} \\ \mathbb{1}_{X_1} + \ldots + \mathbb{1}_{X_t} &\preceq \mathbb{1}_{X_1} \vee \ldots \vee \mathbb{1}_{X_t}. \end{aligned}$$

In order to distinguish conventional division property transitions from the perfect ones (Definition 6), we shall call them "weak" transitions. This also emphasizes that they may lose information about the involved sets.

Definition 10 (Weak transition) Let $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m, u \in \mathbb{F}_2^n, v \in \mathbb{F}_2^m$. We write $u \xrightarrow{F}\!\!\!\!\!\rightarrow\ v$ and call it a *weak transition* if $u' \xrightarrow{F} v'$ for some $u' \succeq u, v' \preceq v$.

Remark 8 The term "weak transition" and the notation $u \xrightarrow{F}\!\!\!\!\!\rightarrow\ v$ are introduced specifically in this survey in order to distinguish the perfect and the conventional division property transition.

Proposition 19 *Let $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$. If $X \subseteq \mathbb{F}_2^n$ satisfies division property $\mathbb{K} \subseteq \mathbb{F}_2^n$, then $Y := F(X) \subseteq \mathbb{F}_2^m$ satisfies division property $\mathbb{K}' \subseteq \mathbb{F}_2^m$, where*

$$\mathbb{K}' = \bigcup_{u \in \mathbb{K}} \{v \in \mathbb{F}_2^m \mid u \xrightarrow{S}\!\!\!\!\!\rightarrow\ v \text{ is a weak transition}\}.$$

Furthermore, the induced $\uparrow\mathbb{K}'$ can not be improved in general.

Proof The proof is by contradiction. Assume that Y does not satisfy $\mathbb{K}'$. Then, there must exist $v \in \mathcal{U}(Y) \setminus \uparrow\mathbb{K}'$. By Proposition 6, there must exist a transition $u \xrightarrow{F} v$ with $u \in \mathcal{U}(X) \subseteq \mathbb{K}$, which implies a weak transition $u \xrightarrow{F}\!\!\!\!\!\rightarrow\ v$, contradicting that $v \in \mathcal{U}(Y)$ does not belong to $\uparrow\mathbb{K}'$.

We now prove the tightness by contradiction. Assume that there exists $v \in \mathbb{K}'$ such that, for all sets $X \subseteq \mathbb{F}_2^n$ satisfying the input division property $\mathbb{K}$, the output set $Y := F(X)$ satisfies some division property $\mathbb{K}''$ such that $v \notin \uparrow\mathbb{K}''$. Since division property is defined up to the upper closure $\uparrow$, we can assume without loss of generality that v is minimal in $\mathbb{K}'$. Therefore, there must exist $u \in \mathbb{K}$ such that $u \xrightarrow{F} v$. Let $X = \mathcal{U}(\{u\})$ so that $\mathcal{U}(X) = \{u\}$. It follows that $Y := F(X)$ must have $v \in \mathcal{U}(Y) \subseteq \mathbb{K}''$, leading to contradiction. □

Remark 9 If the input division property is known to be tight (e.g., $\mathbb{K} = \min(\mathcal{U}(X))$), then the defined output division property $\mathbb{K}'$ can be improved for some cases of $\mathbb{K}$ and F. Indeed, let u, u' be minimal in $\mathbb{K}$, $u \neq u'$ such that $u \xrightarrow{F} v$ and $u' \xrightarrow{F} v$ for some v that is minimal in $\mathbb{K}'$ (as defined in Proposition 19). Assume also that for no other vector $w \in \mathbb{K}$ it holds $w \xrightarrow{F} v$. Then, by Proposition 6, the two propagations cancel and so v does not belong to $\mathcal{U}(F(X))$ for all sets X tightly satisfying $\mathbb{K}$. Therefore, $F(X)$ also satisfies division property $\uparrow(\mathbb{K}') \setminus \{v\}$, which is tighter than $\mathbb{K}'$.

This idea of considering *some* cancellations leads to the so-called "three-subset division property" [7].

The following theorem summarizes several characterizations of the set of all weak transitions through a function. To prove it, we will use the following simple lemma.

Lemma 2 ([16, Lem.1]) *Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$, $u \in \mathbb{F}_2^n$. Then,*

$$\sum_{x \in \mathbb{F}_2^n} x^u f(x) = 1 \tag{2}$$

and u is a minimal such vector if and only if the ANF of f contains the maximal monomial $x^{\neg u}$.

Proof Let X be the support of f. By Proposition 1, (2) holds if and only if $(\neg u) \in \mathcal{A}(\neg X)$. Since u is minimal such vector and $\max(\mathcal{A}(\neg X)) = \max(\mathcal{A}(X))$, it follows that $x^{\neg u}$ is maximal in the ANF of f. □

Theorem 3 ([11, Prop.7],[16, Thm.1]) *Let $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m, u \in \mathbb{F}_2^n, v \in \mathbb{F}_2^m$. The following statements are equivalent:*

1. *$u \xrightarrow{F} v$;*
2. *there exist $u' \succeq u, v' \preceq v$ such that $u' \xrightarrow{F} v'$;*
3. *there exist $u' \succeq u, v' \preceq v$ such that $F^{v'}(x)$ contains the monomial $x^{u'}$;*
4. *$(\neg u, v)$ belongs to $\uparrow\mathcal{U}(\Gamma_F)$;*
5. *$(u, \neg v)$ belongs to $\downarrow\mathcal{A}(\Gamma_F)$.*

Remark 10 As the characterized set of all weak transitions is monotone, the characterizations apply to the sets of all extreme elements (characterized by maximal u

and minimal v). That is, it holds that $u \xrightarrow{F} v$ and (u, v) is (maximal,minimal) such vector, if and only if $(u, \neg v)$ is a minimal vector in the parity set of the graph of F, if and only if $x^{\neg u} y^v$ is a maximal monomial in the ANF of the graph indicator of F.

Proof (1 $\Leftrightarrow$ 2) holds by definition. (2 $\Leftrightarrow$ 3) is proved by Proposition 4. (4 $\Leftrightarrow$ 5) is proved by the relation between $\mathcal{A}$ and $\mathcal{U}$ from Proposition 1 and the fact that $\mathrm{Max}(\mathcal{A}(X)) = \mathrm{Max}(\mathcal{A}(\neg X))$.

(3 $\Leftrightarrow$ 4) will be proved by showing equality of the extreme subsets. The condition $(\neg u, v) \in \mathrm{Min}(\mathcal{U}(\Gamma_F))$ holds if and only if

$$\sum_{(x,y)\in\Gamma_F} x^{\neg u} y^v = \sum_{x\in\mathbb{F}_2^n} x^{\neg u} S^v(x) = 1 \tag{3}$$

and $(\neg u, v)$ is minimal such pair. For any fixed v, by Lemma 2, (3) holds with $\neg u$ minimal if and only if $S^v(x)$ contains the monomial x^u and it is maximal in the ANF of S^v. We obtain that extreme elements in the set from point 3 belong to the set from point 4, and extreme elements in the set from point 4 belong to the set from point 3. It follows that the sets of extreme elements coincide (otherwise, we could map an extreme element from one set to the other, find a covering extreme element, map back and show that the initial element could not have been extreme). □

An interesting straightforward corollary is the antisymmetry of transitions with respect to the inverse map. It is a direct analogue of Proposition 5.

Corollary 3 ([16, Prop.6]) *Let $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^n$ be a bijection, $u \in \mathbb{F}_2^n, v \in \mathbb{F}_2^m$. Then, the transition $u \xrightarrow{F} v$ is valid if and only if the transition $\neg v \xrightarrow{F^{-1}} \neg u$ is valid.*

4.3.1 Minimal and Core Transitions

From the characterizations given in Theorem 3 it is clear that some valid transitions may imply validity of some other transitions. Therefore, it is useful to study minimal sets of transitions that imply all others. In addition, for practical purposes, it is convenient to consider transitions which are minimal only in the output: this idea is behind the *division property propagation table* (DPPT), which is used to encode or evaluate the propagation of conventional division property in practice.

Definition 11 Let $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$. A weak transition $u \xrightarrow{F} v$ is called *minimal* (denoted $u \xrightarrow[\min]{F} v$) if there exists no $v' \prec v$ such that $u \xrightarrow{F} v'$.

Definition 12 (DPPT) Let $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$. The division property propagation table (DPPT) of F is the mapping $\mathrm{DPPT}\colon \mathbb{F}_2^n \to \mathcal{P}(\mathbb{F}_2^m)$ given by

$$\mathrm{DPPT}_F(u) = \{v \mid u \xrightarrow{F} v, \nexists v' \prec v : u \xrightarrow{F} v'\}.$$

To push the idea of the minimal set of transitions further, Udovenko [16] introduced the notion of *core* transitions. It stems naturally from the definition of weak transitions (Definition 10).

Definition 13 Let $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$. A weak transition $u \xrightarrow{F} v$ is called a *core* transition (denoted $u \xrightarrow[\text{core}]{F} v$) if there exists no weak transitions $u' \xrightarrow{F} v'$ with $v' \preceq v, u' \succeq u, (u', v') \neq (u, v)$.

