On the Fast Algebraic Immunity of Majority Functions

Pierrick Méaux

ICTEAM/ELEN/Crypto Group, Université catholique de Louvain, Belgium
pierrick.meaux@uclouvain.be

Abstract. In different contexts such as filtered LFSR, Goldreich's PRG, and FLIP stream ciphers, the security of a cryptographic primitive mostly depends on the algebraic properties of one Boolean function. Since the Seventies, more and more efficient attacks have been exhibited in this context, related to more and more general algebraic properties, such as the degree, the algebraic immunity, and finally, the fast algebraic immunity. Once the properties to estimate the attack complexities are identified, it remains to determine the exact parameters of interesting families of functions with these properties. Then, these functions can be combined in secondary constructions to guarantee the good algebraic properties of a main function. In particular, the family of symmetric functions, and more precisely the subclass of majority functions, has been intensively studied in the area of cryptography, because of their practical advantages and good properties.
The degree of all these functions is known, and they have been proven to reach the optimal algebraic immunity, but still very few is known relatively to its fast algebraic immunity. For a function in $n = 2^m + j$ variables, an upper bound is known for all m and j, proving that these functions do not reach the optimal fast algebraic immunity. However, the exact fast algebraic immunity is known only for very few families indexed by j, where the parameter is exhibited for all members of the family since m is big enough. Recent works gave exact values for $j = 0$ and $j = 1$ (in the first case), and for $j = 2$ and $j = 3$ with $m \geq 2$ (in the second case). In this work, we determine the exact fast algebraic immunity for all possible values of j, for all member of the family assuming $m \geq 1 + \log_2(j+1)$.

Keywords: Boolean Functions, Fast Algebraic Attacks, Symmetric Functions, Majority Functions.

1 Introduction

1.1 Cryptographic Primitives with Security Determined by the Algebraic Properties of One Boolean Function

For some constructions, the security of a cryptographic primitive mostly depends on the algebraic properties of one Boolean function. This strong connection between security and algebraic properties happens when two conditions hold. First, the primitive has a simple structure where the non-linear part is provided by an unique Boolean function. Second, an adversary is only able to obtain a system of equations as the output of this function (non iterated). The typical example of such context relates to stream-cipher encryption schemes, more particularly to the family of designs called (combined)

filtered Linear Feedback Shift Register (LFSR) (*e.g.* [32]). This design is characterized by the combination of linear components applying on the secret key, which output is filtered by a Boolean function of degree at least two, generating the key-stream. An adversary can have access to the key-stream (this corresponds to the known plaintext-ciphertext pairs model) and build the corresponding algebraic system. Solving this system gives the secret variables, *i.e.* the secret key, directly breaking the security (as it enables to decrypt any message encrypted with this key).

Other examples of constructions for which the security mostly depends on the algebraic properties of a Boolean functions are given by the family of stream-ciphers FLIP [29], and the local PseudoRandom Generators (PRG) following the blueprint of Goldreich's PRG [21]. The first example is a recent design of symmetric encryption scheme which goal is to facilitate the use of Fully Homomorphic Encryption [20] for outsourcing computation. In this construction, for each produced key-stream bit, a public reordering of the key-bits is performed, and a fixed Boolean function is applied. The second example relates to cryptographic primitives that can exist in low-complexity classes such as NC0, and these PRG have been the focus of many attentions lastly. These are considered as potential building blocks for indistinguishability obfuscation [1, 24, 25]. For these PRG, each output is the result of a fixed Boolean function applied on a subpart of the seed, the subpart being publicly determined. In these two cases (as for filtered LFSR), an adversary can build an algebraic system of equation in the secret variables, from the main function only, and attacks solving these algebraic systems are amongst the most efficient against those primitives.

Many algorithms are known to solve algebraic systems of equations, one of the most efficient being based on Grobner bases such as F5 [19]. The algorithms based on linearization techniques are generally less efficient but with easier to determine time and data complexities. For these algorithms and corresponding attacks, the complexities can be determined only based on algebraic criteria on the Boolean function. The first attack known to apply on the constructions we listed is an attack based on the algebraic degree of the function, that we note d. Indeed, as all the equations of algebraic system have this degree, it is possible for the attacker to rewrite any monomial of degree at most d in a new variable. Then, it corresponds to a linear system in at most $D = \sum_{i=0}^{d} \binom{n}{d}$ variables, where we note n the number of secret variables. Finally, solving such linear system can be performed in time complexity $O(D^\omega)$ (where ω is the exponent in linear system inversion, such as $\omega = 2.807$ for Strassen's algorithm, and 2.373 for the latest results), giving the cost of these attacks.

A simpler system may be obtained by considering algebraic properties of the function other than its degree. Calling f the Boolean function giving the equations, the Algebraic Immunity (AI) of f is defined as the smallest integer d such that there exists a function g (nonzero) of degree d for which $fg = 0$ on all inputs (or $g(f + 1) = 0$ on all inputs). In 2003, Courtois and Meier [14] showed that an attack can be mounted on filtered LFSR using algebraic systems of degree at most the algebraic immunity of the function f. As the algebraic immunity is at most equal to the degree of a function, the so-called algebraic attack has a better complexity than the one targeting the degree. Recently, the algebraic attacks have been rediscovered in the context of Goldreich's PRG [2], also giving better attacks than the one based on the degree. A more general

algebraic property of f is its Fast Algebraic Immunity (FAI). This property takes in account the degrees of both the function g and of the product fg, where g is a nonzero function of degree smaller than $\mathsf{AI}(f)$. The principle of the corresponding attack is to perform linear combinations of the system's equations to cancel the monomials of degree between $\mathsf{deg}(g)$ and $\mathsf{deg}(fg)$. Hence, it gives an algebraic system of even smaller degree than those obtained from other algebraic properties of f, and which can be more easily linearized. It corresponds to a more efficient class of attack called fast algebraic attack [13].

1.2 Determining Algebraic Properties, and Good Functions

Once the properties required for security are exhibited, two natural questions have to been answered. The first concern is how to (efficiently) determine the parameter of each function relatively to one property. The second one consists in determining which functions have good parameters for these properties.

For the first question, the situation is very different depending on which algebraic property is targeted. The algebraic degree of a Boolean function is directly given by its Algebraic Normal Form (ANF), one of the most common representation. Various works focus on efficient algorithms to determine the degree directly from the truth table of a Boolean function such as [12]. The algebraic immunity is more complex to determine from the ANF, or any other common representation. This parameter for a particular function f can be determined by considering the rank of Reed-Muller codes punctured at the support of f, and at the support of $f + 1$. In [4] a more efficient algorithm using multivariate interpolation is given, and further works are dedicated to assess the complexity of such algorithms [16, 22]. Determining the fast algebraic immunity is a more intricate task, since all known algorithms to do this become less efficient in this case. An algorithm to determine this parameter is also given in [4], with a running time complexity in $\mathcal{O}(DE^2)$ where $D = \binom{n}{d}$ and $E = \binom{n}{e}$, n being the number of variables of the function f, e the degree of g and d the degree of fg. Note that in the worst case this complexity is exponential as d can be as high as $\lceil (n+1)/2 \rceil$, giving $D \approx 2^{n/2}$ (using Stirling approximation). As a result, this kind of algorithms cannot be used to determine the fast algebraic immunity of arbitrary functions of hundreds of variables that are used in some cryptographic constructions.

For the second question, the difficulty of finding functions with good parameters is scaling up with the generality of the property. Any n-variable function with monomials of degree n in its ANF has maximum algebraic degree (or equivalently, odd weight truth table), which allows to easily define families of functions (indexed by n) with good algebraic degree. Finding functions with optimal algebraic immunity is the central topic of different works such as [17, 18]. This line of works shows that exhibiting families of functions with good algebraic immunity is much more complex than exhibiting functions with good algebraic degree. Typical examples of such families are the majority functions and modifications, or Carlet-Feng [9] functions. In comparison, only partial results have been obtained about families of functions with optimal, or good, fast algebraic immunity. In various works, this concern is tackled by showing a lower bound on this parameter for functions already known to have good characteristics relatively to other properties. For example, Carlet-Feng functions have optimal fast

algebraic immunity when they are defined in 2^s+1 variables [27]. T-C-T functions [35] have optimal AI and almost optimal FAI [26], and the functions introduced in [34] have optimal AI and FAI of at least $n-6$.

