On the algebraic immunity of direct sum constructions

Pierrick Méaux

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

Abstract. In this paper, we study sufficient conditions to improve the lower bound on the algebraic immunity of a direct sum of Boolean functions. We exhibit three properties on the component functions such that satisfying one of them is sufficient to ensure that the algebraic immunity of their direct sum exceeds the maximum of their algebraic immunities. These properties can be checked while computing the algebraic immunity and they allow to determine better the security provided by functions central in different cryptographic constructions such as stream ciphers, pseudorandom generators, and weak pseudorandom functions. We provide examples for each property and determine the exact algebraic immunity of candidate constructions.

Keywords: Boolean Functions, Algebraic Immunity, Direct Sum.

1 Introduction.

Since the introduction of algebraic attacks on filtered LFSR ciphers by Courtois and Meier in 2003 [CM03] the algebraic immunity became an important cryptographic criterion for Boolean functions used in stream ciphers. In the context of filtered LFSR the adversary acquires a system of equations $f(L_i(x)) = b_i$ where x is the secret binary vector, f is a known Boolean function, for each i L_i is a public affine transformation and b_i is the binary result. Before the algebraic attacks, a necessary condition on the Boolean functions f was to have high algebraic degree since the system of equations has algebraic degree $\mathsf{deg}(f)$ and solving it allows to determine x. The principle of the algebraic attacks is to find low degree functions not both null g and h such that $fg = h$ hence new equations of shape $h(L_i(x)) = b_i\, g(L_i(x))$ can be created, giving a system with degree $\max(\mathsf{deg}(g), \mathsf{deg}(h))$ instead of $\mathsf{deg}(f)$. It gave birth to the notion of algebraic immunity [MPC04]: for a Boolean function f its algebraic immunity, $\mathsf{AI}(f)$, is defined as the minimum algebraic degree over the functions $g \neq 0$ such that $gf = 0$ for all inputs (or $g(f \oplus 1) = 0$ for all inputs). With this definition, the initial system of equations obtained by the adversary can be converted into a system of equations of algebraic degree $\mathsf{AI}(f)$, and solving it (with Gaussian elimination for example) allows to retrieve x. Since the algebraic immunity is always lower than or equal to the algebraic degree, these attacks lead to forget the criterion of algebraic degree in favor of the algebraic immunity.

Beyond the context of filtered LFSR and stream ciphers, algebraic immunity is an important cryptographic criterion in other contexts. Such kind of algebraic attacks apply each time an adversary can get a system as previously described: with a known Boolean function f, applied only on known affine transformations of the secret vector. Recently, the algebraic immunity has been used to study the security of other cryptographic primitives (sometimes under the name rational degree) such as pseudoradom generators (PRG) of weak pseudorandom functions (WPRF). Revisiting the security of local PRG such as Goldreich's PRG [Gol01], Applebaum and Lovett [AL16] showed that algebraic immunity was one of the main criteria on the predicate to study the security, and since then it is one of the criteria studied for variants of the PRG such as in [GJLS20]. In the recent trend of building conceptually simple WPRF, candidate constructions such as in [BIP+18, BCG+20] fall in the context described above and the algebraic immunity of these functions is discussed to study their security.

The direct sum construction is a standard method to build an $(n+m)$-variable Boolean function from an n-variable and an m-variable Boolean functions. For f defined on $x = x_1, \cdots, x_n$ and g defined on $y = y_1, \cdots, y_m$ different variables, the direct sum $\psi(x, y)$ is defined as $f(x) \oplus g(y)$. It is one of the simplest secondary constructions, methods aiming to built a function (ψ) with prescribed parameters from component functions (f and g) with known parameters. In the domain of Boolean functions used in cryptography this construction can be traced back for example to the first secondary constructions of bent functions [Dil76, Rot76] using bent functions as components. In terms of computations, primitives using direct sums have the advantage of performing the evaluation of f and g independently before applying the sum, which makes it easy to perform in parallel, also for direct sums of more than two components. In terms of cryptographic criteria, the main criteria for Boolean functions used in stream ciphers such as algebraic degree, non-linearity and resiliency order are straightforwardly obtained from the one of f and g. Nevertheless, for the algebraic immunity the parameter of the obtained function is only bounded from above and below:

$$\max(\mathsf{AI}(f), \mathsf{AI}(g)) \leq \mathsf{AI}(f \oplus g) \leq \mathsf{AI}(f) + \mathsf{AI}(g). \tag{1}$$

Examples of functions reaching the lower or the upper bound can be found, and such gap between the two bounds is a drawback when designing cryptographic constructions. As an illustration, taking $n = m$ and f and g two functions with optimal algebraic immunity (that is reaching $\lceil (n+1)/2 \rceil$), the direct sum value can go from simple to double. From a designer point of view, it oscillates between two extremes. If the lower bound is tight then using the direct sum is a waste: one function was already providing the same security. If the upper bound is tight, implementing these two functions in parallel allows to double the degree of the algebraic system targeted by the algebraic attacks, hence squaring the complexity of the attack.