Remark 11 From Theorem 3 it is clear that core transitions correspond to *minimal* vectors of $\mathcal{U}(\Gamma_F)$ or, equivalently, to *maximal* monomials in the ANF of the graph indicator.

Importantly, the set of core transitions fully identifies all three defined kinds of conventional division property transitions through the function and its inverse, if it exists. Of course, the same characterization can be derived from the set of maximal monomials in the ANF of the graph indicator.

Theorem 4 ([16, Thm.2]) *Let $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$ and let $\mathcal{D}_F \coloneqq \mathrm{Min}(\mathcal{U}(\Gamma_F))$. Then,*

1. $u \xrightarrow{F} v$ *is valid if and only if* $(\neg u, v) \in {\uparrow}\mathcal{D}_F$;
2. $u \xrightarrow[min]{F} v$ *is valid if and only if* $(\neg u, v) \in \mathrm{Min}_2({\uparrow}\mathcal{D}_F)$;
3. $u \xrightarrow[core]{F} v$ *if and only if* $(\neg u, v) \in \mathcal{D}_F$.

If, in addition, $n = m$ and F is bijective:

4. $v \xrightarrow{F^{-1}} u$ *if and only if* $(u, \neg v) \in {\uparrow}\mathcal{D}_F$;
5. $v \xrightarrow[min]{F^{-1}} u$ *if and only if* $(u, \neg v) \in \mathrm{Min}_1({\uparrow}\mathcal{D}_F)$;
6. $v \xrightarrow[core]{F^{-1}} u$ *if and only if* $(u, \neg v) \in \mathcal{D}_F$.

Here, the subscript of Min *defines the coordinate on which the min-set is computed (for vectors (u, v),* Min_1 *operates on u and* Min_2 *operates on v).*

Remark 12 Clearly, *core* transitions are antisymmetric in the same way as general transitions (Corollary 3). More precisely, the transition $u \xrightarrow[\text{core}]{F} v$ is valid if and only if $\neg v \xrightarrow[\text{core}]{F^{-1}} \neg u$ is valid (if and only if $(\neg u, v) \in \mathcal{D}_F$). However, *minimal* transitions are not generally antisymmetric and are dependent on the direction of F.

Proposition 20 ([16, Prop.8]) *Let $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$. Define the following sets:*

$$\mathbb{I}_F \coloneqq \{(u, v) \in \mathbb{F}_2^n \times \mathbb{F}_2^m \mid \neg u \not\xrightarrow{F} v\},$$
$$\mathbb{M}_F \coloneqq \{(u, v) \in \mathbb{F}_2^n \times \mathbb{F}_2^m \mid \neg u \xrightarrow{F} v, \nexists v' \prec v : \neg u \xrightarrow{F} v'\},$$

$$\mathbb{R}_F \coloneqq \{(u,v) \in \mathbb{F}_2^n \times \mathbb{F}_2^m \mid \neg u \xrightarrow{F} v, \exists v' \prec v : \neg u \xrightarrow{F} v'\},$$

corresponding to the sets of all invalid, minimal, and redundant weak transitions respectively (after the logical negation of the first coordinate).

The sets $\mathbb{I}_F, \mathbb{M}_F, \mathbb{R}_F$ form a partition of $\mathbb{F}_2^n \times \mathbb{F}_2^m$. Moreover, $\mathbb{I}_F$ is a lower set, $\mathbb{M}_F$ is a convex set, $\mathbb{R}_F$ is an upper set.

Proof The sets form a partition by their construction. The set $\mathbb{I}_F$ is a lower set as the complement of the upper set $\uparrow\mathcal{D}_F$.

If $\mathbb{M}_F$ is not convex, then there must exist pairs $(u,v) \in \uparrow\mathcal{D}_F \setminus \mathbb{M}_F$, and (u_1,v_1), $(u_2,v_2) \in \mathbb{M}_F$, such that $(u_1,v_1) \prec (u,v) \prec (u_2,v_2)$. Since $(u,v) \in \uparrow\mathcal{D}_F$, there must exist $v' \prec v$ such that $(u,v') \in \mathbb{M}_F$. It follows that $\uparrow\mathcal{D}_F$ contains $(u_2,v') \succ (u,v')$ and $v' \prec v \preceq v_2$, which contradicts that $(u_2,v_2) \in \mathbb{M}_F$.

For any fixed u, if $(u,v) \in \mathbb{R}_F$, then it is clear from definition of $\mathbb{R}_F$ that $(u,v') \in \mathbb{R}_F$ for all $v' \succeq v$. Therefore, if $\mathbb{R}_F$ is not an upper set, there must exist pairs $(u,v) \in \mathbb{R}_F, (u',v) \in (\uparrow\mathcal{D}_F \setminus \mathbb{R}_F = \mathbb{M}_F)$ such that $u \prec u'$. By definition of $\mathbb{R}_F$, there must exist $v' \prec v$ such that $(u,v') \in \mathbb{M}_F$. It follows that $(u',v') \in \uparrow\mathcal{D}_F$ while $(u',v) \in \mathbb{M}_F$ is such that $v' \prec v$, contradicting the definition of $\mathbb{M}_F$.

□

Note that all three sets can be derived from the set $\mathcal{D}_F$ of core transitions (see Theorem 4):

$$\begin{aligned}\mathbb{I}_F &= \overline{\uparrow\mathcal{D}_F},\\ \mathbb{M}_F &= \mathrm{Min}_v(\uparrow\mathcal{D}_F),\\ \mathbb{R}_F &= \overline{\mathbb{I}_F \cup \mathbb{M}_F} = \uparrow\mathcal{D}_F \setminus \mathbb{M}_F.\end{aligned}$$

4.4 Linear Combinations at the Input and at the Output

As noticed by Lambin, Derbez and Fouque [29], the result of conventional division property propagation through a function may change significantly, if an invertible linear map is composed with the function before computing the propagation table (and its inverse is composed with the consequent function to preserve the functional equivalence of the analyzed primitive). This setting is also useful for finding integral distinguishers which take an arbitrary affine subspace as an input set, as opposed to bit-aligned cubes. In [29], it was suggested to exhaust all invertible linear maps to be composed with the analyzed function. While this is a viable approach for 4-bit S-boxes, it is hardly scalable for bigger functions. Later, Derbez and Fouque [30] showed that exhausting all possible choices of one linear *component* (a linear combination) is sufficient to determining whether an integral distinguisher can be found using this method. More precisely, integral distinguishers with an input affine subspace of codimension 1 can be checked. For the input side, the linear component defines the orthogonal subspace on which the inputs are fixed to a constant. For the output side, the linear component defines the output linear component which is used in the integral distinguisher. Udovenko [16] showed that these two cases are almost identical, due to the antisymmetry of transitions with respect to the inverse mapping (see Corollary 3).

Theorem 5 ([16, Thm.3]) *Let S be a permutation of $\mathbb{F}_2^n$, $\alpha, u, v \in \mathbb{F}_2^n$, $L_\alpha \in GL_n(\mathbb{F}_2)$ be such that $L_\alpha(x) = (\langle \alpha, x \rangle, \ldots)$. Then,*

$$v \xrightarrow{\langle \alpha, S \rangle} (1) \quad \text{if and only if} \quad v \in {\downarrow}\mathcal{A}(\langle \alpha, S \rangle).$$

Similarly,

$$(0, 1, \ldots, 1) \xrightarrow{S \circ L_\alpha^{-1}} u \quad \text{if and only if} \quad \neg u \in {\downarrow}\mathcal{A}(\langle \alpha, S^{-1} \rangle).$$

Proof The first statement follows from Theorem 3 (point 2). The second statement follows from Theorem 3 and Corollary 3 (using $(S \circ L_\alpha^{-1})^{-1} = L_\alpha \circ S^{-1}$). □

Remark 13 In the case S is not bijective, the Boolean function $z \mapsto \left\langle a, S^{-1}(z) \right\rangle$ can be replaced by $z \mapsto \sum_{x \,:\, S(x)=z} \langle a, x \rangle$.

4.5 Propagation through a Linear Map

We now consider conventional division property propagation through linear maps. First, we will show that, in the case of linear maps, minimal transitions coincide with core transitions.

Proposition 21 *Let M be a linear map given by a linear matrix $M \in \mathbb{F}_2^{m \times n}$, $u \in \mathbb{F}_2^n, v \in \mathbb{F}_2^m$. Then, (i) : $u \xrightarrow[core]{M} v$ if and only if $u \xrightarrow[min]{M} v$ (ii) : only if $\mathrm{wt}(u) = \mathrm{wt}(v)$. Furthermore, (iii) : $u \xrightarrow{M} v$ implies $\mathrm{wt}(u) \leq \mathrm{wt}(v)$.*

Proof (iii) : By definition of $u \xrightarrow{M} v$, there must exist $u' \succeq u, v' \preceq v$ such that $u \preceq u' \xrightarrow{M} v' \preceq v$. By Proposition 12, $u' \xrightarrow{M} v'$ implies $\mathrm{wt}(u') \leq \mathrm{wt}(v')$. Therefore, we have $\mathrm{wt}(u) \leq \mathrm{wt}(u') \leq \mathrm{wt}(v') \leq \mathrm{wt}(v)$.