Another approach to build sufficiently good functions for the algebraic criteria is to combine good functions with others such that the properties of the combination can only increase. This principle is used for Goldreich's PRG [21], and for FLIP [29], where a function with good (F)AI is one of the constituent of a direct sum giving the main function. In these cases, the algebraic degree, AI, and FAI of the main functions are at least equal to the one of the function with good algebraic properties. Then, determining the exact algebraic parameters of a good function directly leads to a lower bound on the parameters of the more complex function, which is the main component of the cryptographic primitive.

1.3 Majority Functions, and their Fast Algebraic Immunity

The family of majority functions has been the center of various studies in the context of cryptographic constructions. It is one of the first examples of functions proven to reach the optimal algebraic immunity [18]. As it is the most particular case of symmetric functions, this family is known and studied in diverse contexts, for example as easy-to-compute functions with branching programs. Despite their optimal algebraic immunity, majority functions are not good for all cryptographic properties (for non-linearity or resilience for example [15, 18]), hence they are not directly used as filtering functions. Nevertheless, they are used in diverse constructions combined with other functions, for example in the XOR-MAJ functions [2,3], or in the Caesar competitor ACORN [23]. In the context of Goldreich's PRG [2, 3], or Improved Filter Permutator [28] (a paradigm of stream-cipher), the main function is the result of a secondary construction (a direct sum) where one of the two components is a majority function. The degree, AI and FAI of the majority function used are then giving lower bounds of the same parameters for the main function.

The cryptographic properties of majority functions and more generally of symmetric functions are investigated in many works such as [5, 6, 7, 18, 30, 31, 33]. From these series of works some properties of symmetric functions are better known that others. The balancedness, and more generally resilience of (non affine) symmetric function is still the object of conjectures. On the opposite extreme the algebraic degree is the property which is better known for this class of functions. As majority functions are known to have optimal algebraic immunity since a decade, and balanced for n odd, there are more results on the parameters of this subclass. The resilience, nonlinearity and algebraic immunity are exhibited for all elements of this family, but it is still not the case for the fast algebraic immunity.

A fundamental result regarding the algebraic properties of majority function is given in [4]. Despite having optimal algebraic immunity, majority functions do not reach the optimal fast algebraic immunity. More specifically, in this paper the authors show an upper bound smaller than n for majority functions, considering as function g a particular symmetric function. However, this result does not allow us to obtain the exact fast algebraic immunity of majority function, which affects their use as building blocks in secondary constructions. Up to now, very little has been shown for these exact

parameters. Two papers focus on this specific question, proving the exact parameter for particular subclasses. [36] gave the first relative result, where the exact FAI is exhibited for the subclasses of majority function in 2^m and $2^m + 1$ variables. More recently, the parameter of two other subclasses have been determined in [11], where the cases 2^m+2 and $2^m + 3$ are tackled for $m \geq 2$.

1.4 Our Contributions

In this article we exhibit the fast algebraic immunity of all majority function in $2^m + j$ variables with $m \geq 2$ and $0 \leq j < 2^{m-1}$. These results are obtained using different properties relatively to the annihilators of threshold functions (a subclass of symmetric functions containing majority functions), to their algebraic immunity, and to the structure of their algebraic normal form. These results give lower bounds on the complexity of fast algebraic attack on Goldreich's PRG or FILIP stream-cipher instantiated with XOR-MAJ functions [2, 3, 28].

To obtain these results, we first focus on the minimal degree of the nonzero annihilators of majority functions and their complementary. Combining it with the upper bound of [4], we can derive an interval for the FAI of any majority function in $n \geq 2$ variables. We show how this interval directly implies the result of [36]. Then, we use and combine different results on the representation of symmetric functions. These results allow us to exhibit an expression of the algebraic normal form of any threshold function. Finally, we show that for many values of n the functions reach the upper bound of [4]. To get this result, we show that multiplying the majority function by a low degree function cannot degrade too much the degree of the product, which corresponds to the case of low FAI. This is done by partitioning the majority function in two parts, one with the monomials of degree less that t, the other one with the monomials of degree equal or greater than t. We show that there is a quantified gap between the highest degree appearing in the low degree part, and the lowest degree appearing in the high degree part. We prove that the upper part corresponds to a threshold function. Moreover, using results on the degree of the annihilators of threshold functions we can derive results on the initial (partitioned) majority function, allowing us to state the final result.

1.5 Paper Organization

The article is organized in the following way: In Section 2 we define the notations and preliminary notions necessary to follow the main results. Section 3 is dedicated to the lower and upper bound of the FAI of majority functions. Section 4 shows results relative to the ANF of threshold functions. In Section 5 we present the main theorem, giving the exact FAI of several families of majority functions. Finally, Section 6 concludes on the results and open problems relative to this work.

2 Preliminaries

In addition to classic notations we use $[n]$ to denote the subset of all integers between 1 and n: $\{1, \ldots, n\}$. For readability we use the notation $+$ instead of $\oplus$ to denote the

addition in $\mathbb{F}_2$ and $\sum$ instead of $\bigoplus$. Let $v \in \mathbb{F}_2^n$, we refer to the element v as a word, or a Boolean vector of length n, we denote its coefficient v_i (for $i \in [n]$). When we consider $v \in \mathbb{F}_2^n$ as an integer we refers to the integer $\sum_{i=1}^{n} 2v_i^{i-1}$. The Hamming weight (or weight) of v is $\mathsf{wH}(v) = \#\{v_i \neq 0 \,|\, i \in [n]\}$. We denote $\overline{v} \in \mathbb{F}_2^n$ the complementary of v: $\forall i \in [n]$, $\overline{v_i} = 1 - v_i$. We call support of v and denote $\mathsf{supp}(v)$ the set of elements i in $[n]$ such that $v_i \neq 0$.

2.1 Boolean Functions, and order on $\mathbb{F}_n^n$

Definition 1 (Boolean Function). *A Boolean function f with n variables is a function from $\mathbb{F}_2^n$ to $\mathbb{F}_2$.*

Definition 2 (Algebraic Normal Form (ANF)). *We call Algebraic Normal Form of a Boolean function f its n-variable polynomial representation over $\mathbb{F}_2$ (i.e. belonging to $\mathbb{F}_2[x_1, \ldots, x_n]/(x_1^2 + x_1, \ldots, x_n^2 + x_n)$):*

$$f(x) = \sum_{I \subseteq [n]} a_I \left(\prod_{i \in I} x_i \right) = \sum_{I \subseteq [n]} a_I x^I,$$

where $a_I \in \mathbb{F}_2$.

Definition 3 (Order $\preceq$). *We denote $\preceq$ the partial order on $\mathbb{F}_2^n$ defined as:*

$$a \preceq b \Leftrightarrow \forall i \in [n], a_i \leq b_i,$$

where $\leq$ denotes the usual order on $\mathbb{Z}$ and the elements a_i and b_i of $\mathbb{F}_2$ are identified to 0 or 1 in $\mathbb{Z}$.

Property 1 (Corollary of Lucas's Theorem (*e.g.* [8]))**.** *Let u,$v \in \mathbb{F}_2^n$:*

$$u \preceq v \Leftrightarrow \binom{v}{u} \equiv 1 \mod 2,$$

where the inputs of the binomial coefficient are the integers whose binary decomposition corresponds to u and v.

2.2 Algebraic Immunity and Fast Algebraic Immunity

Definition 4 (Algebraic Immunity and Annihilators). *The algebraic immunity of a Boolean function $f \in \mathcal{B}_n$, denoted as $\mathsf{AI}(f)$, is defined as:*

$$\mathsf{AI}(f) = \min_{g \neq 0} \{\mathsf{deg}(g) \mid fg = 0 \text{ or } (f+1)g = 0\},$$

where $\mathsf{deg}(g)$ is the algebraic degree of g. The function g is called an annihilator of f (or $f + 1$).