In this work we aim at improving the bounds of Equation 1, finding conditions on f and g to tighten the gap, more particularly to improve the lower bound. The motivation is twofold. First, the direct sum is used in diverse cryptographic constructions *e.g.* [AL16, MJSC16, HKM17, MCJS19, BCG$^+$20] where the algebraic immunity is relevant for the security. A better lower bound can improve the security estimation towards algebraic attacks of candidate constructions, or boost the efficiency by using functions in a lower number of variables for the same security level.Then, since the direct sum is (one of) the simplest secondary constructions and since algebraic immunity is a standard criterion it would be natural to fully understand the behavior of this criterion on this construction. A full characterization of the algebraic immunity of direct sums requires to find properties on f and g sufficient to transform Equation 1 into an equality. We progress on this characterization by determining different properties on the component functions sufficient to improve the lower bound.

1.1 Contributions.

We show how 3 properties on the components functions allow to improve the lower bound, and we give examples of constructions where these improvements apply, focusing on cases where it is sufficient to determine the exact algebraic immunity.

First, we show that the algebraic degree can be sufficient to improve the lower bound by 1. Extending this result to direct sums of more than two components allows to estimate better the algebraic immunity, from the degree of each component. It enables us to determine the exact algebraic immunity of the candidate WPRF of [BCG$^+$20], and provide a simpler proof for the functions used in [MJSC16].

Then, we prove that the value of the difference $\mathit{\Delta}_{\mathsf{AN}}(f)$ between the smallest degree of functions $g \neq 0$ such that $fg = 0$ and of functions $h \neq 0$ such that $h(f \oplus 1) = 0$ is a criterion allowing to improve the bound. We exhibit conditions where the algebraic immunity of the direct sum is improved by at least the quantity $\mathit{\Delta}_{\mathsf{AN}}(f)$, giving the new upper bound:

$$\mathsf{AI}(\psi) \geq \max\left(\mathsf{AI}(f) + \min\{\mathit{\Delta}_{\mathsf{AN}}(f), \mathsf{AI}(g)\}, \mathsf{AI}(g) + \min\{\mathit{\Delta}_{\mathsf{AN}}(g), \mathsf{AI}(f)\}\right), \tag{2}$$

and we exhibit constructions for which this result gives the exact algebraic immunity.

Finally, extending the property on the degree, we show that the invertibility of a matrix defined by the higher degree coefficients of one of the component functions allows to improve the bound. We study the *AI increasing functions* having such property, and provide a family of such functions. We combine the two approaches on the $\mathit{\Delta}_{\mathsf{AN}}(f)$ and on the AI increasing functions, it gives the best improvement on the lower bound of Equation 1 from our study, that we state in Theorem 1.

1.2 Related works.

The general bounds of Equation 1 appears in many works, with examples reaching the extremes. We highlight the progresses on the upper bound which are complementary to our results towards the full characterization. In [BP05] the degree is taken into account to improve the upper bound to:

$$\mathsf{AI}(f \oplus g) \leq \min\left(\max\left[\mathsf{deg}(f), \mathsf{deg}(g)\right], \mathsf{AI}(f) + \mathsf{AI}(g)\right).$$

In [CM20] necessary conditions are studied to build direct sums with optimal algebraic immunity, it provides families reaching the upper bound.

The role of the difference $\mathit{\Delta}_{\mathsf{AN}}(f)$ has been studied previously without naming this quantity. In [Riz10], in addition to the algebraic immunity criterion the author studies the impact of the *complementary algebraic immunity* defined as the maximum between two quantities: the minimum degree over the functions $g \neq 0$ such that $fg = 0$ and the minimum degree over the functions $h \neq 0$ such that $h(f \oplus 1) = 0$. Accordingly, $\mathit{\Delta}_{\mathsf{AN}}(f)$ is the difference between the complementary AI and the AI. Rizomiliotis shows that the $\mathit{\Delta}_{\mathsf{AN}}(f)$ quantity allows to improve the known lower bounds on the r-th order nonlinearity of f. When $\mathit{\Delta}_{\mathsf{AN}}(f) = 0$ either the bound of [Car06] or of [Mes08] are tight (depending on r), and when $\mathit{\Delta}_{\mathsf{AN}}(f) \neq 0$ it provides a better bound than the former ones. In [CM20], towards proving the cryptographic parameters of particular direct sums the authors prove that when the AI and the complementary AI of a function differ, the direct sum with a non-constant function increases the AI. With our definitions it means a non-null $\mathit{\Delta}_{\mathsf{AN}}(f)$ is sufficient to improve the lower bound of Equation 1, and it corresponds to a particular case of our new bound given in Equation 2.

1.3 Paper organization.

The article is organized in the following way: In Section 2 we give the notations and properties of Boolean functions and cryptographic criteria used in the rest of the paper. Section 3 is dedicated to the improvement from the degree and applications to multiple direct sums. The results obtained from the $\mathit{\Delta}_{\mathsf{AN}}()$ are presented in Section 4. In Section 5 we combine both approaches, giving the main theorem and a study on AI increasing functions. Section 6 concludes on the results from this work and the open questions.

2 Preliminaries.

For readability we use the notation $+$ instead of $\oplus$ to denote addition in $\mathbb{F}_2$, and $[n]$ to denote $\{1, \ldots, n\}$ and more generally $[a, b]$ for the set of integers c such that $a \leq c \leq b$. For a binary vector a we denote $\mathsf{w_H}(a)$ its Hamming weight. $\log$ refers to the logarithm in basis 2.