(ii) : From (iii) we already know that $\mathrm{wt}(u) \leq \mathrm{wt}(v)$. Now, assume that $u \xrightarrow[\min]{M} v$ and $\mathrm{wt}(u) < \mathrm{wt}(v)$. It follows that $M^v(x)$ contains a monomial multiple of x^u. It is easy to show that there must exist $v' \prec v$ with $\mathrm{wt}(v') = \mathrm{wt}(u)$ such that $M^{v'}(x)$ contains the monomial x^u, contradicting that $u \xrightarrow[\min]{M} v$. Indeed, all monomials in $M^v(x)$ are multiples of some monomials from $M^{v'}(x)$.

(ii) : It is left to prove that, for linear maps, minimal transitions coincide with core transitions. Let $u \xrightarrow[\min]{M} v$ that is not a core transition. Then, there must exist $u' \succ u$ such that $u' \xrightarrow[\text{core}]{M} v$ (and $u' \xrightarrow[\min]{M} v$ as a special case). However, it is then $\mathrm{wt}(u) < \mathrm{wt}(u') = \mathrm{wt}(v)$, which contradicts $\mathrm{wt}(u) = \mathrm{wt}(v)$ following from $u \xrightarrow[\min]{M} v$ using (ii). □

The result of Zhang and Rijmen [26] (Theorem 2) directly applies to minimal transitions of conventional division property (as originally intended in their work). Indeed, core and minimal transitions $u \xrightarrow[\text{core}]{M} v$ must have $\mathrm{wt}(u) = \mathrm{wt}(v)$ and also $u \xrightarrow{M} v$, which makes Theorem 2 applicable.

Theorem 6 *Let M be a linear map given by a linear matrix $M \in \mathbb{F}_2^{m \times n}$, $u \in \mathbb{F}_2^n, v \in \mathbb{F}_2^m$. Then, $u \xrightarrow[core]{M} v$ if and only if $u \xrightarrow[min]{M} v$ if and only if $M_{v,u}$ is an invertible square matrix, where $M_{v,u}$ denotes the submatrix of M with rows selected by the vector v and columns selected by the vector u.*

5 Upper Bounds on the Degree of a Composition of Functions

A classic problem in cryptanalysis of symmetric-key primitives is determining the algebraic degree of a block cipher (more precisely, it's maximum over all possible keys). Low algebraic degree can be used in the higher-order differential attack, introduced by Knudsen [1]. Most, if not all, block ciphers are built from an iterative structure - a composition of simple round functions. Therefore, the problem reduces to finding upper bounds on the algebraic degree of a composition of functions.

In this section, we briefly recall classic upper bounds on the algebraic degree and describe their relations with variants of division property. While bit-based division property is much more fine-grained than a degree upper bound, computing its propagation is generally a hard problem. Degree-based bounds, on the other hand, can be usually derived using pen-and-paper, allowing quick degree estimations.

Chen, Xiang, Zeng and Zhang [15] studied relationships between the bit-/word-/state-based division property (perfect and conventional), the naive and the Boura-Canteaut degree bounds, concluding that division property is superior over those bounds. Udovenko [16] further studied relationships with the recent bounds by Carlet.

5.1 Classic Methods

The most generic and straightforward upper bound is obtained by simply multiplying the degrees of composed functions. It can not be improved in general without having additional information about the functions.

Proposition 22 *Let $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m, g\colon \mathbb{F}_2^m \to \mathbb{F}_2$. Then,*

$$\deg g \circ F \leq \deg g \cdot \deg F.$$

The first nontrivial upper bound was obtained by Canteaut and Videau [31] under an additional constraint of divisibility of all Walsh coefficients by a power of 2.

Theorem 7 ([31, Cor.1]) *Let* $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m, g\colon \mathbb{F}_2^m \to \mathbb{F}_2$ *such that the Walsh spectrum of* F *is divisible by* 2^ℓ*, for a integer* $\ell \geq 1$*. Then,*

$$\deg G \circ F \leq n - \ell + \deg g.$$

Remark 14 The Walsh spectrum of $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$ is the multiset given by

$$\left\{ \sum_{x \in \mathbb{F}_2^n} (-1)^{\langle \alpha, F(x) \rangle + \langle \beta, x \rangle} \;\middle|\; \alpha \in \mathbb{F}_2^m \setminus \{\underline{0}\}, \beta \in \mathbb{F}_2^n \right\}.$$

Later, Boura and Canteaut [12], among other results, showed a new bound based on the *degree of the inverse* of one of the functions, which is a generalization of the result specific to SPN functions by Boura, Canteaut and De Cannière [32].

Theorem 8 ([12, Cor.3.2]) *Let* $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^n$ *be a bijection,* $g\colon \mathbb{F}_2^n \to \mathbb{F}_2$*. Then,*

$$\deg g \circ F \leq n - \left\lceil \frac{n - \deg g}{\deg F^{-1}} \right\rceil.$$

This bound is quite surprising as it shows that a block cipher may need nearly twice as many rounds (compared to the naive bound) to reach full degree, as the following example shows.

Example 8 To illustrate the idea, consider an SPN-based block cipher with $2r$ rounds, n-bit block and an m-bit S-box, $m \geq 3$. The degree of the S-box is at most $m-1 := d$. In the first half of the rounds, let us apply the naive degree bound (Proposition 22), which increases the degree bound by a factor d per round. In the second half of the rounds, let us apply the Boura-Canteaut bound, which decreases the degree "deficit" by a factor d per round. To get full degree $n-1$ (degree "deficit" 1) for the full $2r$-round permutation, we need $d^r + d^r \geq n$. Let r be the smallest number such that the condition holds, i.e., $2(r-1)$ rounds are not enough to reach full degree. However, using only the naive bound, the upper bound may reach full degree already after $r+1$ rounds, instead of $2r$: $d^{r+1} \geq 2d^r \geq n$.

More detailed analysis of this phenomenon in the framework of Feistel Networks was done in [33], and in the framework of SPNs in [13].

More recently, Carlet [14] showed several bounds based on the degrees of the involved graph indicators. We reproduce here one general theorem which is particularly relevant for division property.

Theorem 9 ([14, Theorems 1, 3, 4]) *Let* $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m, g\colon \mathbb{F}_2^m \to \mathbb{F}_2$. *Then,*

$$\deg g \circ F \leq \deg g + \deg \mathbb{1}_{\Gamma_F} - m.$$

More generally, let $G_1, \ldots, G_r$ *be such that* $G_i\colon \mathbb{F}_2^{m_{i-1}} \to \mathbb{F}_2^{m_i}$ *and let* $g : \mathbb{F}_2^{m_r} \to \mathbb{F}_2$. *Then,*

$$\deg g \circ G_r \circ \ldots \circ G_1 \leq \deg g + \sum_{i=1}^{r} (\deg \mathbb{1}_{\Gamma_{G_i}} - m_i).$$

Bounds from this theorem are not comparable to the previous bounds (in the common case of the composition of two functions). That is, for any of the classic bounds, there exists a case where it will be strictly stronger than the other bounds. A more complete study and comparison of classic bounds (excluding the division property) can be found in [14].

5.2 Formulation with Conventional Division Property

In the seminal paper [4], Todo provides a propagation rule of conventional division property through a vectorial Boolean function.

Proposition 23 ([4]) *Let* $X \subseteq \mathbb{F}_2^n$ *satisfy division property* D_k^n, *and let* $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$. *Then,* $F(X) \subseteq \mathbb{F}_2^m$ *(elements with odd-multiplicity) satisfies* $D_{k'}^m$ *with* $k' = \left\lceil \frac{k}{\deg F} \right\rceil$.

Since 1-dimensional division property corresponds to an upper bound on the algebraic degree of the set (i.e., we have $\deg X \leq n-k, \deg F(X) \leq m-k'$), we can conclude the following upper bound on the degree of the output set.

Corollary 4 *Let* $X \subseteq \mathbb{F}_2^n, F\colon \mathbb{F}_2^n \to \mathbb{F}_2^m$. *Then,*

$$\deg F(X) \leq m - \left\lceil \frac{n - \deg X}{\deg F} \right\rceil.$$

This bound clearly resembles the Boura-Canteaut bound (Theorem 8). However, the proposition relates the degrees of *sets*, rather than the degrees of *functions*. Furthermore, it has $\deg F$ instead of $\deg F^{-1}$ in the denominator and even does not require F to be invertible. In fact, Corollary 4 can be interpreted exactly as the Boura-Canteaut when F is invertible: let $F' = F^{-1}, g = \mathbb{1}_X$ so that $\mathbb{1}_{F(X)} = \mathbb{1}_X \circ F^{-1} = g \circ F'$.