We also use the notation $\mathsf{AN}(f)$ for the minimum algebraic degree of nonzero annihilator of f:

$$\mathsf{AN}(f) = \min_{g \neq 0} \{\mathsf{deg}(g) \mid fg = 0\}.$$

Property 2 (Algebraic Immunity Properties (*e.g.* [8]))**.** *Let f be a Boolean function:*

- *The null and the all-one functions are the only functions such that* $\mathsf{AI}(f) = 0$*,*
- *All monomial (non constant) functions f are such that* $\mathsf{AI}(f) = 1$*,*
- *For all non constant f it holds that:* $\mathsf{AI}(f) \leq \mathsf{AN}(f) \leq \mathsf{deg}(f)$*,*
- $\mathsf{AI}(f) \leq \lfloor \frac{n+1}{2} \rfloor$.

Definition 5 (Fast Algebraic Immunity [4]). *The fast algebraic immunity of a Boolean function $f \in \mathcal{B}_n$, denoted as* $\mathsf{FAI}(f)$*, is defined as:*

$$\mathsf{FAI}(f) = \min \left\{ 2\mathsf{AI}(f), \min_{1 \leq \mathsf{deg}(g) < \mathsf{AI}(f)} [\mathsf{deg}(g) + \mathsf{deg}(fg)] \right\}.$$

Property 3 (Fast Algebraic Immunity Properties (*e.g.* [8]))**.** *Let f be a Boolean function:*

- $\mathsf{FAI}(f) = \mathsf{FAI}(f+1)$,
- $\mathsf{FAI}(f) \leq n$*,*
- $\mathsf{FAI}(f) \geq \mathsf{AN}(f+1) + 1$.

Remark 1. The last item comes from the fact that $\mathsf{deg}(fg)$ is at least the degree of $\mathsf{AN}(f+1)$ as by construction fg is a nonzero annihilator of $f+1$.

2.3 Symmetric Functions

Symmetric functions are functions such that changing the order of the inputs does not change the output. They have been the focus of many studies *e.g.* [6, 7, 18, 30, 31, 33]. These functions can be described more succinctly through the simplified value vector.

Definition 6 (Simplified Value Vector). *Let f be a symmetric function in n variables, we define its simplified value vector:*

$$\mathbf{s}_f = [w_0, w_1, \ldots, w_n]$$

of length n, where for each $k \in \{0, \ldots, n\}$, $w_k = f(x)$ for x such that $\mathsf{w_H}(x) = k$, i.e. w_k is the value of f on all inputs of Hamming weight k.

Definition 7 (Elementary Symmetric Functions). *Let $n \in \mathbb{N}^*$, let $i \in \{0, \cdots, n\}$, the elementary symmetric function of degree i in n variables, denoted σ_i, is the function which ANF contains all monomials of degree i and no monomials of other degrees. The $n+1$ elementary symmetric functions in n variables form a basis of the symmetric functions in n variables.*

We define the sub-family of threshold functions, and then a sub-family of threshold functions of particular interest: the family of majority functions.

Definition 8 (Threshold Function). *For any positive integers $d \leq n+1$ we define the Boolean function $\mathsf{T}_{d,n}$ as:*

$$\forall x \in \mathbb{F}_2^n, \quad \mathsf{T}_{d,n}(x) = \begin{cases} 0 & \text{if } \mathsf{w_H}(x) < d, \\ 1 & \text{otherwise.} \end{cases}$$

Definition 9 (Majority Function). *For any positive integer n we define the Boolean function MAJ_n as:*

$$\forall x \in \mathbb{F}_2^n, \quad \mathsf{MAJ}_n(x) = \begin{cases} 0 & \text{if } \mathsf{wH}(x) \leq \lfloor \frac{n}{2} \rfloor, \\ 1 & \text{otherwise.} \end{cases}$$

Note that for a threshold function, we have $w_k = 0$ for $k < d$ and 1 otherwise, so the simplified value vector of a threshold function $\mathsf{T}_{d,n}$ is the $n+1$-length vector of d consecutive 0's and $n+1-d$ consecutive 1's.

Remark 2 (Convention on Majority). Note that for n even it gives $\mathsf{MAJ}_n = \mathsf{T}_{\frac{n}{2}+1,n}$ and for n odd $\mathsf{MAJ}_n = \mathsf{T}_{\frac{n+1}{2},n}$. In the case of n even, the choice of $\mathsf{T}_{\frac{n}{2},n}$ or $\mathsf{T}_{\frac{n}{2}+1,n}$ as the majority function is arbitrary, some papers considers the second choice. As shown by the following proposition, it does not matter in this work as both functions have the same behavior relatively to fast algebraic immunity. We include the proof of [10] for the ease of the reader.

Proposition 1 ([10]). *Let $n \in \mathbb{N}^*$ and $d \in [0, n+1]$, for all $x \in \mathbb{F}_2^n$ let $1_n + x$ denote the element $(1 + x_1, \ldots, 1 + x_n) \in \mathbb{F}_2^n$, then the following relation holds for $\mathsf{T}_{d,n}$ and $\mathsf{T}_{n-d+1,n}$:*

$$\forall x \in \mathbb{F}_2^n, \quad 1 + \mathsf{T}_{d,n}(1_n + x) = \mathsf{T}_{n-d+1,n}(x).$$

Proof. We use the simplified value vector formalization (see Definition 6) to show this result. For all elements $x \in \mathbb{F}_2^n$, we have $\mathsf{wH}(x+1_n) = \mathsf{wH}(1_n) - \mathsf{wH}(x) = n - \mathsf{wH}(x)$. So denoting w'_k the coefficients of the simplified value vector of $\mathsf{T}_{d,n}(1_n + x)$ we get: $w'_k = w_{n-k}$ for all $k \in [0, n]$. It gives a vector symmetric to the first simplified value vector, *i.e.* with the elements from 0 to $n - d$ being 1 and from $n - d + 1$ to n being 0.

For all $x \in \mathbb{F}_2^n$, $1 + \mathsf{T}_{d,n}(1_n + x) = \overline{\mathsf{T}_{d,n}(1_n + x)}$, its complement to 1. Then, denoting $w"_k$ the coefficients of the simplified value vector of $1 + \mathsf{T}_{d,n}(1_n + x)$ we get: $w"_k = \overline{w_{n-k}}$ for all $k \in [0, n]$. It gives a vector which is the complement of the precedent simplified value vector to the $(n+1)$-length all-1 vector, *i.e.* with the elements from 0 to $n - d$ being 0 and from $n - d + 1$ to n being 1. This simplified value vector is the one of $\mathsf{T}_{n-d+1,n}$, finishing the proof.

□

More precisely, for the case $d = n/2$ it gives that $\mathsf{T}_{\frac{n}{2},n}$ and $\mathsf{T}_{\frac{n}{2}+1,n}$ are extended affine equivalent, then having the same degree, algebraic immunity, and fast algebraic immunity.

3 Lower and Upper Bound on The Fast Algebraic Immunity of Majority Functions

In the following we will express n as $2^m + 2k + \varepsilon$ where $0 \leq k < 2^{m-1}$, and $\varepsilon = 0$ or 1, as this writing will be convenient to highlight the properties of the functions used. First we recall the result of [4] giving the upper bound on MAJ_n, using the formulation developed in [36] and [11]:

Lemma 1 (Upper Bound on $\mathsf{FAI}(\mathsf{MAJ}_n)$, [4] Theorem 2). *Let $n \geq 2$ such that $n = 2^m + 2k + \varepsilon$, $m \geq 1$, $0 \leq k < 2^{m-1}$, and $\varepsilon \in \{0, 1\}$, Then:*

$$\mathsf{FAI}(\mathsf{MAJ}_n) \leq 2^{m-1} + 2k + 2.$$

For the lower bound we recall the result of [10] on the nonzero annihilator of minimal degree of threshold functions and then we combine it with Property 3. We include the proof of [10] for the ease of the reader.