2.1 Boolean functions and cryptographic criteria.

We recall the definition of Boolean function and representations we will use in the following sections. We recall the cryptographic criterion of algebraic immunity, or AI, at the center of this paper, together with less usual notations relatively to the annihilators of a Boolean function. For further backgrounds on Boolean functions used in cryptography, or connections between the AI criterion and and other cryptographic criteria, we refer the reader to [Car21].

Definition 1 (Boolean function). *A Boolean function f in n variables (an n-variable Boolean function) is a function from $\mathbb{F}_2^n$ to $\mathbb{F}_2$. The set of all Boolean functions in n variables is denoted by $\mathcal{B}_n$.*

The following representation is commonly used, and its basic properties also.

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 $\mathbb{F}_2[x_1,\dots,x_n]/(x_1^2+x_1,\dots,x_n^2+x_n)$):*

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

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

- *The algebraic degree of f equals the global degree of its ANF:* $\mathsf{deg}(f) = \max_{\{I \mid a_I=1\}} |I|$ *(with the convention that* $\mathsf{deg}(0) = -\infty$*).*
- *Any term $\prod_{i\in I} x_i$ in such an ANF is called a monomial and its degree equals $|I|$. A function with only one non-zero coefficient a_I, where I is non-empty, is called a monomial function.*

We will also use the following generalization of the ANF:

Definition 3 (Partitioned Algebraic Normal Form ([CM20])). *We call (n,m)-Partitioned Algebraic Normal Form of an $(n+m)$-variable Boolean function f its polynomial representation over $\mathbb{F}_2$ (i.e. belonging to $\big(\mathbb{F}_2[x_1,\dots,x_n]/(x_1^2+x_1,\dots,x_n^2+x_n)\big)[y_1,\dots,y_m]/(y_1^2+y_1,\dots,y_m^2+y_m)$):*

$$f(x,y) = \sum_{I\subseteq[m]} a_I(x_1,\dots,x_n) \left(\prod_{i\in I} y_i\right) = \sum_{I\subseteq[m]} a_I(x)\, y^I,$$

where $a_I \in \mathbb{F}_2[x_1,\dots,x_n]/(x_1^2+x_1,\dots,x_n^2+x_n)$ and $x = (x_1,\dots,x_n)$.

We call partitioned-(n,m)-ANF coefficients the coefficients a_I, or simply PANF coefficients when n and m are clearly identified.

Our study focuses on the cryptographic criterion of 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:*

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

where $\mathsf{deg}(g)$ *is the algebraic degree of g. The function g is called an annihilator of f (or $f+1$). We additionally use the notation* $\mathsf{AN}(f)$ *for the minimum algebraic degree of non-null annihilators of f, and* $\Delta_{\mathsf{AN}}(f)$ *for the difference between* $\mathsf{AN}(f)$ *and* $\mathsf{AN}(f+1)$*:*

$$\mathsf{AN}(f) = \min_{g\neq 0}\{\mathsf{deg}(g) \mid fg = 0\}, \quad \Delta_{\mathsf{AN}}(f) = |\mathsf{AN}(f) - \mathsf{AN}(f+1)|.$$

Note that, for every Boolean function f, the functions f and $f+1$ are mutual annihilators, and:

Property 1 (Algebraic immunity's properties)**.**

- *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:* $\mathsf{AI}(f)\leq\mathsf{AN}(f)\leq\mathsf{deg}(f)$.

2.2 Direct sum construction, families of functions.

The direct sum construction is a secondary construction, allowing to obtain a function with targeted parameters from two (or more) component functions with already known parameters.

Definition 5 (Direct Sum). *Let f be a Boolean function of n variables and g a Boolean function of m variables, f and g depending on distinct variables, the direct sum ψ of f and g is defined by:*

$$\psi(x,y)=f(x)+g(y),\quad \text{where } x\in\mathbb{F}_2^n \text{ and } y\in\mathbb{F}_2^m.$$

We note $\psi=\mathsf{DS}(f,g)$, and extend the notation for direct sums of $t>2$ functions f_1 to f_t as $\mathsf{DS}(f_1,\dots,f_t)$ and $\mathsf{DS}^t(f)$ when $f_1=\ldots=f_t$.

Lemma 1 (Direct sum and algebraic immunity). *Let $n,m\in\mathbb{N}$, $f\in\mathcal{B}_n$ and $g\in\mathcal{B}_m$, if $\psi=\mathsf{DS}(f,g)$ then:* $\max(\mathsf{AI}(f),\mathsf{AI}(g))\leq\mathsf{AI}(\psi)\leq\mathsf{AI}(f)+\mathsf{AI}(g)$.

Lemma 2 (Annihilators of direct sums and PANF coefficients ([CM20], Lemma 10)). *Let f be a Boolean function in the variables $x_1,\dots x_n$ and g be a Boolean function in the variables $y_1,\dots y_m$. Let $\psi=\mathsf{DS}(f,g)$, $\varepsilon\in\{0,1\}$, and h a function in $x_1,\dots,x_n,y_1,\dots,y_m$ with (n,m)-partitioned algebraic normal form: $h(x,y)=\sum_{I\subseteq[m]}h_I(x)y^I$. If h is an annihilator of $\psi+\varepsilon$ then the following relation holds on its PANF coefficients:*

$$\forall I\subseteq[m],\quad h_I(x)\left(f+\varepsilon+\sum_{J\subseteq I}g_J\right)=\sum_{J\subsetneq I}h_J(x)\sum_{K\subseteq J}g_{K\cup\{I\setminus J\}},\tag{3}$$

where the coefficients g_I correspond to the (standard) ANF coefficients of g.