The latter identity $\mathbb{1}_{F(X)} = \mathbb{1}_X \circ F^{-1}$ also implies the complementary analogue of the basic upper bound (Proposition 22): $\deg F(X) \leq \deg X \cdot \deg F^{-1}$, which can be translated into the division property terminology.

Corollary 5 ([16, Prop.3.22]) *Let* $X \subseteq \mathbb{F}_2^n$ *satisfy division property* D_k^n, *and let* $F\colon \mathbb{F}_2^n \to \mathbb{F}_2^n$ *be a bijection. Then,* $F(X) \subseteq \mathbb{F}_2^n$ *(elements with odd-multiplicity) satisfies* $D_{k'}^n$ *with*

$$k' \geq n - (n-k)\deg F^{-1}.$$

Remark 15 Todo also noticed that division property D_n^n is preserved by bijections (since the input and output sets have to be exactly $\mathbb{F}_2^n$). This corollary includes this case.

The two rules from the corollaries mirror the two classic degree-based upper bounds - the naive bound and the Boura-Canteaut bound. However, division property can be more precise: instead of computing k' solely from k and the degree of the function (or of its inverse), it can be computed from the DPPT of the function. More precisely, the value

$$\begin{aligned} D_F(k) &:= \min\{\mathrm{wt}(v) \mid u \xrightarrow{F} v, \mathrm{wt}(u) \geq k\} \\ &= \min\{\mathrm{wt}(v) \mid F^v(x) \text{ contains } x^u, \mathrm{wt}(u) \geq k\}. \end{aligned}$$

is optimal (maximum possible) for k', and the degree-based bounds do not always reach it (see Example 9 below). This parameter, as a function of F and k, was introduced in [28].

By Theorem 3, the value of $D_F(k)$ can be also characterized by

$$D_F(k) = \min\{\mathrm{wt}(v) \mid \text{the ANF of } \mathbb{1}_{\Gamma_F} \text{ contains } x^u y^{\neg v}, \mathrm{wt}(u) \geq k\}.$$

In other words, the 1-dimensional division property propagation table is fully characterized by the pairs of degrees (one per each variable) of maximal monomials of the graph indicator $\mathbb{1}_{\Gamma_F}(x, y)$ of the considered function F. This information is more detailed then simply the degree of the graph indicator, used in the Carlet's bound.

5.3 Conventional Division Trails and Indicator Monomial Trails

Udovenko [16] exhibited a close connection of the Carlet's graph indicator method with conventional division property propagation. In fact, the latter can be seen as a more fine-grained variant of the former, requiring more information about the composed function, more computational effort to compute the upper bound, but (possibly) resulting in a stronger bound.

The result relies on the following representation of the graph indicator of a composition as the sum of the products of the involved indicators over all possible values of all intermediate variables.

Proposition 24 ([14, 34]) *Let* $G_i \colon \mathbb{F}_2^{m_{i-1}} \to \mathbb{F}_2^{m_i}$, $i \in \{1, \ldots, r\}$, *let* $F = G_r \circ \ldots \circ G_1$. *Then,*

$$\mathbb{1}_{\Gamma_F}(x, z) = \sum_{\substack{(y_1, \ldots, y_{r-1}) \\ \in \mathbb{F}_2^{m_1} \times \ldots \times \mathbb{F}_2^{m_{r-1}}}} \mathbb{1}_{\Gamma_{G_1}}(x, y_1) \cdot \mathbb{1}_{\Gamma_{G_2}}(y_1, y_2) \cdot \ldots \cdot \mathbb{1}_{\Gamma_{G_r}}(y_{r-1}, z).$$

The following theorem shows that conventional division property trails essentially correspond to chains of monomials from the involved graph indicators.

Theorem 10 ([16, Thm.4]) *Let F, G_i be defined as above. Let I be the formal expansion (i.e., no +-cancellations) of the product*

$$\mathbb{1}_{\Gamma_{G_1}}(x, y_1) \cdot \mathbb{1}_{\Gamma_{G_2}}(y_1, y_2) \cdot \ldots \cdot \mathbb{1}_{\Gamma_{G_r}}(y_{r-1}, z).$$

Then, I contains a monomial multiple of

$$x^u y_1^{\underline{1}} y_2^{\underline{1}} \ldots y_{r-1}^{\underline{1}} z^{\neg v} \tag{4}$$

if and only if there exists a valid division trail

$$u \xrightarrow{G_1} w_1 \xrightarrow{G_2} \ldots \xrightarrow{G_{r-1}} w_{r-1} \xrightarrow{G_r} v. \tag{5}$$

Proof The proof relies on Theorem 3, applied to each link in the indicator chain and division trail:

$$u \xrightarrow{G_1} w_1 \quad \Leftrightarrow \quad \mathbb{1}_{\Gamma_{G_1}}(x, y_1) \text{ contains a monomial multiple of } x^u y_1^{\neg w_1}, \tag{6}$$

$$w_1 \xrightarrow{G_2} w_2 \quad \Leftrightarrow \quad \mathbb{1}_{\Gamma_{G_2}}(y_1, y_2) \text{ contains a monomial multiple of } y_1^{w_1} y_2^{\neg w_2},$$

$$\vdots$$

$$w_{r-1} \xrightarrow{G_r} v \quad \Leftrightarrow \quad \mathbb{1}_{\Gamma_{G_{r-1}}}(y_{r-1}, z) \text{ contains a monomial multiple of } y_{r-1}^{w_{r-1}} z^{\neg v} \tag{7}$$

The formal expansion I contains a monomial multiple of (4) if and only if there exists a chain of monomials $x^{u'} y_1^{w_1'}$, $y_1^{w_1} y_2^{w_2'}$, …, $y_{r-1}^{w_{r-1}} z^{\neg v'}$ from the ANFs of $\mathbb{1}_{\Gamma_{G_1}}(x, y_1)$, $\mathbb{1}_{\Gamma_{G_2}}(y_1, y_2)$, …, $\mathbb{1}_{\Gamma_{G_r}}(y_{r-1}, z)$ respectively such that these monomials multiply to (4). In other words, the conditions are $u' \succeq u, \neg v' \succeq \neg v$, and

$$w_1' \vee w_1 = \underline{1} \in \mathbb{F}_2^{m_1},$$

$$\vdots$$

$$w_{r-1}' \vee w_{r-1} = \underline{1} \in \mathbb{F}_2^{m_r}.$$

Equivalently, $w_1' \succeq \neg w_1$, …, $w_{r-1}' \succeq \neg w_{r-1}$. In other words, $\mathbb{1}_{\Gamma_{G_1}}(x, y_1)$ contains a monomial multiple of $x^u y_1^{\neg w_1}$, …, $\mathbb{1}_{\Gamma_{G_{r-1}}}(y_{r-1}, z)$ contains a monomial multiple of $y_{r-1}^{w_{r-1}} z^{\neg v}$. By Theorem 3 (see (6)-(7) above), these are equivalent to the validity of the division trail (5).

□

In order to relate conventional division property propagation to Carlet's Theorem 9, we will show that the latter can be proven in terms of the formal expansion of the graph indicator product, i.e., that it does not rely on any +-cancellations. We recall the statement.

Theorem 9. Let $G_1, \ldots, G_r$ be such that $G_i \colon \mathbb{F}_2^{m_{i-1}} \to \mathbb{F}_2^{m_i}$ and let $g : \mathbb{F}_2^{m_r} \to \mathbb{F}_2$. Then,

$$\deg g \circ G_r \circ \ldots \circ G_1 \leq \deg g + \sum_{i=1}^{r} (\deg \mathbb{1}_{\Gamma_{G_i}} - m_i).$$

Proof Let $F \coloneqq g \circ G_r \circ \ldots \circ G_1 : \mathbb{F}_2^n \to \mathbb{F}_2$. We will use the fact that the indicator of a Boolean function $y = f(x)$ is simply $y + f(x)$. We have

$$\deg F \leq \deg \mathbb{1}_{\Gamma_F} \leq d,$$

where d is the maximum integer such that the monomial $x^d y_1^{\frac{1}{1}} \ldots y_r^{\frac{1}{r}}$ belongs to the formal expansion

$$I \coloneqq \mathbb{1}_{\Gamma_{G_1}}(x, y_1) \cdot \mathbb{1}_{\Gamma_{G_2}}(y_1, y_2) \cdot \ldots \cdot \mathbb{1}_{\Gamma_{G_r}}(y_{r-1}, y_r) \cdot \mathbb{1}_{\Gamma_g}(y_r, z).$$

Here, we used the expression of $\mathbb{1}_{\Gamma_F}(x, z)$ from Proposition 24. It follows that

$$\deg x^d y_1^{\frac{1}{1}} \ldots y_r^{\frac{1}{r}} = d + \sum_{i=1}^{r} m_r \leq \sum_{i=1}^{r} \deg \mathbb{1}_{\Gamma_{G_i}} + \deg_{\mathbb{1}_{\Gamma_g}}.$$

Note that $\deg_{\mathbb{1}_{\Gamma_g}} = \deg g$ for non-constant g. For constant g the theorem follows from the fact that the sum $\sum_{i=1}^{r} \left(\deg \mathbb{1}_{\Gamma_{G_i}} - m_i\right)$ is always non-negative. We conclude that

$$\deg F \leq \sum_{i=1}^{r} \left(\deg \mathbb{1}_{\Gamma_{G_i}} - m_i\right) + \deg g.$$

□

We conclude that bit-based conventional division property is never worse than Carlet's bound (Theorem 9). Furthermore, we will now show that already 1-dimensional (state-based) table-based division property propagation is never worse than Carlet's bound and, in fact, is often better (at the cost of requiring more information about the graph indicator's ANF and being computational).