Lemma 2 (AN of Threshold Functions, [10]). *Let n be a nonzero positive integer, $1 \leq d \leq n$, the threshold function $\mathsf{T}_{d,n}$ has the following property:*

$$\mathsf{AN}(\mathsf{T}_{d,n}) = n - d + 1, \text{ and } \mathsf{AN}(1 + \mathsf{T}_{d,n}) = d.$$

Proof. Applying the transformation $x \mapsto x + 1_n$, where 1_n is the all-1 vector of length n, changes $\mathsf{T}_{d,n}$ into the indicator of the set of vectors of Hamming weight at most $n - d$. The relations between the expressions of the coefficients of the ANF $\sum_{I \subseteq [n]} a_I x^I$ by means of the values of the function, namely, $a_I = \sum_{\mathsf{supp}(x) \subseteq I} f(x)$ and $f(x) = \sum_{I \subseteq \mathsf{supp}(x)} a_I$, show that the annihilators of this indicator are all the linear combinations over $\mathbb{F}_2$ of the monomials of degrees at least $n - d + 1$. Hence, the annihilators of $\mathsf{T}_{d,n}$ are obtained from these latter linear combinations by the transformation $x \mapsto x + 1_n$. They can have every algebraic degree at least $n - d + 1$. And the annihilators of $1 + \mathsf{T}_{d,n}$ are similarly the linear combinations over $\mathbb{F}_2$ of the monomials of degrees at least d. They can have every algebraic degree at least d. Hence $\mathsf{AN}(\mathsf{T}_{d,n}) = n - d + 1$, $\mathsf{AN}(1 + \mathsf{T}_{d,n}) = d$, and $\mathsf{AI}(\mathsf{T}_{d,n}) = \min(d, n - d + 1)$. □

Lemma 3 (Lower Bound on $\mathsf{FAI}(\mathsf{MAJ}_n)$). *Let $n > 2$ such that $n = 2^m + 2k + \varepsilon$, $m \geq 1$, $0 \leq k < 2^{m-1}$, and $\varepsilon \in \{0, 1\}$, Then:*

$$\mathsf{FAI}(\mathsf{MAJ}_n) \geq 2^{m-1} + k + 2.$$

Proof. First, for such values of n note that MAJ_n is the threshold function $\mathsf{T}_{d,n}$ with $d = 2^{m-1} + k + 1$. Using Lemma 2 gives $\mathsf{AN}(\mathsf{T}_{d,n}) = 2^{m-1} + k + \varepsilon$ and $\mathsf{AN}(1 + \mathsf{T}_{d,n}) = 2^{m-1} + k + 1$. As for any function f $\mathsf{AI}(f) = \min(\mathsf{AN}(f), \mathsf{AN}(f + 1))$, it leads to $\mathsf{AI}(\mathsf{MAJ}_n) = 2^{m-1} + k + \varepsilon$. The third item of Property 3 enables to obtain:

$$\mathsf{FAI}(\mathsf{MAJ}_n) \geq \min(2^m + 2k + 2\varepsilon, 2^{m-1} + k + 2).$$

When $m \geq 2$ or $\varepsilon = 1$ this minimum is reached by $2^{m-1} + k + 2$, then gives the final result for $n > 2$.

□

Note that the difference between the upper and lower bound is only of k, then they coincide when $k = 0$, giving the result previously obtained in [36]:

Corollary 1 (Fast Algebraic Immunity of MAJ_n for $n = 2^m + \varepsilon$). *Let $n = 2^m + \varepsilon$, $n > 2$, $m \geq 1$, $\varepsilon \in \{0, 1\}$, then $\mathsf{FAI}(\mathsf{MAJ}_n) = 2^{m-1} + 2$.*

The intuition behind our main result is that many functions are reaching the upper bound. A way of proving this consists in showing that for all functions of degree at most k, the product with the majority function is higher that the lower bound given by $\mathsf{AN}(\mathsf{MAJ}_n)$. A difference of k with this bound is sufficient to show that $\mathsf{FAI}(\mathsf{MAJ}_n)$ reaches the upper bound. In order to determine when this difference happens, we will use the properties of algebraic immunity of various threshold functions. Such strategy is possible since considering the ANF of a majority function restricted to its monomials of degree superior to a fixed level corresponds to the ANF of another threshold function.

4 Exhibiting The Algebraic Normal Form of Threshold Functions

In order to prove our main result we need to exhibit the ANF of threshold function as a sum of elementary symmetric functions, a form which is easier to use to reach our objective. First we show some properties on the ANF of these functions, and then we give a general expression. In this part we use the fact that some properties of symmetric functions can be derived from the periodicity of their simplified value vector [6].

4.1 Some Properties on the ANF of Threshold Functions

We first recall how a symmetric function can be written as a sum of elementary symmetric function, using Lucas's theorem and the simplified value vector (*e.g.* [8]).

Lemma 4 (Symmetric Function as a Sum of Elementary Symmetric Function, [8] p144).
Let f be a symmetric n-variable Boolean function with simplified value vector $\mathbf{s}_f = [w_0, w_1, \ldots, w_n]$, then $\forall x \in \mathbb{F}_2^n$:

$$f(x) = \sum_{i=0}^{n} \lambda_i \sigma_i(x), \quad \textit{where } \lambda_i = \sum_{j \preceq i} w_j.$$

Let us denote $D = 2^{\lceil \log d \rceil}$, we show that the indices of the elementary symmetric functions appearing in the ANF of the threshold function $\mathsf{T}_{d,n}$ follow a period of D.

Lemma 5 (ANF of Threshold Functions and Periodicity). *Let n and d be two integers such that $0 < d \leq n+1$, let $D = 2^{\lceil \log d \rceil}$, if $\mathsf{T}_{d,n} = \sum_{i'=0}^{n} \lambda_{i'} \sigma_{i'}$ then the following relation holds on its coefficients:*

$$\forall i' \in [n],\ \lambda_{i'} = \lambda_i \,|\, i' \equiv i \mod D, \textit{ and } i \in [D].$$

Proof. Let us consider the integer $i' \in [n]$, it can be written as $i' = i + kD$ with $i \in [D]$ and $k \in \mathbb{N}$. Using Lemma 4 and partitioning the sum in intervals of size D (except the last one stopping at $i + kD$ since no bigger element complies the order relation) we get:

$$\lambda_{i'} = \sum_{\substack{j \preceq i+kD \\ 0 \leq j < D}} w_j + \cdots + \sum_{\substack{j \preceq i+kD \\ \ell D \leq j < (\ell+1) D}} w_j + \cdots + \sum_{\substack{j \preceq i+kD \\ kd \leq j \leq i+kD}} w_j. \tag{1}$$

We first consider the case $i \neq D$. In the set $[0, i]$ the number of elements j such that $0 \leq j \leq i$, $j \preceq i$ is even (it can be seen using Property 1 and $\sum_{j=0}^{i} \binom{i}{j} = 2^i \equiv 0 \mod 2$). Since $j > i$ implies $j \not\preceq i$, then an even number of elements of [0,D-1] are such that $j \preceq i$. When $j = \ell D + r$ with $0 \leq r < D$ we have $j \preceq i + kD \Leftrightarrow r \preceq i$ and $\ell \preceq k$, allowing us to conclude that all the sums from the second to the last contain an even number of w_j. Since all these coefficients are for $j \geq D$ and we set $D \geq d$, and by the definition of threshold function $\mathsf{T}_{d,n}$, all these coefficients are equal to 1, giving:

$$\lambda_{i'} = \sum_{\substack{j \preceq i + kD \\ 0 \leq j < D}} w_j = \lambda_i.$$

For the remaining case, $i = D$, we need to show that $\lambda_{kD} = \lambda_D$ for $k \in \mathbb{N}^*$ (note that λ_0 is not considered). Since i' is a multiple of D, at most one coefficient of each sum can contribute to the total sum, and in this case Equation 1 gives:

$$\lambda_{kD} = \sum_{\substack{j \preceq kD \\ j=0}} w_j + \cdots + \sum_{\substack{j \preceq kD \\ j=\ell D}} w_j + \cdots + \sum_{\substack{j \preceq kD \\ j=kD}} w_j.$$

The number of elements $j \preceq kD$ such that j is a multiple of D corresponds to the number of elements of $[0, k]$ preceding k, which is even as $k > 0$. All coefficients $w_{\ell D}$ with $\ell \in [k]$ are equal to 1 and w_0 is equal to 0 as $0 < d \leq D$, therefore (since $0 \preceq kD$ for all k), it implies $\lambda_{kD} = 1$ for $k \in \mathbb{N}^*$. We can then conclude that $\lambda_{kD} = \lambda_D$ for $k \in \mathbb{N}^*$.