We define two families of Boolean functions that we will often use for concrete examples of constructions where our results allow to improve on the lower bound of Lemma 1. First, the direct sum of monomials family consists in functions obtained by iterating the direct sum construction on monomials, giving functions with very sparse ANF. This family has been introduced in [MJSC16] in the context of homomorphic transciphering [NLV11], as Boolean functions efficient to evaluate by fully homomorphic encryption schemes.

Definition 6 (Direct sum of monomials & direct sum vector [MJSC16]). *Let f be a Boolean function of n variables, we call f a Direct Sum of Monomials (or DSM) if the following holds for its ANF: $\forall(I,J)$ such that $a_I=a_J=1$, $I\cap J\in\{\emptyset,I\cup J\}$.*

Let f a DSM, we define its direct sum vector: $\mathbf{m}_f=[m_1,m_2,\dots,m_k]$ of length $k=\mathsf{deg}(f)$, where m_i is the number of monomials of degree i of f: for $i>0$, $m_i=|\{a_I=1,\ \text{such that } |I|=i\}|$.

A sub-family of particular interest of DSM is the family of triangular functions:

Definition 7 (Triangular functions [MJSC16]). *Let $k \in \mathbb{N}^*$. The k-th triangular function T_k is the following direct sum of monomials of $k(k+1)/2$ variables:*

$$T_k(x_1, \dots, x_{k(k+1)/2}) = \sum_{i=1}^{k} \prod_{j=1}^{i} x_{j+i(i-1)/2}.$$

It can also be defined from its direct sum vector which is the all-1 vector of length k: $\mathbf{m}_{T_k} = [1, 1, \dots, 1]$.

Lemma 3 (Algebraic immunity of DSM ([CM19], Theorem 1)). *Let $f \in \mathcal{B}_n$ if f is a direct sum of monomials with associated direct sum vector $\mathbf{m}_f = [m_1, \dots, m_k]$, then $\mathsf{AI}(f) = \min_{0 \le d \le k} \left(d + \sum_{i=d+1}^{k} m_i\right)$.*

The second family of Boolean functions we will use to illustrate our results are the threshold functions. These functions are symmetric, the output of f is identical on inputs with the same Hamming weight. A threshold function of threshold d gives 1 only on inputs with Hamming weight at least d, it is a generalization of majority functions (where the threshold is $n/2$) which have been one the first family of Boolean functions known for their optimal algebraic immunity [BP05, DMS06]. Threshold functions are easy to compute and most of their relevant cryptographic parameters have been studied (AI, nonlinearity and resiliency in [CM19, CM20], fast algebraic immunity in [Méa20]). We will use that this family contains element with any possible value of $\Delta_{\mathsf{AN}}(f)$.

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

$$\forall x = (x_1, \dots, x_n) \in \mathbb{F}_2^n, \quad \mathsf{T}_{d,n}(x) = \begin{cases} 0 & \textit{if } \mathsf{w_H}(x) < d, \\ 1 & \textit{otherwise.} \end{cases}$$

Lemma 4 (Algebraic immunity and annihilators of threshold functions, ([CM19], Lemma 6)). *Let $n > 0$ and $1 \le d \le n$. The threshold function $\mathsf{T}_{d,n}$ has the following algebraic criteria:*

$\mathsf{AI}(\mathsf{T}_{d,n}) = \min(d, n-d+1)$, $\mathsf{AN}(\mathsf{T}_{d,n}) = n-d+1$, $\mathsf{AN}(1+\mathsf{T}_{d,n}) = d$, *and* $\Delta_{\mathsf{AN}}(f) = |n+1-2d|$.

3 Improvement from the degree

In this part we show that the algebraic degree of one of the two components of a direct sum is sufficient to improve on the lower bound of Lemma 1. We generalize this result to the direct sum of multiple functions, which allows to study the algebraic immunity of some weak PRF candidates of [BCG$^+$20] in Section 3.1 and give a simple proof for the algebraic immunity of DSM in Section 3.2.

Lemma 5. *Let $n, m \in \mathbb{N}^*$, let $f \in \mathcal{B}_n$, $g \in \mathcal{B}_m$ and $\psi = \mathsf{DS}(f, g)$. If $\mathsf{AI}(f) < \mathsf{deg}(g)$ then $\mathsf{AI}(\psi) > \mathsf{AI}(f)$.*