Proposition 25 *Bounds by conventional state-level division property with table-based propagation (i.e., based on D_F) are never worse than bounds from Theorem 9.*

Proof Let $G_1, \ldots, G_r, g$ be defined as in Theorem 9, $f = g \circ G_r \circ \ldots \circ G_1$. Theorem 9 states that

$$\deg_f \leq d \coloneqq \deg g + \sum_{i=1}^{r} (\deg \mathbb{1}_{\Gamma_{G_i}} - m_i).$$

We will prove that there doesn't exist a weak division trail (under the weight projection $\mathrm{wt}_1 = \mathrm{wt}$)

$$d + 1 \xrightarrow{G_1} w_1 \xrightarrow{G_2} \ldots \xrightarrow{G_r} w_r \xrightarrow{g} 1,$$

where $w_i \in [0, m_i]$ for all i. This will prove that the division property implies the bound $\deg_f \leq d$.

The proof is by contradiction. Assume that such a weight trail exists. Note that it does not imply existence of the full binary trail, only separate trails for each of its steps:

$$\begin{aligned} u \xrightarrow{G_1} v_1 \quad &\text{with} \quad \mathrm{wt}(u) = d+1 \text{ and } \mathrm{wt}(v_1) = w_1, \\ v_1' \xrightarrow{G_2} v_2 \quad &\text{with} \quad \mathrm{wt}(v_1') = w_1 \text{ and } \mathrm{wt}(v_2) = w_2, \\ &\vdots \\ v_{r-1}' \xrightarrow{G_2} v_r \quad &\text{with} \quad \mathrm{wt}(v_{r-1}') = w_{r-1} \text{ and } \mathrm{wt}(v_r) = v_1, \\ v_r' \xrightarrow{g} 1 \quad &\text{with} \quad \mathrm{wt}(v_r') = w_r, \end{aligned}$$

where $u \in \mathbb{F}_2^n, v_i \in \mathbb{F}_2^{m_i}$ for all i. By Theorem 3, each such transition implies a particular monomial (multiple) in the respective graph indicator's ANF, implying a degree bound:

$$\begin{aligned} d+1+(m_1-w_1) &\le \deg \mathbb{1}_{\Gamma_{G_1}}, \\ w_1+(m_2-w_2) &\le \deg \mathbb{1}_{\Gamma_{G_2}}, \\ &\vdots \\ w_{r-1}+(m_r-w_r) &\le \deg \mathbb{1}_{\Gamma_{G_r}}, \\ w_r &\le \deg \mathbb{1}_{\Gamma_g}. \end{aligned}$$

Summing the bounds leads to

$$d+1+\sum_{i=1}^{r} m_i \le \sum_{i=1}^{r} \deg \mathbb{1}_{\Gamma_{G_i}} + \deg_{\mathbb{1}_{\Gamma_g}}.$$

By substituting the definition of d, it must be

$$\deg g + 1 \le \deg_{\mathbb{1}_{\Gamma_g}},$$

which implies that g is constant. However, for constant g the only weak transition to 1 is $0 \xrightarrow{g} 1$, and the only weak transitions to 0 are $0 \xrightarrow{G_i} 0$, implying $d+1 = w_1 = \ldots = w_r = 0$, which is not possible due to $d \ge 0$ (defining the degree of the constant zero function to be $-\infty$ causes contradiction as well). We obtain a contradiction with existence of such a weight trail.

□

5.4 Exposition of Compositional Bounds through Bounds on Monomials of the Graph Indicator

In this section, we will show that Carlet's graph indicator method allows illustrative comparison of classic compositional bounds and 1-dimensional conventional division property. A key idea is to derive bounds on degrees of monomials in the graph indicator directly from a given classic bound. The motivation comes from the fact that possible degrees of monomials in the graph indicator's ANF define 1-dimensional division property propagation. This allows comparison of bounds in the same setting.

The obtained bounds can then be illustrated graphically on an example function, exhibiting visually the impossible degree pairs that are or are

not removed by each of the bounds. We will focus on the classic bounds for compositions of two functions.

Let $F : \mathbb{F}_2^n \to \mathbb{F}_2^m$. Let $x^u y^v$ be a maximal monomial in $\mathbb{1}_{\Gamma_F}(x, y)$. Define

$$g : \mathbb{F}_2^m \to \mathbb{F}_2 : y \mapsto y^{\neg v}.$$

Note that $\deg g = m - \mathrm{wt}(v)$. Recall that

$$z + g(F(x)) + 1 = \mathbb{1}_{\Gamma_{g \circ F}}(x, z) = \\ = \sum_{y \in \mathbb{F}_2^m} \mathbb{1}_{\Gamma_F}(x, y) \cdot \mathbb{1}_{\Gamma g}(y, z) = \sum_{y \in \mathbb{F}_2^m} \mathbb{1}_{\Gamma_F}(x, y) \cdot (z + y^{\neg v} + 1).$$

Observe that the monomial $x^u y^v y^{\neg v} = x^u y^{\underline{1}}$ belongs to the polynomial $\mathbb{1}_{\Gamma_F}(x, y) \cdot \mathbb{1}_{\Gamma g}(y, z)$ as it occurs exactly once in the formal expansion (due to the maximality of $x^u y^v$). It follows that x^u belongs to $g(F(x))$ or is constant (if $u = \underline{0}$). Hence, $\mathrm{wt}(u) \leq \deg g \circ F$. We can now apply classic bounds to the composition $g \circ F$ and obtain bounds on degrees of the monomial $x^u y^v$ from the graph indicator of F.

1. By the naive bound (Proposition 22), we have

$$\mathrm{wt}(u) \;\leq\; \deg g \circ F \;\leq\; \deg g \cdot \deg F \;=\; (m - \mathrm{wt}(v)) \cdot \deg F.$$

Therefore,

$$\mathrm{wt}(u) + \mathrm{wt}(v) \cdot \deg F \leq m \cdot \deg F. \tag{8}$$

2. By the Boura-Canteaut bound (Theorem 8), we have

$$\mathrm{wt}(u) \;\leq\; \deg g \circ F \;\leq\; m - \frac{m - \deg g}{\deg F^{-1}} \;=\; m - \frac{\mathrm{wt}(v)}{\deg F^{-1}}.$$

Therefore,

$$\mathrm{wt}(u) \cdot \deg F^{-1} + \mathrm{wt}(v) \leq m \cdot \deg F^{-1}. \tag{9}$$

3. By the Carlet bound (Theorem 9), we have

$$\mathrm{wt}(u) \;\leq\; \deg g \circ F \;\leq\; \deg \mathbb{1}_{\Gamma_F} - m + \deg f \;=\; \deg \mathbb{1}_{\Gamma_F} - \mathrm{wt}(v).$$

Therefore,

$$\mathrm{wt}(u) + \mathrm{wt}(v) \leq \deg \mathbb{1}_{\Gamma_F}. \tag{10}$$

Remark 16 The third bound is trivial by definition. What is important here is that it is equivalent to the Carlet's bound. Similarly, the second bound can be obtained from the first one applied directly to the inverse of F. However, it is insightful to derive it directly from the Boura-Canteaut bound.

Note that by the monotonicity of bounds and by the fact that $x^u y^v$ was chosen as a maximal monomial in the graph indicator's ANF, these bounds ((8), (9), (10)) hold for all monomials $x^u y^v$ of $\mathbb{1}_{\Gamma_F}(x, y)$ in general.