□

As a result we can link the presence of an elementary symmetric function in the ANF of a threshold function to a property on a small set.

Proposition 2. *Let n and d be two integers such that $0 < d \leq n + 1$, let $D = 2^{\lceil \log d \rceil}$, let $\mathsf{T}_{d,n} = \sum_{i'=0}^{n} \lambda_{i'} \sigma_{i'}$. For all $i' \in [n]$ let i be the integer such that $i' \equiv i \mod D$ and $i \in [D]$, then:*

$$\forall i' \in [n],\ \lambda_{i'} = \#\{j \preceq i \,|\, j \in [d, D]\} \mod 2.$$

Proof. From Lemma 5 we know that $\lambda_{i'} = \lambda_i$, and using Lemma 4, $\lambda_i = \sum_{j \preceq i} w_j$, since $w_j = 0$ for $j < d$ and $w_j = 1$ for $j \geq d$, and $i \leq D$ gives $\lambda_i = \sum_{d \leq j \leq D \,|\, j \preceq i} 1$, which is equivalent to the final result.

□

Finally, we highlight that for majority functions, only the set $[d, D]$ is important to exhibit the ANF.

Proposition 3. *Let $n > 2$ and $\mathsf{MAJ}_n = \mathsf{T}_{d,n} = \sum_{i'=0}^{n} \lambda_{i'} \sigma_{i'}$, $D = 2^{\lceil \log d \rceil}$, then: $\lambda_{i'} = 1 \Rightarrow d \leq i' \leq D$.*

Proof. Let us write n as $2^m + 2k + \varepsilon$ such as $m \geq 1$, $0 \leq k < 2^{m-1}$, and $\varepsilon \in \{0, 1\}$. This fixes $d = 2^{m-1} + k + 1$ and $D = 2^m$. We first show that none of the indices

smaller than d can appear in the ANF, and that the same holds for all indexes between $D+1$ and n.

When $d > 0$ we have $\lambda_0 = w_0 = 0$. For all $i \in [d-1]$ we have $\{j \preceq i \,|\, j \in [d, D]\} = \emptyset$, so Proposition 2 gives $\lambda_i = 0$. Hence, no coefficient smaller than $2^{m-1} + k + 1$ appears in the ANF of a majority function.

Now, note that all the elements in $[D+1, n]$ have their representative in $[D]$ (the element with same congruence modulus D) belonging to $[2k+\varepsilon]$ since $n = D+2k+\varepsilon$. Because $k < 2^{m-1}$, we have $2k+\varepsilon < 2^{m-1}+k+1 = d$ and combining Lemma 5 with Proposition 3 implies that $\forall i' \in [D+1, n], \lambda_{i'} = 0$. Hence, no coefficient bigger than 2^m appears in the ANF of a majority function. In conclusion only coefficients between $2^{m-1} + k + 1$ and 2^m appear in the ANF.

□

4.2 A Simple Expression of the Algebraic Normal Form of Threshold Functions

In this section, we give a simple expression of the ANF of threshold functions using some set representation. This expression will be used to exhibit the main results of Section 5. We begin with a preliminary lemma relative to the order $\preceq$ which simplifies the proof of the main theorem of this section.

Lemma 6. *Let $a, b-1 \in \mathbb{F}_2^n$ then $a \preceq \overline{b-1} \Leftrightarrow a \preceq a+b-1$.*

Proof. We first show $a \preceq \overline{b-1} \Rightarrow a \preceq a+b-1$. $a \preceq \overline{b-1}$ implies $\mathsf{supp}(a) \cap \mathsf{supp}(b-1) = \emptyset$, so $a \preceq a+b-1$. We prove the other direction by contrapositive:

$$a \not\preceq \overline{b-1} \Rightarrow \exists i \in [n] \,|\, a_i = 1 \text{ and } \overline{b-1}_i = 0.$$

Taking the smallest index i with this property, $a_i = 1$ and $(b-1)_i = 1$ and it is the first carry, hence $a_i = 1$ and $(a+b-1)_i = 0$ giving $a \not\preceq a+b-1$, finishing the proof.

□

Theorem 1 (Algebraic Normal Form of Threshold Functions). *Let n and d be two integers such that $0 < d \leq n+1$, let $D = 2^{\lceil \log d \rceil}$. Let $\mathsf{T}_{d,n} = \sum_{i'=0}^{n} \lambda_{i'} \sigma_{i'}$, and let S_d denote the set $\{v \in [0, D-1] \,|\, v \preceq D-d\}$ also equal to $\{v \in \mathbb{F}_2^{\lceil \log d \rceil} \,|\, v \preceq \overline{d-1}\}$ where $d-1$ is considered over $\log(D)-1$ bits. The following relation holds:*

$$\lambda_{i'} = 1 \Leftrightarrow i' \in S'_d,$$

where $S'_d = \{kD + d + v \,|\, k \in \mathbb{N}, v \in S_d\} \cap [n] = \{kD - v \,|\, k \in \mathbb{N}^*, v \in S_d\} \cap [n]$. Or equivalently:

$$\mathsf{T}_{d,n} = \sum_{i \in S'_d} \sigma_i.$$

Proof. We will first show that σ_0 is never in the ANF of these threshold functions, then that the two definitions of S_d are equivalent, as the two definitions of S'_d, and finally that the appearance in the ANF is equivalent to the membership in S'_d.

For $d > 0$, $w_0 = \lambda_0 = 0$ and no threshold functions with $d > 0$ can have σ_0 in the ANF. By definition, $\mathsf{T}_{0,n}$ is the constant n-variable function 1, for which the ANF is $1 = \sigma_0$.

For the set equivalences, we consider both the integer representation and the Boolean vector representation. We first focus on the set S_d, $v \in [0, D-1]$ means that v is a positive integer smaller than $D = 2^{\lceil \log d \rceil}$ which is equivalent to a Boolean vector of length $\lceil \log d \rceil$. Then, as $D = 2^{\lceil \log d \rceil}$, $D - d = D - 1 - (d-1)$ where $D-1$ and $d-1$ can both be written on $\lceil \log d \rceil - 1$ bits and $D-1$ is the all 1 vector of this length. Therefore $D - d = \overline{d-1}$ on $\log(D) - 1$ bits, which proves the equivalent representations of S_d. Note that $v \preceq D - d \Leftrightarrow v = D - d - v' \,|\, v' \preceq D - d$, and $\{d + v \,|\, v \in S_d\} = \{D - v' \,|\, v' \in S_d\}$ implies the equivalence of the two definitions of S'_d.