Proof. We prove the result by contradiction, showing that $\mathsf{AI}(\psi) \le \mathsf{AI}(f)$ is impossible. First, Lemma 1 shows the impossibility if $\mathsf{AI}(f) < \mathsf{AI}(g)$, hence we assume $\mathsf{AI}(f) \ge \mathsf{AI}(g)$ in the following, which implies $\mathsf{AI}(\psi) \ge \mathsf{AI}(f)$ by Lemma 1 hence $\mathsf{AI}(\psi) = \mathsf{AI}(f)$. Let h be the annihilator of $\psi + \varepsilon$ ($\varepsilon \in \{0, 1\}$) defining the algebraic immunity, then Lemma 2 gives relations on the PANF coefficients of h and since $\mathsf{deg}(h) = \mathsf{AI}(f)$ for all $I \subseteq [m]$ we have the relation: $\mathsf{deg}(h_I(x)) \le \mathsf{AI}(f) - |I|$. Since h is not null, at least one of the $h_I(x)$ is not null. Let us denote I_0 a set such that for all I such that $|I| < |I_0|$, $h_I(x) = 0$, and $h_{I_0}(x) \neq 0$. Then Equation 3 for I_0 gives $h_{I_0}(x)(f + \varepsilon + \sum_{J \subseteq I_0} g_J) = 0$. Hence, $h_{I_0}(x)$ annihilates either f or $f+1$, and the only possibility is therefore $I_0 = \emptyset$ and $\mathsf{deg}(h_\emptyset(x)) = \mathsf{AI}(f)$.

Let us denote $d = \mathsf{deg}(g)$, since $\mathsf{deg}(g) > \mathsf{AI}(f)$ there exists a set $I_d \subseteq [m]$ such that $|I_d| = d$ and $g_{I_d} = 1$. $h_{I_d}(x)$ is forced to be null (otherwise $\mathsf{deg}(h) > \mathsf{AI}(f)$), and Equation 3 for I_d leads to:

$$0 = h_\emptyset(x) + \sum_{\emptyset \subsetneq J \subsetneq I_d} h_J(x) \sum_{K \subseteq J} g_{K \cup \{I_d \setminus J\}}.$$

Since $\mathsf{deg}(h_\emptyset(x)) = \mathsf{AI}(f)$ and $\forall J$ such that $\emptyset \subsetneq J \subseteq [m]$ we have $\mathsf{deg}(h_J(x)) < \mathsf{AI}(f)$, we get a contradiction. It allows to conclude, $\mathsf{AI}(\psi) > \mathsf{AI}(f)$.

□

Lemma 5 can be used iteratively to bound the AI of the direct sum of multiple functions as summarized in the following lemma. In particular, iterating the direct sum of the same function allows the algebraic immunity to reach the algebraic degree.

Lemma 6. *Let $t \in \mathbb{N}^*$, and $f_1, \ldots, f_t$ be t Boolean functions, if for $r \in [t]$ there exists r different indexes $i_1, \cdots, i_r$ of $[t]$ such that $\forall j \in [r], \mathsf{deg}(f_{i_j}) \geq j$ then $\mathsf{AI}(\mathsf{DS}(f_1, \ldots, f_t)) \geq r$.*

Proof. We denote $\psi_1 = f_{i_1}$ and then for $j \in [2, r]$ we denote $\psi_j = \mathsf{DS}(\psi_{j-1}, f_{i_j})$, in these terms $\psi_r = \mathsf{DS}(f_{i_1}, \ldots, f_{i_r})$, we prove by recurrence that $\mathsf{AI}(\psi_j) \geq j$. For the initialization step, $\mathsf{AI}(\psi_2) \geq \max(\mathsf{AI}(f_{i_1}), \mathsf{AI}(f_{i_2}))$ from Lemma 1, hence $\mathsf{AI}(f_{i_1}) \geq 2$ or $\mathsf{AI}(f_{i_1}) \geq 2$ imply $\mathsf{AI}(\psi_2) \geq 2$. The only remaining possibility is $\mathsf{AI}(f_{i_1}) \leq 1$ and $\mathsf{AI}(f_{i_2}) \leq 1$, and since $\mathsf{deg}(f_{i_2}) \geq 2$ we can apply Lemma 5, giving $\mathsf{AI}(\psi_2) > \mathsf{AI}(f_{i_1})$, and therefore $\mathsf{AI}(\psi_2) \geq 2$ since $\mathsf{AI}(f_{i_1}) \geq 1$. It concludes the initialization step.

For the recurrence step, we focus on the algebraic immunity of $\psi_{j+1} = \mathsf{DS}(\psi_j, f_{i_{j+1}})$. Using the recurrent hypothesis, $\mathsf{AI}(\psi_j) \geq j$, if $\mathsf{AI}(\psi_j) \geq j+1$ it implies $\mathsf{AI}(\psi_{j+1}) \geq j+1$, otherwise $\mathsf{AI}(\psi_j) = j$ and since $\mathsf{deg}(f_{i_{j+1}}) \geq j+1$ Lemma 5 gives $\mathsf{AI}(\psi_{j+1}) \geq j+1$ also in this case. It concludes the recurrence: $\mathsf{AI}(\psi_r) \geq r$. Since $\mathsf{DS}(f_1, \ldots, f_t)$ is the direct sum of ψ_r and other functions, Lemma 1 it allows to conclude $\mathsf{AI}(\mathsf{DS}(f_1, \ldots, f_t)) \geq \mathsf{AI}(\psi_r) \geq r$.