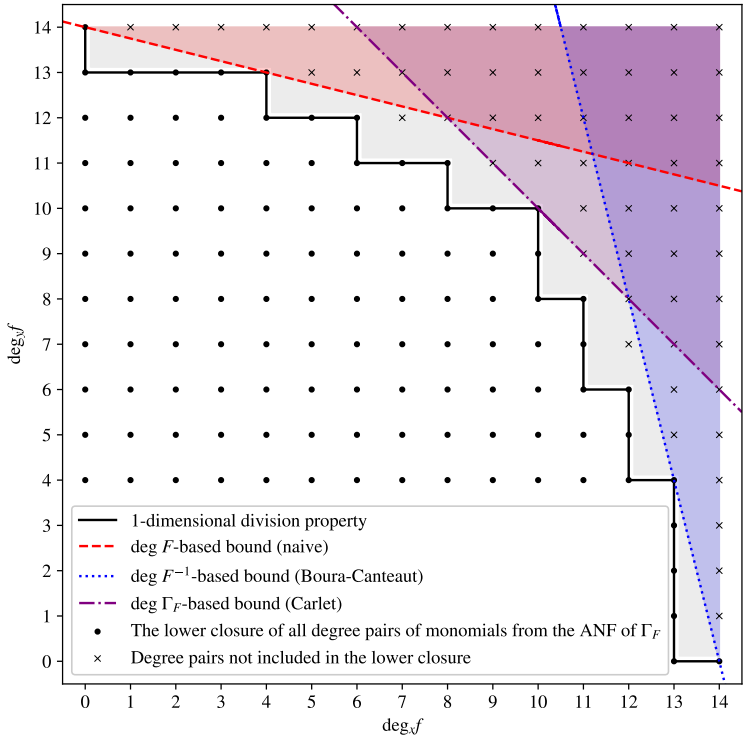


Fig. 3 Comparison of bounds on degrees on x and y of monomials $f(x, y) = x^u y^v$ of the graph indicator $\mathbb{1}_{\Gamma_F}(x, y)$, where $F : \mathbb{F}_2^{14} \to \mathbb{F}_2^{14}$ is defined as $F : (\mathbb{F}_{2^7})^2 \to (\mathbb{F}_{2^7})^2 : (x_1, x_2) \mapsto (x_1^3, x_2^{1/3})$, so that $\deg F = \deg F^{-1} = 4, \deg \mathbb{1}_{\Gamma_F} = 20$. Shaded areas highlight the points not satisfying the respective bounds (exclusive of the respective lines themselves).

Example 9 We will now illustrate these bounds on a simple example. Let $F : \mathbb{F}_2^{14} \to \mathbb{F}_2^{14}$ be defined as

$$F : (\mathbb{F}_{2^7})^2 \to (\mathbb{F}_{2^7})^2 : (x_1, x_2) \mapsto (x_1^3, x_2^{1/3}),$$

so that $\deg F = \deg F^{-1} = 4$, $\deg \mathbb{1}_{\Gamma_F} = 20$.

The three bounds together with the actual maximal degree pairs are displayed graphically on Figure 3. The black dots correspond to the set of pairs

$$\downarrow\left\{(\deg_x f, \deg_y f) \in [0,14]^2 \mid f \text{ a monomial in } \mathbb{1}_{\Gamma_F}(x,y)\right\} = \underset{14,14}{\mathrm{wt}}(\downarrow\mathcal{A}(\Gamma_F)),$$

i.e., to pairs of degrees on x and y of monomials having a monomial multiple present in the ANF of the graph indicator of F. The black crosses correspond to pairs of degrees outside of this set. The red dashed line corresponds to the bound (8) derived from the naive degree bound. The blue dotted line corresponds to the bound (9) derived from the Boura-Canteaut degree bound. The purple dash-dot line corresponds to the bound (10) derived from the Carlet's degree bound. The shaded areas above the lines (exclusive of the respective lines) cover points not satisfying the bounds.

Note that the plot is symmetric along the $\deg_y f = \deg_x f$ line, due to the construction of F implying $\deg F = \deg F^{-1}$. This makes the naive and the Boura-Canteaut bounds fully symmetric (but not equivalent), which is not the case in general.

The figure shows that each bound is tight, i.e., that they all reach equality at some points, and that none of them are redundant. Furthermore, even all three bounds jointly do not precisely model the right set of degree pairs: the points with coordinates $(7,12)$ and $(12,7)$ do not belong to the set $\downarrow \mathrm{wt}_{14,14}(\mathcal{A}(\Gamma_F))$, but are not rejected by any of the three bounds.

In Section 5.2 it was stated that degree-based bounds for 1-dimensional (state-based) conventional division property are not always precise. The points $(7,12)$ and $(12,7)$ mentioned above prove this statement. Indeed, the point $(7,12)$ implies $D_F(7) = 14 - (12-1) = 3$. At the same time, the bounds from Proposition 23 and Corollary 4 state respectively that $D_F(7) \geq \left\lceil \frac{7}{\deg F} \right\rceil = 2$ and $D_F(7) \geq n - (n-7)\deg F^{-1} = -14$.

Finally, we note that even 1-dimensional (state-based) conventional division property propagation (using D_F rather than degree-based bounds) utilizes full knowledge of the degrees of the maximal monomials in the graph indicator's ANF (by Theorem 3). Hence, it is illustrated as a precise boundary (black solid line) on the figure.

6 Security Arguments

The perfect division property can be used to provide security arguments for ciphers against attacks. One important attack vector is higher-order differential cryptanalysis [35], which exploits a low algebraic degree of a cipher or, more generally, a missing monomial in the ANF. While the standard definition of the algebraic degree of a vectorial Boolean function considers the *maximum* degree over all its *coordinates*, the security of a cipher depends on the *minimum* degree over all possible (nonzero) *linear combinations* of the coordinates. Indeed, an adversary can choose any linear combination of output bits they want to attack.

Definition 14 (Degree of a Block Cipher [36]) Let $E\colon \mathbb{F}_2^n \times \mathbb{F}_2^\ell \to \mathbb{F}_2^n$ be a block cipher, where the ANF can be written as $E_k(x) = \sum_{u\in\mathbb{F}_2^n} p_u(k)x^u$. We define the

algebraic degree of E by

$$\deg(E) \coloneqq \max_{k\in\mathbb{F}_2^\ell} \deg(E_k) = \max_{u\in\mathbb{F}_2^n, p_u(k)\neq 0} \mathrm{wt}(u).$$

Furthermore, we define the *minimum degree* of E by

$$\mathrm{minDeg}(E) \coloneqq \min_{\beta\in\mathbb{F}_2^n,\beta\neq\underline{0}} \max_{k\in\mathbb{F}_2^\ell} \langle\beta, E_k\rangle = \min_{\beta\in\mathbb{F}_2^n} \max_{u\in\mathbb{F}_2^n,\langle\beta,p_u(k)\rangle\neq 0} \mathrm{wt}(u).$$

Let $E\colon \mathbb{F}_2^n \times \mathbb{F}_2^\ell \to \mathbb{F}_2^n$ be a block cipher. We can write the ANF of E_k as

$$E_k(x) = \sum_{u\in\mathbb{F}_2^n} p_u(k) x^u,$$

where $p_u\colon \mathbb{F}_2^\ell \to \mathbb{F}_2^n$ for all $u \in \mathbb{F}_2^n$ and

$$p_u(k) = \sum_{v\in\mathbb{F}_2^\ell} \lambda_{u,v} k^v,\ \lambda_{u,v} \in \mathbb{F}_2^n.$$

When the input division property (u, v) propagates to e_i for E, i.e.,

$$(u, v) \xrightarrow{E} e_i,$$

this means that $\lambda_{u,v}^{(i)} = 1$ and $p_u(k) \neq 0$. We refer to u as the input monomial and v as the key pattern. The propagation above implies that the monomial x^u occurs in the ANF of $E_k^{(i)}$ for at least one key k.

To show that E has an algebraic degree of at least d, it is sufficient to find a $u \in \mathbb{F}_2^n$ where $\mathrm{wt}(u) \geq d$, a key pattern $v \in \mathbb{F}_2^\ell$ and an output bit $i \in \{1, \ldots, n\}$, so that

$$(u, v) \xrightarrow{E} e_i.$$

In the case of the minimum degree, we have to find n pairs of input monomials and key patterns $(u_1, v_1), \ldots, (u_n, v_n)$ and compute the corresponding entries in the ANF for all output bits. When the matrix

$$\begin{pmatrix} \lambda_{u_1,v_1}^{(1)} & \lambda_{u_2,v_2}^{(1)} & \cdots & \lambda_{u_n,v_n}^{(1)} \\ \lambda_{u_1,v_1}^{(2)} & \lambda_{u_2,v_2}^{(2)} & \cdots & \lambda_{u_n,v_n}^{(2)} \\ & \vdots & & \\ \lambda_{u_1,v_1}^{(n)} & \lambda_{u_2,v_2}^{(n)} & \cdots & \lambda_{u_n,v_n}^{(n)} \end{pmatrix}$$

has rank n, the minimum degree of E is at least $\min_{i\in\{1,\ldots,n\}} \mathrm{wt}(u_i)$.

The integral attacks can be seen as a generalization of higher-order differentials. While for the latter one, an adversary sums over

$$\sum_{x\preceq u} \langle\beta, E_k(x)\rangle$$

where the weight of u is larger then the algebraic degree of E, the adversary sums over an arbitrary set $S \subset \mathbb{F}_2^n$:

$$\sum_{x \in S} \langle \beta, E_k(x) \rangle.$$

This can be considered as a linear combination of coefficients of the ANF of $\langle \beta, E_k(x) \rangle = \sum_{u \in \mathbb{F}_2^n} \lambda_u x^u$:

$$\begin{aligned}
\sum_{x \in S} \langle \beta, E_k(x) \rangle &= \sum_{x \in S} \sum_{u \in \mathbb{F}_2^n} \lambda_u x^u \\
&= \sum_{v \in \mathcal{U}(S)} \sum_{x \preceq v} \sum_{u \preceq x} \lambda_u \\
&= \sum_{v \in \mathcal{U}(S)} \sum_{u \in \mathbb{F}_2^n} |\{x \in \mathbb{F}_2^n \mid u \preceq x \preceq v\}| \cdot \lambda_u \\
&= \sum_{v \in \mathcal{U}(S)} \lambda_v.
\end{aligned}$$

When this sum is always key-dependent, i.e. there does not exists an integral attack on a certain block cipher, we say that the cipher fulfills the integral-resistance property.