Using Proposition 2, for all $i' \in [n]$ we have $\lambda_{i'} = 1$ if and only if the set $\{j \preceq i \,|\, j \in [d, D]\}$ has odd cardinality. Property 1 gives that the Boolean value $j \preceq i$ equals the parity of the binomial coefficient i choose j. Using Pascal's identity as $d > 0$ we get:

$$\begin{aligned}\sum_{j=d}^{D} \binom{i}{j} &\equiv \sum_{j=d}^{D} \left[\binom{i-1}{j} + \binom{i-1}{j-1} \right] \equiv \binom{i-1}{d-1} + \binom{i-1}{D} \mod 2, \\ &\equiv \binom{i-1}{d-1} \equiv \binom{i-1}{i-d} \mod 2.\end{aligned}$$

The parity of the cardinality of the set is then the Boolean value $i - d \preceq i - 1$. If $i - d < 0$ then $\lambda_{i'} = 0$ from Proposition 3, and both $i - d$ and $i - 1$ are non negative integers smaller than D (the case $i' = 0$ has already been considered). Therefore, we can identify a to $i-d$ and b to i in Lemma 6, which gives that $\lambda_{i'}$ is equal to the Boolean value $i - d \preceq \overline{d-1}$. It enables us to conclude:

$$\begin{aligned}\lambda_{i'} = 1 &\Leftrightarrow i' \in [n], i' = kD + i, i \in [D], k \in \mathbb{N}, i - d \in S_d, \\ &\Leftrightarrow i' \in [n], i' = kD + d + v, d + v \in [D], k \in \mathbb{N}, v \in S_d, \\ &\Leftrightarrow i' \in S'_d.\end{aligned}$$

□

Remark 3. Note that the threshold functions in n variables form a basis of the n-variable symmetric function, then this representation with the sets S'_d can be used to obtain the ANF of any symmetric function. For example, the ANF of the indicator function of the elements of Hamming weight d, $\varphi_{d,n} = \mathsf{T}_{d,n} + \mathsf{T}_{d+1,n}$, is given by $S'_d \Delta S'_{d+1}$, where Δ denotes the symmetric difference of sets.

5 Exact Fast Algebraic Immunity of Several Families

In this section we show that many majority functions reach the upper bound of [4]. To do so, we use the ANF formulation of Theorem 1 to show a gap between two consecutive elementary symmetric functions appearing in the ANF of a majority function, σ_a and

σ_b. We use the fact that for functions of degree smaller than this gap the product with the part of degree up to a has a smaller degree than b. Then, when the function obtained by the part of degree at least b is a threshold function, we can determine a lower bound on the degree of its product with a function of degree at most the gap. By construction, this lower bound also applies to the degree of the product of the majority function with a function of degree at most the gap, giving a lower bound on the fast algebraic immunity.

We write the integer n as in Section 3: $n = 2^m + 2k + \varepsilon$ but with $k < 2^{m-2}$ this time, and we show that in this case the ANF of MAJ_n has no monomials of degree between $2^m - 2^{m-2} - 1$ and $2^m - 2^{m-2} + k$.

Lemma 7. *Let $n > 2$ be an integer such that $n = 2^m + 2k + \varepsilon, m \geq 2, 0 \leq k < 2^{m-2}, \varepsilon \in \{0,1\}$. Let $\mathsf{MAJ}_n = \sum_{i'=0}^{n} \lambda_{i'}\sigma_{i'}$, the following holds:*

$$\{i' \in [2^m - 2^{m-2}, 2^m - 2^{m-2} + k + 1] \,|\, \lambda_{i'} = 1\} = \{2^m - 2^{m-2}, 2^m - 2^{m-2} + k + 1\}.$$

Proof. Using Theorem 1 we know that $S'_d = \{\ell D - v \,|\, \ell \in \mathbb{N}^*, v \in S_d\} \cap [n]$, and using Proposition 3:

$$S'_d = \{2^m - v \,|\, v \in S_d\}.$$

Since $d = 2^{m-1} + k + 1$, we have $S_d = \{v \in \mathbb{F}_2^m \,|\, v \preceq \overline{2^{m-1} + k + 1 - 1}\}$. Then, $\overline{2^{m-1} + k} = 2^m - 1 - (2^{m-1} + k) = 2^{m-1} - k - 1$. In the following we show which integers are covered or not by $2^{m-1} - k - 1$ in the set we are interested in.

Writing $2^{m-1} - k - 1$ as $2^{m-2} + 2^{m-2} - k - 1$, and since $2^{m-2} - k - 1 < 2^{m-2}$ (2^{m-2} and $2^{m-2} - k - 1$ have disjoint support) we see that $2^{m-2} - k - 1 \preceq 2^{m-2} + 2^{m-2} - k - 1$ (and $2^{m-2} \preceq 2^{m-2} + 2^{m-2} - k - 1$). It means that $2^{m-2} - k - 1 \preceq 2^{m-1} - k - 1$ (and $2^{m-2} \preceq 2^{m-1} - k - 1$), hence $2^{m-2} - k - 1$ and 2^{m-2} are both in S_d.

The elements in $]2^{m-2} - k - 1, 2^{m-2}[$ are smaller than 2^{m-2} so they are covered by $2^{m-2} + 2^{m-2} - k - 1$ if and only if the are covered by $2^{m-2} - k - 1$, which is not the case as they are bigger than this number. Finally $S_d \cap]2^{m-2} - k - 1, 2^{m-2}[= \emptyset$. □

In the following we prove that the function obtained by considering only the monomials of degree at least $2^m - 2^{m-2} + k + 1$ of the MAJ_n function corresponds to the threshold function with $d = 2^m - 2^{m-2} + k + 1$.

Lemma 8. *Let $n > 2$ be an integer such that $n = 2^m + 2k + \varepsilon, m \geq 2, 0 \leq k < 2^{m-2}, \varepsilon \in \{0,1\}$. Let $\mathsf{MAJ}_n = \sum_{i'=0}^{n} \lambda_{i'}\sigma_{i'}$, then the following holds:*

$$\forall x \in \mathbb{F}_2^n, \sum_{i'=2^m-2^{m-2}+k+1}^{n} \lambda_{i'}\sigma_{i'}(x) = \mathsf{T}_{2^m-2^{m-2}+k+1,n}(x).$$

Proof. In order to prove that these functions coincide we can use the formalization of Theorem 1, it consists in showing:

$$S'_{2^{m-1}+k+1} \cap [2^m - 2^{m-2} + k + 1, n] = S'_{2^m-2^{m-2}+k+1}. \tag{2}$$

First, we show that both sets are subsets of $[2^m - 2^{m-2} + k + 1, 2^m]$. For the first one it is a direct consequence of Proposition 3 since $S'_{2^{m-1}+k+1} \subseteq [2^{m-1} + k + 1, 2^m]$.

For the other set, due to the periodicity proven in Lemma 5, there are elements greater than 2^m in $S'_{2^m-2^{m-2}+k+1}$ only if there are indexes i in $[2k+\varepsilon]$ such that the ANF coefficients of $\mathsf{T}_{2^m-2^{m-2}+k+1,n}$ are equal to 1. It is not the case as $k < 2^{m-2}$, implying that $2k+\varepsilon < 2^{m-2}+k+1 < 2^m - 2^{m-2} + k + 1$.

Then, writing $[2^m - 2^{m-2} + k + 1, 2^m]$ as $[2^m - v \,|\, v \in [0, 2^{m-2} - k - 1]]$, using the definitions of the sets S_d given in Theorem 1, Equation 2 can be simplified to:

$$\{v \in [0, 2^{m-2} - k - 1] \,|\, v \preceq 2^{m-1} - k - 1\} = \{v \in [0, 2^m - 1] \,|\, v \preceq 2^{m-2} - k - 1\}.$$

We show that both sets are equal to $\{v \in [0, 2^{m-2} - k - 1] \,|\, v \preceq 2^{m-2} - k - 1\}$. For the first one, note that $v < 2^{m-2}$ then $v \preceq 2^{m-1} - k - 1 \Rightarrow v \preceq 2^{m-2} - k - 1$. For the other one, note that $v > 2^{m-2} - k - 1 \Rightarrow v \not\preceq 2^{m-2} - k - 1$. The equivalence of these sets finishes the proof.

□

The next lemma shows a lower bound on the degree of a (particular case of) majority function multiplied by a low degree function. This lemma is the corner stone of the main theorem.