□

Corollary 1. *Let $d \in \mathbb{N}$, $d \geq 2$, let $f_1, \ldots, f_d$ Boolean functions, if $\forall i \in [d]$ $\mathsf{deg}(f_i) = d$ then $\mathsf{AI}(\mathsf{DS}(f_1, \ldots, f_d)) = d$.*

Proof. We denote $\psi = \mathsf{DS}(f_1, \ldots, f_d)$. Applying Lemma 6 on the indexes $1, \ldots, d$ gives $\mathsf{AI}(\psi) \geq d$, and since $\mathsf{deg}(\psi) = d$ it gives $\mathsf{AI}(\psi) \leq d$, allowing to conclude $\mathsf{AI}(\psi) = d$. □

3.1 Determining the algebraic immunity of [BCG$^+$20] WPRF candidates

In [BCG$^+$20] Boyle *et al.* introduce and study the primitive of pseudorandom correlated function. They give efficient constructions based on candidate weak pseudorandom functions (WPRF) from the class of complexity depth-2 $\mathsf{AC}^0[\oplus]$ (XOR of conjunctions of input variables and their negations). Such candidates WPRF can be written as $f_K(x)$, a family of $|x|$-variable Boolean functions indexed by K a binary key. In the following we recall the candidate weak PRF families, and prove the algebraic immunity of f_K when it was left as a conjecture or open.

We summarize the WPRF candidates of [BCG$^+$20] and the knowledge on their algebraic immunity in Table 1. The results on f_K^1 and f_K^3 were obtained using that in both case the function can be written as $\mathsf{DS}(f, g)$ where g is affine equivalent to T_D: since $\mathsf{AI}(T_D) = D$ ([MJSC16] Lemma 6, or as a particular case of Lemma 3), applying Lemma 1 gives the lower bound. We conclude on the algebraic immunity of the four families of Boolean functions in Proposition 1.

$f_K(x)$	AI status
$K \in \mathbb{F}_2^{\omega D(D+1)/2},\ f_K^1 := \bigoplus_{i=1}^{D} \bigoplus_{j=1}^{\omega} \bigwedge_{k=1}^{i} (x_{i,j,k} \oplus K_{i,j,k})$	$\mathsf{AI}(f_K^1) \geq D$, proven
$K \in \mathbb{F}_2^{\omega D(D+1)/2},\ f_K^2 := \bigoplus_{i=1}^{D} \bigoplus_{j=1}^{\omega} \bigwedge_{k=1}^{i} (x_{j,k} \oplus K_{i,j,k})$	$\mathsf{AI}(f_K^2) = D$, conjectured
$K \in \mathbb{F}_2^{\omega D(D+1)/2+D},\ f_K^3 := (\bigoplus_{i=1}^{D} \bigoplus_{j=1}^{\omega} \bigwedge_{k=1}^{i} (x_{j,k} \oplus K_{i,j,k})) \oplus T_D(y \oplus K_y)$	$\mathsf{AI}(f_K^3) \geq D$, proven
$K \in \mathbb{F}_2^{\omega D},\ f_K^4 := \bigoplus_{i=1}^{D} \bigoplus_{j=1}^{\omega} \bigwedge_{k=1}^{i} (x_{j,k} \oplus K_{j,k})$	open

Table 1. Candidate WPRF of [BCG$^+$20] and algebraic immunity. $f_K(x)$ refers to the description of the WPRF family, and "AI status" to the results on its algebraic immunity. For the four candidates $D, \omega \in \mathbb{N}^*$, $D < \omega$.

Proposition 1. *Let $D, \omega \in \mathbb{N}^*$, $D < \omega$, and for $i \in [4]$ f_K^i the Boolean function defined in Table 1, then $\mathsf{AI}(f_K^i) = D$.*

Proof. First, note that all these functions have algebraic degree D for any choice of K, therefore with AI at most d, which already allows to conclude for the cases f_K^1 and f_K^3 since these functions are direct sums with one component being the triangular function T_D (Lemma 1 and Lemma 3). Then, f_K^2 and f_K^4 are both obtained by the direct sum of ω functions of degree D. In both cases we can rewrite the function as $\mathsf{DS}(f, g)$ where f is the part containing the sum over the indexes $j \in [D]$ and g the indexes $j \in [D+1, \omega]$. Applying Corollary 1 $\mathsf{AI}(f) = D$, and combining it with Lemma 1: $\mathsf{AI}(\mathsf{DS}(f, g)) \geq D$ which allows to conclude $\mathsf{AI}(f_k^2) = \mathsf{AI}(f_K^4) = D$. □

3.2 A simpler proof for the AI of DSM

Using Lemma 6 we can derive a lower bound on the algebraic immunity of any direct sum of t functions based on the degree of its components, similarly to the formula of Lemma 3.

Proposition 2. *Let $t \in \mathbb{N}^*$, $f_1, \ldots, f_t$ t Boolean functions, and $\psi = \mathsf{DS}(f_1, \ldots, f_t)$. We note $d_i = |\{f_j \mid \deg(f_j) = i,\ j \in [t]\}|$ and $k = \max\{\deg(f_i) \mid i \in [t]\}$, the following bound applies for ψ:*

$$\mathsf{AI}(\psi) \geq \min_{0 \leq i \leq k} \left(i + \sum_{j>i} d_j \right).$$

Proof. First, in the particular case $k \leq 0$, all functions are constants, it gives $\mathsf{AI}(\psi) = 0$ (Property 1 item 1). Then, for the rest of the proof we can assume that at least on function is not constant. Let $e = \min_{0 \leq i \leq k}(i + \sum_{j>i} d_j)$ and let $i^* \in [0, k]$ an integer such that $e = i^* + \sum_{j>i^*} d_j$ (by definition at least on integer fulfills this property). We show that e different indexes can be taken to apply Lemma 6.