Definition 15 (Integral-Resistance Property [36]) Let $E \colon \mathbb{F}_2^n \times \mathbb{F}_2^\ell \to \mathbb{F}_2^n$ be a block cipher. We say that E fulfills the *integral-resistance property* when for every proper non-empty subset $M \subset \mathbb{F}_2^n$ and every non-zero output mask $\beta \in \mathbb{F}_2^n$ the sum

$$\sum_{x \in M} \langle \beta, E_k(x) \rangle$$

is key-dependent.

Definition 16 (Integral-Resistance Matrix [36]) Let $E \colon \mathbb{F}_2^n \times \mathbb{F}_2^\ell \to \mathbb{F}_2^n$ be a block cipher with the corresponding ANF

$$E_k(x) = \sum_{u \in \mathbb{F}_2^n} \sum_{v \in \mathbb{F}_2^\ell} \lambda_{u,v} x^u k^v.$$

Further, let $v_1, \dots v_s \in \mathbb{F}_2^\ell$ be key patterns. We call the following matrix over $\mathbb{F}_2^{n^2 \times s}$

$$\mathcal{I}(E) = \begin{pmatrix} \lambda^{(1)}_{\neg e_1, v_1} & \lambda^{(1)}_{\neg e_1, v_2} & \cdots & \lambda^{(1)}_{\neg e_1, v_s} \\ \lambda^{(2)}_{\neg e_1, v_1} & \lambda^{(2)}_{\neg e_1, v_2} & \cdots & \lambda^{(2)}_{\neg e_1, v_s} \\ & \vdots & & \\ \lambda^{(n)}_{\neg e_1, v_1} & \lambda^{(n)}_{\neg e_1, v_2} & \cdots & \lambda^{(n)}_{\neg e_1, v_s} \\ \lambda^{(1)}_{\neg e_2, v_1} & \lambda^{(1)}_{\neg e_2, v_2} & \cdots & \lambda^{(1)}_{\neg e_2, v_s} \\ \lambda^{(2)}_{\neg e_2, v_1} & \lambda^{(2)}_{\neg e_2, v_2} & \cdots & \lambda^{(2)}_{\neg e_2, v_s} \\ & \vdots & & \\ \lambda^{(j)}_{\neg e_i, v_1} & \lambda^{(j)}_{\neg e_i, v_2} & \cdots & \lambda^{(j)}_{\neg e_i, v_s} \\ & \vdots & & \\ \lambda^{(n-1)}_{\neg e_n, v_1} & \lambda^{(n-1)}_{\neg e_n, v_2} & \cdots & \lambda^{(n-1)}_{\neg e_n, v_s} \\ \lambda^{(n)}_{\neg e_n, v_1} & \lambda^{(n)}_{\neg e_n, v_2} & \cdots & \lambda^{(n)}_{\neg e_n, v_s} \end{pmatrix}$$

an *integral-resistance matrix*. A column i corresponds to a key pattern v_i and each row corresponds to a combination of an output bit and a monomial of weight $n-1$.

Now, the integral-resistance matrix is sufficient to prove the integral-resistance property for a block cipher, assuming an independent whitening key added at the input (see also Proposition 18).

Proposition 26 *[36] Let $E\colon \mathbb{F}_2^n \times \mathbb{F}_2^\ell \to \mathbb{F}_2^n$ be a block cipher and $\mathcal{I}(E)$ be a corresponding integral-resistance matrix. If $\mathcal{I}(E)$ has rank n^2 and k_0 is an independent whitening key, $E_k(x + k_0)$ fulfills the integral-resistance property.*

7 Conclusions

In this survey, we explained the state-of-the-art variants of the division property, explained the connection with the ANF and how the division property can be used to find both attacks and security arguments for symmetric cryptographic primitives. Our focus was on the underlying theory and on a clear and precise notation. We hope that in particular readers with a background in Boolean functions find our survey helpful.

We also like to note that there are several important topics that we did not cover. Those topics are mainly concerned with an efficient computation of the division property using either dedicated algorithms [4, 7, 30, 37] or general tools, in particular mixed integer linear programming tools [8, 10, 17, 38–43], SAT/SMT solvers [16, 27, 44–48]. It is actually this practical computational aspect that is of great importance for all variants of the division property. As described, the security of symmetric primitives is the concrete security against concrete attacks. The division property is a powerful set of techniques to compute concrete properties of those ciphers.

Acknowledgements

The authors thank Claude Carlet for proposing the idea of this survey and the anonymous reviewers for their valuable suggestions which helped to greatly improve this work. The work was supported by the Luxembourg National Research Fund's (FNR) and the German Research Foundation's (DFG) joint project APLICA (C19/IS/13641232).

References

[1] Knudsen, L.R.: Truncated and higher order differentials. In: Preneel, B. (ed.) FSE'94. LNCS, vol. 1008, pp. 196–211. Springer (1995). https://doi.org/10.1007/3-540-60590-8_16

[2] Knudsen, L.R., Wagner, D.: Integral cryptanalysis. In: Daemen, J., Rijmen, V. (eds.) FSE 2002. LNCS, vol. 2365, pp. 112–127. Springer (2002). https://doi.org/10.1007/3-540-45661-9_9

[3] Dinur, I., Shamir, A.: Cube attacks on tweakable black box polynomials. In: Joux, A. (ed.) EUROCRYPT 2009. LNCS, vol. 5479, pp. 278–299. Springer (2009). https://doi.org/10.1007/978-3-642-01001-9_16

[4] Todo, Y.: Structural evaluation by generalized integral property. In: Oswald, E., Fischlin, M. (eds.) EUROCRYPT 2015, Part I. LNCS, vol. 9056, pp. 287–314. Springer (2015). https://doi.org/10.1007/978-3-662-46800-5_12

[5] Todo, Y.: Integral cryptanalysis on full MISTY1. In: Gennaro, R., Robshaw, M.J.B. (eds.) CRYPTO 2015, Part I. LNCS, vol. 9215, pp. 413–432. Springer (2015). https://doi.org/10.1007/978-3-662-47989-6_20

[6] Todo, Y.: Integral cryptanalysis on full MISTY1. Journal of Cryptology **30**(3), 920–959 (2017). https://doi.org/10.1007/s00145-016-9240-x

[7] Todo, Y., Morii, M.: Bit-based division property and application to simon family. In: Peyrin, T. (ed.) FSE 2016. LNCS, vol. 9783, pp. 357–377. Springer (2016). https://doi.org/10.1007/978-3-662-52993-5_18

[8] Hao, Y., Leander, G., Meier, W., Todo, Y., Wang, Q.: Modeling for three-subset division property without unknown subset - improved cube attacks against Trivium and Grain-128AEAD. In: Canteaut, A., Ishai, Y. (eds.) EUROCRYPT 2020, Part I. LNCS, vol. 12105, pp. 466–495. Springer (2020). https://doi.org/10.1007/978-3-030-45721-1_17

[9] Hao, Y., Leander, G., Meier, W., Todo, Y., Wang, Q.: Modeling for three-subset division property without unknown subset. Journal of Cryptology **34**(3), 22 (2021). https://doi.org/10.1007/s00145-021-09383-2

[10] Xiang, Z., Zhang, W., Bao, Z., Lin, D.: Applying MILP method to searching integral distinguishers based on division property for 6 lightweight block ciphers. In: Cheon, J.H., Takagi, T. (eds.) ASIACRYPT 2016, Part I. LNCS, vol. 10031, pp. 648–678. Springer (2016). https://doi.org/10.1007/978-3-662-53887-6_24

[11] Boura, C., Canteaut, A.: Another view of the division property. In: Robshaw, M., Katz, J. (eds.) CRYPTO 2016, Part I. LNCS, vol. 9814, pp. 654–682. Springer (2016). https://doi.org/10.1007/978-3-662-53018-4_24

[12] Boura, C., Canteaut, A.: On the Influence of the Algebraic Degree of F^{-1} on the Algebraic Degree of G ∘ F. IEEE Transactions on Information Theory **59**(1), 691–702 (2013)

[13] Biryukov, A., Khovratovich, D., Perrin, L.: Multiset-algebraic cryptanalysis of reduced Kuznyechik, Khazad, and secret SPNs. IACR Trans. Symm. Cryptol. **2016**(2), 226–247 (2016). https://doi.org/10.13154/tosc.v2016.i2.226-247. https://tosc.iacr.org/index.php/ToSC/article/view/572