Lemma 9. *Let $n > 2$ be an integer such that $n = 2^m + 2k + \varepsilon$, with $m \geq 2, 0 \leq k < 2^{m-2}$, and $\varepsilon \in \{0, 1\}$. For all nonzero $n-$variable function g such that $\deg(g) \leq k$, the following holds:*

$$\mathsf{deg}(g\mathsf{MAJ}_n) \geq 2^m - 2^{m-2} + k + 1.$$

Proof. We begin by writing MAJ_n as the sum of a function of degree less than or equal to $2^m - 2^{m-2} + k$ and the the part of MAJ_n of degree at least $2^m - 2^{m-2} + k + 1$. Note that in terms of ANF it corresponds to a partition of the ANF of MAJ_n, we use the following notation:

$$\mathsf{MAJ}_n = f_{\leq t} + f_{>t},$$

where $f_{\leq t}$ is the degree less than or equal to $2^m - 2^{m-2} + k$ part, and $f_{>t}$ is the remaining part: all the monomials of degree between $2^m - 2^{m-2} + k + 1$ and $2^m + 2k + \varepsilon$. Applying Lemma 7, $\mathsf{deg} f_{\leq t} = 2^m - 2^{m-2}$, and applying Lemma 8, $f_{>t} = \mathsf{T}_{2^m-2^{m-2}+k+1,n}$.

Then, for any nonzero function g such that $\deg(g) \leq k$, the degree of $gf_{\leq t}$ is at most $2^m - 2^{m-2} + k$. Using Lemma 2 on $\mathsf{T}_{2^m-2^{m-2}+k+1,n}$, and since $\mathsf{AI}(\mathsf{T}_{2^m-2^{m-2}+k+1,n}) = \min(2^m - 2^{m-2} + k + 1, 2^{m-2} + k + \varepsilon) > k$, we have:

$$\mathsf{deg}(g\mathsf{T}_{2^m-2^{m-2}+k+1,n}) \geq \mathsf{AN}(1 + \mathsf{T}_{2^m-2^{m-2}+k+1,n}) = 2^m - 2^{m-2} + k + 1.$$

The degree of $gf_{\leq t}$ being at most $2^m - 2^{m-2} + k$ and the one of $gf_{>t}$ at least $2^m - 2^{m-2} + k + 1$, it gives for all g nonzero of degree at most f:

$$\mathsf{deg}(g\mathsf{MAJ}_n) = \mathsf{deg}(g(f_{\leq t} + f_{>t})) \geq 2^m - 2^{m-2} + k + 1.$$

□

This lemma enables us to state the theorem for the exact fast algebraic immunity of several families of majority functions:

Theorem 2 (Exact Fast Algebraic Immunity of Majority Functions). *Let* MAJ_n *be the* $n-variable$ *majority function such that* $n = 2^m + 2k + \varepsilon$*, where* $m \geq 2$*,* $0 \leq k < 2^{m-2}$*, and* $\varepsilon \in \{0,1\}$*:*

$$\mathsf{FAI}(\mathsf{MAJ}_n) = 2^{m-1} + 2k + 2.$$

Proof. First we recall the definition of the fast algebraic immunity, applied on MAJ_n:

$$\mathsf{FAI}(\mathsf{MAJ}_n) = \min\left\{2\mathsf{AI}(\mathsf{MAJ}_n), \min_{1\leq \mathsf{deg}(g)<\mathsf{AI}(\mathsf{MAJ}_n)}[\mathsf{deg}(g) + \mathsf{deg}(g\mathsf{MAJ}_n)]\right\}.$$

Lemma 2 gives $2\mathsf{AI}(\mathsf{MAJ}_n) = 2^m + 2k + 2\varepsilon$ which is greater than the upper bound stated in Lemma 1. Then, we consider the degree of the right term depending on the degree of g. The degree of $g\mathsf{MAJ}_n$ for $1 \leq \mathsf{deg}(g) < \mathsf{AI}(\mathsf{MAJ}_n)$ is at least $2^{m-1}+k+1$ using Property 3 and Lemma 2. Then we have two cases:

- If $\mathsf{deg}(g) > k$, then $\mathsf{deg}(g\mathsf{MAJ}_n)+\deg g \geq 2^{m-1}+k+1+k+1 = 2^{m-1}+2k+2$, which is the upper bound of $\mathsf{FAI}(\mathsf{MAJ}_n)$ from [4].
- If $\mathsf{deg}(g) \leq k$, then applying Lemma 9, $\mathsf{deg}(g\mathsf{MAJ}_n) \geq 2^m - 2^{m-2} + k + 1 = 2^{m-1} + 2^{m-2} + k + 1$. As $k < 2^{m-2}$, it gives $\mathsf{deg}(g\mathsf{MAJ}_n) \geq 2^{m-1} + 2k + 2$, also reaching the upper bound.

We can therefore conclude: $\mathsf{FAI}(\mathsf{MAJ}_n) = 2^{m-1} + 2k + 2$.

□

6 Conclusion

In conclusion, in this article we developed different techniques to determine the exact fast algebraic immunity of majority functions over n bits:

- First, we gave an expression of the ANF of any threshold function, using a particular set representation to determine more easily its decomposition in terms of elementary symmetric functions.
- Then, we showed some gaps between two consecutive degrees appearing in the ANF of a threshold function. We used these gaps to bound the degree of the products with functions of upper bounded degree.
- Finally, we exhibited that a subpart of a threshold function can correspond to another threshold function of higher threshold. The known properties of this second function can then be used to derive properties on the first one.

We used these techniques on a sub-case of majority functions, writing any integer n as $2^m + 2k + \varepsilon$ with m the biggest power of 2 smaller than or equal to n and ε used for the parity of n, our results applies for $m \geq 2$ and $k < 2^{m-2}$. Since $m \geq 2$, Theorem 2 gives the exact fast algebraic immunity of all function with $n \in [2^m, 2^m + 2^{m-1} - 1]$, which corresponds to half of the functions. In the formulation of [11, 36], the majority functions are considered in terms of family indexed by j such that $n = 2^m + j$. It corresponds to a more asymptotic oriented definition, where the result applies since m is bigger than a fixed value. In these terms, Theorem 2 applies to any j, giving that the

exact FAI of the majority in 2^m+j variables is $2^{m-1}+j+2$ for j even and $2^{m-1}+j+1$ for j odd, and the result applies since $m \geq \log(j+1)+1$.

Concerning the remaining cases, for $m < 2$, it corresponds to $n \in \{1,2,3\}$, where $k = 0$, the case $n = 3$ is taken in Corollary 1. The other cases are exceptions, for $n = 2$ the minimum is given by $2\mathsf{AI}(\mathsf{MAJ}_2) = 2$ which is optimal. For $n = 1$ the definition of FAI is not accurate as the formula would give 2 (whereas the FAI is supposed to be at most n). For $m \geq 2$, note that no majority function such that $2^{m-2} < k < 2^{m-1}$ can reach the upper bound of Lemma 1. Indeed, These majority functions have degree 2^m and then for any degree 1 function g:

$$\mathsf{deg}(g\mathsf{MAJ}_n) + \mathsf{deg}(g) \leq 2^m + 2, \text{ and } 2^m + 2 < 2^{m-1} + 2k + 2,$$

where we only use that $\mathsf{AI}(\mathsf{MAJ}_n) > 1$, and $\mathsf{AN}(1 + \mathsf{MAJ}_n) = 2^{m-1} + k + 1$ (from Lemma 2). This explains why no such values of k are taken in consideration it the main theorem. Nevertheless, for the remaining values the techniques we developed could still be used, but targeting a tighter upper bound.

As noted in Remark 3, the $n-$variable threshold functions form a basis of $n-$variable symmetric functions. As such, the results we presented allow to obtain the exact fast algebraic immunity of several families of majority functions. We hope than such techniques can be used to determine more precisely the algebraic properties (such as ANF, degree, AI, and FAI) of all symmetric functions.

7 Acknowledgements

The author is a beneficiary of a FSR Incoming Post-doctoral Fellowship.