First, for all $i \in [0, i^* - 1]$ by definition of i^* we get the inequality $i + \sum_{j>i} d_j \geq i^* + \sum_{j>i^*} d_j$ and therefore $\forall i \in [0, i^* - 1],\ i^* - i \leq \sum_{j>i}^{i^*} d_j$, which guarantees the existence of i^* different indexes $i_1, \ldots, i_{i^*}$ in $[t]$ such that $\forall j \in [i^*]\ \deg(f_{i_j}) \geq j$. Then, for all $i \in [i^* + 1, k]$ by definition of i^* we get the inequality $i + \sum_{j>i} d_j \geq i^* + \sum_{j>i^*} d_j$ and therefore $\forall i \in [i^* + 1, k],\ i - i^* \geq \sum_{j>i^*}^{i} d_j$. Since there are exactly $s = \sum_{j>i^*}^{k} d_j$ functions of degree in $[i^* + 1, k]$ the later property guarantees the existence of s different indexes $j_1, \ldots, j_s$ in $[t]$ such that $\forall k \in [s]\ \deg(f_{j_k}) \geq i^* + k$.

Finally, since the indexes i_j are used for functions of degree at most i^* and the indexes j_k for functions of degree at least $i^* + 1$ all these indexes are different and we can apply Lemma 6: $\mathsf{AI}(\psi) \geq e$.

□

We determine sufficient conditions for the bound of Proposition 2 to be an equality, and we remark that DSM functions are one of these cases.

Corollary 2. *Let $t \in \mathbb{N}^*$, $f_1, \dots, f_t$ t Boolean functions, and $\psi = \mathsf{DS}(f_1, \dots, f_t)$. We denote $d_i = |\{f_j \mid \mathsf{deg}(f_j) = i, j \in [t]\}|$, $k = \max\{\mathsf{deg}(f_i) \mid i \in [t]\}$, $e = \min_{0 \le i \le k} \Big(i + \sum_{j>i} d_j\Big)$, and $i^* = \max\{i \in [0,k] \mid (i + \sum_{j>i} d_j) = e\}$.*

If $\forall f \in \{f_j \mid \mathsf{deg}(f_j) > i^*, j \in [t]\}$ $\mathsf{AI}(f) = 1$ then $\mathsf{AI}(\psi) = e$.

Proof. First, for the particular case $k \le 0$, all functions are constants and $\mathsf{AI}(\psi) = 0$ as seen in the proposition, therefore we can assume in the following that at least on the function f_i is not constant. Proposition 2 gives $\mathsf{AI}(\psi) \ge e$, and the condition on the functions of degree greater than i^* allows to prove the existence of an annihilator of degree e of ψ or $\psi + 1$. More precisely, let denote $S_{\le} = \{f_j \mid \mathsf{deg}(f_j) \le i^*, j \in [t]\}$ and $S_{>} = \{f_j \mid \mathsf{deg}(f_j) > i^*, j \in [t]\}$ and similarly $\psi_{\le} = \mathsf{DS}(f \in S_{\le})$, $\psi_{>} = \mathsf{DS}(f \in S_{>})$. We separate the cases based on the value of i^*.

If $i^* = k$, then $e = k$, $\psi_{\le} = \psi$ hence $1 + \psi_{\le}$ is a degree e annihilator of ψ. If $i^* \in [k-1]$, since $\psi_{\le}$ has degree i^* (direct sum of functions of degree at most i^* and at least one has degree i^* otherwise the minimum of $i + \sum_{j>i} d_j$ cannot be reached in i^*), the function $1 + \psi_{\le}$ is an annihilator of degree i^* of $\psi_{\le}$. Then, for all element either $f \in S_{>}$ f or $f + 1$ has an annihilator of degree 1, hence the product g of these $e - i^*$ annihilators is a degree $e - i^*$ annihilator of $\psi_{>}$ or $1 + \psi_{>}$. Finally, since $\psi = \mathsf{DS}(\psi_{\le}, \psi_{>})$, the function $g \cdot (1 + \psi_{\le})$ is a degree e annihilator of ψ. If $i^* = 0$, we build an annihilator g of $\psi_{>}$ as in the former case, and therefore g is a degree e annihilator of ψ or $\psi + 1$.

□

Remark 1. Since monomial functions have algebraic immunity 1 (as recalled in Property 1), DSM fulfill the conditions of Corollary 2 therefore it implies Lemma 3.

4 Improvement from the difference of annihilators of minimal degree

In this section we exhibit conditions where the value of $\Delta_{\mathsf{AN}}(f)$ is sufficient to improve the lower bound of Lemma 1 for $\mathsf{AI}(\mathsf{DS}(f,g))$. Then in Section 4.1 we use these results to determine the AI of direct sums of threshold functions.

First we give another characterization of $\mathsf{AN}(f)$.