[14] Carlet, C.: Graph indicators of vectorial functions and bounds on the algebraic degree of composite functions. IEEE Transactions on Information Theory, 1–1 (2020). https://doi.org/10.1109/TIT.2020.3017494

[15] Chen, S., Xiang, Z., Zeng, X., Zhang, S.: On the relationships between different methods for degree evaluation. IACR Trans. Symm. Cryptol. **2021**(1), 411–442 (2021). https://doi.org/10.46586/tosc.v2021.i1.411-442

[16] Udovenko, A.: Convexity of division property transitions: Theory, algorithms and compact models. In: Tibouchi, M., Wang, H. (eds.) ASIACRYPT 2021, Part I. LNCS, vol. 13090, pp. 332–361. Springer (2021). https://doi.org/10.1007/978-3-030-92062-3_12

[17] 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.) ASIACRYPT 2020, Part I. LNCS, vol. 12491, pp. 446–476. Springer (2020). https://doi.org/10.1007/978-3-030-64837-4_15

[18] Carlet, C.: Boolean Functions for Cryptography and Coding Theory. Cambridge University Press (2021). https://doi.org/10.1017/9781108606806

[19] Jean, J.: TikZ for Cryptographers. https://www.iacr.org/authors/tikz/ (2016)

[20] Shannon, C.: A Mathematical Theory of Cryptography, (1945)

[21] Ferguson, N., Kelsey, J., Lucks, S., Schneier, B., Stay, M., Wagner, D., Whiting, D.: Improved cryptanalysis of Rijndael. In: Schneier, B. (ed.) FSE 2000. LNCS, vol. 1978, pp. 213–230. Springer (2001). https://doi.org/10.1007/3-540-44706-7_15

[22] Todo, Y., Aoki, K.: Fast fourier transform key recovery for integral attacks. IEICE Trans. Fundam. Electron. Commun. Comput. Sci. **98-A**(9), 1944–1952 (2015)

[23] Ye, C.-D., Tian, T.: Revisit division property based cube attacks: Key-recovery or distinguishing attacks? IACR Trans. Symm. Cryptol. **2019**(3), 81–102 (2019). https://doi.org/10.13154/tosc.v2019.i3.81-102

[24] Hebborn, P., Lambin, B., Leander, G., Todo, Y.: Lower bounds on the degree of block ciphers. In: Moriai, S., Wang, H. (eds.) ASIACRYPT 2020, Part I. LNCS, vol. 12491, pp. 537–566. Springer (2020). https://doi.org/10.1007/978-3-030-64837-4_18

[25] Hebborn, P.: Security assessment and design of symmetric ciphers for emerging applications. doctoralthesis, Ruhr-Universität Bochum, Universitätsbibliothek (2022). https://doi.org/10.13154/294-8588

[26] Zhang, W., Rijmen, V.: Division cryptanalysis of block ciphers with a binary diffusion layer. IET Information Security **13**(2), 87–95 (2018)

[27] Hu, K., Wang, Q., Wang, M.: IACR transactions class documentation. IACR Trans. Symm. Cryptol. **2020**(1), 396–424 (2020). https://doi.org/10.13154/tosc.v2020.i1.396-424

[28] Göloğlu, F., Rijmen, V., Wang, Q.: On the division property of S-boxes. Cryptology ePrint Archive, Report 2016/188. https://eprint.iacr.org/2016/188 (2016)

[29] Lambin, B., Derbez, P., Fouque, P.-A.: Linearly equivalent S-boxes and the division property. Designs, Codes and Cryptography **88**(10), 2207–2231 (2020)

[30] Derbez, P., Fouque, P.-A.: Increasing precision of division property. IACR Trans. Symm. Cryptol. **2020**(4), 173–194 (2020). https://doi.org/10.46586/tosc.v2020.i4.173-194

[31] Canteaut, A., Videau, M.: Degree of composition of highly nonlinear functions and applications to higher order differential cryptanalysis. In: Knudsen, L.R. (ed.) EUROCRYPT 2002. LNCS, vol. 2332, pp. 518–533. Springer (2002). https://doi.org/10.1007/3-540-46035-7_34

[32] Boura, C., Canteaut, A., De Cannière, C.: Higher-order differential properties of Keccak and Luffa. In: Joux, A. (ed.) FSE 2011. LNCS, vol. 6733, pp. 252–269. Springer (2011). https://doi.org/10.1007/978-3-642-21702-9_15

[33] Perrin, L., Udovenko, A.: Algebraic insights into the secret Feistel network. In: Peyrin, T. (ed.) FSE 2016. LNCS, vol. 9783, pp. 378–398. Springer (2016). https://doi.org/10.1007/978-3-662-52993-5_19

[34] Carlet, C.: Handling vectorial functions by means of their graph indicators. IEEE Transactions on Information Theory **66**(10), 6324–6339 (2020). https://doi.org/10.1109/TIT.2020.2981524

[35] Lai, X.: Higher order derivatives and differential cryptanalysis. In: Communications and Cryptography. The Springer International Series in Engineering and Computer Science, vol. 276, pp. 227–233. Springer (1994)

[36] Hebborn, P., Lambin, B., Leander, G., Todo, Y.: Strong and tight security guarantees against integral distinguishers. In: Tibouchi, M., Wang, H. (eds.) ASIACRYPT 2021, Part I. LNCS, vol. 13090, pp. 362–391. Springer (2021). https://doi.org/10.1007/978-3-030-92062-3_13

[37] Todo, Y., Morii, M.: Compact representation for division property. In: Foresti, S., Persiano, G. (eds.) CANS 16. LNCS, vol. 10052, pp. 19–35. Springer (2016). https://doi.org/10.1007/978-3-319-48965-0_2

[38] Todo, Y., Isobe, T., Hao, Y., Meier, W.: Cube attacks on non-blackbox polynomials based on division property. In: Katz, J., Shacham, H. (eds.) CRYPTO 2017, Part III. LNCS, vol. 10403, pp. 250–279. Springer (2017). https://doi.org/10.1007/978-3-319-63697-9_9

[39] Sasaki, Y., Todo, Y.: New algorithm for modeling s-box in MILP based differential and division trail search. In: SECITC. Lecture Notes in Computer Science, vol. 10543, pp. 150–165. Springer (2017)

[40] Wang, S., Hu, B., Guan, J., Zhang, K., Shi, T.: MILP-aided method of searching division property using three subsets and applications. In: Galbraith, S.D., Moriai, S. (eds.) ASIACRYPT 2019, Part III. LNCS, vol. 11923, pp. 398–427. Springer (2019). https://doi.org/10.1007/978-3-030-34618-8_14

[41] Sun, L., Wang, W., Wang, M.: MILP-aided bit-based division property for primitives with non-bit-permutation linear layers. IET Inf. Secur. **14**(1), 12–20 (2020)

[42] Delaune, S., Derbez, P., Gontier, A., Prud'homme, C.: A simpler model for recovering superpoly on trivium. In: AlTawy, R., Hülsing, A. (eds.)

SAC 2021. LNCS, vol. 13203, pp. 266–285. Springer (2022). https://doi.org/10.1007/978-3-030-99277-4_13

[43] ElSheikh, M., Youssef, A.M.: On MILP-based automatic search for bit-based division property for ciphers with (large) linear layers. In: Baek, J., Ruj, S. (eds.) ACISP 21. LNCS, vol. 13083, pp. 111–131. Springer (2021). https://doi.org/10.1007/978-3-030-90567-5_6

[44] Sun, L., Wang, W., Wang, M.: Automatic search of bit-based division property for ARX ciphers and word-based division property. In: Takagi, T., Peyrin, T. (eds.) ASIACRYPT 2017, Part I. LNCS, vol. 10624, pp. 128–157. Springer (2017). https://doi.org/10.1007/978-3-319-70694-8_5

[45] Han, Y., Li, Y., Wang, M.: Automatical method for searching integrals of ARX block cipher with division property using three subsets. In: Naccache, D., Xu, S., Qing, S., Samarati, P., Blanc, G., Lu, R., Zhang, Z., Meddahi, A. (eds.) ICICS 18. LNCS, vol. 11149, pp. 647–663. Springer (2018). https://doi.org/10.1007/978-3-030-01950-1_38

[46] Eskandari, Z., Kidmose, A.B., Kölbl, S., Tiessen, T.: Finding integral distinguishers with ease. In: Cid, C., Jacobson Jr:, M.J. (eds.) SAC 2018. LNCS, vol. 11349, pp. 115–138. Springer (2019). https://doi.org/10.1007/978-3-030-10970-7_6

[47] Hu, K., Wang, M.: Automatic search for a variant of division property using three subsets. In: Matsui, M. (ed.) CT-RSA 2019. LNCS, vol. 11405, pp. 412–432. Springer (2019). https://doi.org/10.1007/978-3-030-12612-4_21

[48] Ghosh, S., Dunkelman, O.: Automatic search for bit-based division property. In: LATINCRYPT. Lecture Notes in Computer Science, vol. 12912, pp. 254–274. Springer (2021)