References

1. Ananth, P., Sahai, A.: Projective arithmetic functional encryption and indistinguishability obfuscation from degree-5 multilinear maps. In: Coron, J., Nielsen, J.B. (eds.) EUROCRYPT 2017, Part I. pp. 152–181. LNCS, Springer, Heidelberg (May 2017)
2. Applebaum, B., Lovett, S.: Algebraic attacks against random local functions and their countermeasures. In: Wichs, D., Mansour, Y. (eds.) 48th ACM STOC. ACM Press (Jun 2016)
3. Applebaum, B., Lovett, S.: Algebraic attacks against random local functions and their countermeasures. SIAM J. Comput. pp. 52–79 (2018)
4. Armknecht, F., Carlet, C., Gaborit, P., Künzli, S., Meier, W., Ruatta, O.: Efficient computation of algebraic immunity for algebraic and fast algebraic attacks. In: Vaudenay, S. (ed.) EUROCRYPT 2006. LNCS, vol. 4004. Springer, Heidelberg (May / Jun 2006)
5. Braeken, A., Preneel, B.: On the algebraic immunity of symmetric boolean functions. In: Progress in Cryptology - INDOCRYPT 2005, 6th International Conference on Cryptology in India, Bangalore, India, December 10-12, 2005, Proceedings. pp. 35–48 (2005)
6. Canteaut, A., Videau, M.: Symmetric boolean functions. IEEE Trans. Information Theory pp. 2791–2811 (2005)
7. Carlet, C.: On the degree, nonlinearity, algebraic thickness, and nonnormality of boolean functions, with developments on symmetric functions. IEEE Trans. Information Theory pp. 2178–2185 (2004)

8. Carlet, C.: Boolean Functions for Cryptography and Error-Correcting Codes, p. 257397. Encyclopedia of Mathematics and its Applications, Cambridge University Press (2010)
9. Carlet, C., Feng, K.: An infinite class of balanced functions with optimal algebraic immunity, good immunity to fast algebraic attacks and good nonlinearity. In: Pieprzyk, J. (ed.) ASIACRYPT 2008. LNCS, vol. 5350, pp. 425–440. Springer, Heidelberg (Dec 2008)
10. Carlet, C., Méaux, P.: Boolean functions for homomorphic-friendly stream ciphers. In: Preprint (to appear) (2019)
11. Chen, Y., Guo, F., Zhang, L.: Fast algebraic immunity of $2^m + 2$ and $2^m + 3$ variables majority function. Cryptology ePrint Archive, Report 2019/286 (2019)
12. Climent, J.J., Garca, F., Requena, V.: The degree of a boolean function and some algebraic properties of its support. vol. 45, pp. 25–36 (05 2013). https://doi.org/10.2495/DATA130031
13. Courtois, N.: Fast algebraic attacks on stream ciphers with linear feedback. In: Boneh, D. (ed.) CRYPTO 2003. LNCS, vol. 2729, pp. 176–194. Springer, Heidelberg (Aug 2003)
14. Courtois, N., Meier, W.: Algebraic attacks on stream ciphers with linear feedback. In: Biham, E. (ed.) EUROCRYPT 2003. LNCS, vol. 2656. Springer, Heidelberg (May 2003)
15. Cusick, T.: Simpler proof for nonlinearity of majority function. arXiv:1710.02034v2 (2018)
16. Dalai, D.K.: Computing the rank of incidence matrix and the algebraic immunity of Boolean functions. Cryptology ePrint Archive, Report 2013/273 (2013), http://eprint.iacr.org/2013/273
17. Dalai, D.K., Gupta, K.C., Maitra, S.: Cryptographically significant Boolean functions: Construction and analysis in terms of algebraic immunity. In: Gilbert, H., Handschuh, H. (eds.) FSE 2005. LNCS, vol. 3557, pp. 98–111. Springer, Heidelberg (Feb 2005)
18. Dalai, D.K., Maitra, S., Sarkar, S.: Basic theory in construction of boolean functions with maximum possible annihilator immunity. Designs, Codes and Cryptography (2006)
19. Faugère, J.C.: A new efficient algorithm for computing Grobner bases without reduction to zero. In: Workshop on application of Groebner Bases 2002. Catania, Spain (2002), https://hal.inria.fr/inria-00100997
20. Gentry, C.: Fully homomorphic encryption using ideal lattices. In: Mitzenmacher, M. (ed.) 41st ACM STOC. pp. 169–178. ACM Press (May / Jun 2009)
21. Goldreich, O.: Candidate one-way functions based on expander graphs. Electronic Colloquium on Computational Complexity (ECCC) **7**(90) (2000)
22. Hawkes, P., Rose, G.G.: Rewriting variables: The complexity of fast algebraic attacks on stream ciphers. In: Franklin, M. (ed.) CRYPTO 2004. LNCS, vol. 3152, pp. 390–406. Springer, Heidelberg (Aug 2004)
23. Hongjun, W.: A lightweight authenticated cipher (2016), https://competitions.cr.yp.to/round3/acornv3.pdf
24. Lin, H., Tessaro, S.: Indistinguishability obfuscation from trilinear maps and block-wise local PRGs. In: Katz, J., Shacham, H. (eds.) CRYPTO 2017, Part I. pp. 630–660. LNCS, Springer, Heidelberg (Aug 2017)
25. Lin, H., Vaikuntanathan, V.: Indistinguishability obfuscation from DDH-like assumptions on constant-degree graded encodings. In: Dinur, I. (ed.) 57th FOCS. pp. 11–20. IEEE Computer Society Press (Oct 2016). https://doi.org/10.1109/FOCS.2016.11
26. Liu, M., Lin, D.: Almost perfect algebraic immune functions with good nonlinearity. In: 2014 IEEE International Symposium on Information Theory. pp. 1837–1841 (June 2014). https://doi.org/10.1109/ISIT.2014.6875151
27. Liu, M., Zhang, Y., Lin, D.: Perfect algebraic immune functions. In: Wang, X., Sako, K. (eds.) ASIACRYPT 2012. LNCS, vol. 7658, pp. 172–189. Springer, Heidelberg (Dec 2012)
28. Méaux, P., Carlet, C., Journault, A., Standaert, F.X.: Improved filter permutators: Combining symmetric encryption design, boolean functions, low complexity cryptography, and homomorphic encryption, for private delegation of computations. Cryptology ePrint Archive, Report 2019/483 (2019), https://eprint.iacr.org/2019/483

29. Méaux, P., Journault, A., Standaert, F.X., Carlet, C.: Towards stream ciphers for efficient FHE with low-noise ciphertexts. In: Fischlin, M., Coron, J.S. (eds.) EUROCRYPT 2016, Part I. LNCS, vol. 9665, pp. 311–343. Springer, Heidelberg (May 2016)
30. Qu, L., Feng, K., Liu, F., Wang, L.: Constructing symmetric boolean functions with maximum algebraic immunity. IEEE Trans. Information Theory pp. 2406–2412 (2009)
31. Qu, L., Li, C., Feng, K.: A note on symmetric boolean functions with maximum algebraic immunity in odd number of variables. IEEE Transactions on Information Theory **53** (2007)
32. R. Geffe, P.: How to protect data with ciphers that are really hard to break. Electronics pp. 99–101 (01 1973)
33. Sarkar, P., Maitra, S.: Balancedness and correlation immunity of symmetric boolean functions. Discrete Mathematics pp. 2351 – 2358 (2007)
34. Tang, D., Carlet, C., Tang, X., Zhou, Z.: Construction of highly nonlinear 1-resilient boolean functions with optimal algebraic immunity and provably high fast algebraic immunity. IEEE Transactions on Information Theory pp. 6113–6125 (Sep 2017). https://doi.org/10.1109/TIT.2017.2725918
35. Tang, D., Carlet, C., Tang, X.: Highly nonlinear boolean functions with optimal algebraic immunity and good behavior against fast algebraic attacks. IEEE Transactions on Information Theory pp. 653–664 (01 2013). https://doi.org/10.1109/TIT.2012.2217476
36. Tang, D., Luo, R., Du, X.: The exact fast algebraic immunity of two subclasses of the majority function. IEICE Transactions pp. 2084–2088 (2016)