Lemma 7. *Let $n \in \mathbb{N}^*$, $k \in \mathbb{N}$, $D = \sum_{i=0}^{k} \binom{n}{i}$ and $f \in \mathcal{B}_n$. $\mathsf{AN}(f) > k$. is equivalent to: for each D-uple $(g_I)_{I \subseteq [n], 0 \le |I| \le k}$ in $\mathbb{F}_2^D \setminus \{0\}$ there exists at least one $I' \in [n]$ such that:*

$$\sum_{\substack{J \subseteq I' \\ 0 \le |J| \le k}} g_J \sum_{K \subseteq J} f_{K \cup \{I' \setminus J\}} = 1. \tag{4}$$

Proof. Let $g, h \in \mathcal{B}_n$ such that $fg = h$, we focus on the relations of the ANF coefficients f_I, g_I, and h_I where $I \in [n]$. We get:

$$\left(\sum_{I \subseteq [n]} f_I x^I\right) \left(\sum_{I \subseteq [n]} g_I x^I\right) = \sum_{I \subseteq [n]} h_I x^I,$$

developing the product and identifying we obtain:

$$\forall I \subseteq [n], \quad h_I = \sum_{J \subseteq I} g_J \sum_{K \subseteq J} f_{K \cup \{I \setminus J\}}.$$

If $\mathsf{AN}(f) > k$, then for all non-null function g of degree at most k we have $fg \neq 0$, which implies that at least one ANF coefficient of $h = fg$ is not null in this case. Since the functions of $\mathcal{B}_n$ of degree at most k are the 2^D function with ANF coefficients null for all subsets of cardinal greater than k, and the null function is the one with all ANF coefficients null, it gives:

$$\forall (g_I)_{I\subseteq[n], 0\leq|I|\leq k} \in \mathbb{F}_2^D \setminus \{0\},\ \exists I' \in [n] \mid \sum_{\substack{J\subseteq I' \\ 0\leq|J|\leq k}} g_J \sum_{K\subseteq J} f_{K\cup\{I'\setminus J\}} = 1.$$

If for a binary D-uple $(g_I)_{I\subseteq[n], 0\leq|I|\leq k}$ there exists at least one $I' \in [n]$ satisfying Equation 4, then the function g with ANF coefficients given by the D-uple for $|I| < k$ and 0 elsewhere is a function of degree at most k such that the product with f is not null. Since the $2^D - 1$ non-null D-uple are in bijection with the $2^D - 1$ non-null functions of degree at most k, the property holding for each one of the D-uples gives that no (non-null) function of degree at most f annihilates f, *i.e.* $\mathsf{AN}(f) > k$.

□

Using the PANF relations and the characterization of Lemma 7 allows to prove an improvement up to $\Delta_{\mathsf{AN}}(f)$ on the AI of the direct sum:

Lemma 8. *Let $n, m \in \mathbb{N}^*$, $k \in \mathbb{N}$, $f \in \mathcal{B}_n$ and $g \in \mathcal{B}_m$, if $\mathsf{AI}(g) > k$ and $\Delta_{\mathsf{AN}}(f) > k$ then $\mathsf{AI}(\mathsf{DS}(f,g)) > \mathsf{AI}(f) + k$.*

Proof. Without loss of generality we choose f such that $\mathsf{AN}(f+1) > \mathsf{AN}(f) + k$. We do the proof by contradiction, assuming there exists a function $h \in \mathcal{B}_{n+m}$ non-null of degree at most $\mathsf{AI}(f) + k$ such that $h(f+g) = 0$ or $h(f+g+1) = 0$. Note that we can restrict our study to the case $h(f+g) = 0$ since the case $h(f+g+1) = 0$ can be written as $h(f+g') = 0$ where $g' = g+1$ is still an m-variable function of algebraic immunity greater than k.

Applying Lemma 2, we obtain the following relations:

$$\forall I \subseteq [m], \quad h_I(x)f = \sum_{J\subseteq I} h_J(x) \sum_{K\subseteq J} g_{K\cup\{I\setminus J\}}. \tag{5}$$

The constraint $\mathsf{deg}(h) \leq \mathsf{AI}(f) + k$ leads to $\mathsf{deg}(h_I(x)) \leq \mathsf{AN}(f) + k - |I|$ for all $I \subseteq [m]$. The right hand side of Equation 5 is an annihilator of $f+1$ (for all $I \subseteq [m]$), since the $h_I(x)$ have degree at most $\mathsf{AN}(f) + k$ and $\mathsf{AN}(f+1) > \mathsf{AN}(f) + k$ it forces the right hand side to be null for all the equations.

For $|I| > k$ all $h_I(x)$ are null due to the degree constraint, therefore all the right hand sides are combinations of the PANF coefficients related to subset of cardinal at most k. Hence we can rewrite Equation 5 as:

$$\forall I \subseteq [m], \quad h_I(x)f = \sum_{\substack{J\subseteq I \\ 0\leq|J|\leq k}} h_J(x) \sum_{K\subseteq J} g_{K\cup\{I\setminus J\}} = 0. \tag{6}$$

Since h is non-null there exists at least one $z = (x', y') \in \mathbb{F}_2^{n+m}$ such that $h(z) = 1$ and therefore at least one $J \subseteq [m]$ such that $h_J(x') = 1$. Hence for this particular $x' \in \mathbb{F}_2^n$ Equation 6 leads to:

$$\forall I \subseteq [m], \qquad \sum_{\substack{J\subseteq I \\ 0\leq|J|\leq k}} h_J(x') \sum_{K\subseteq J} g_{K\cup\{I\setminus J\}} = 0,$$

where the binary values $h_J(x')$ are not all null. Since $\mathsf{AI}(g) > k$ applying Lemma 7 we know that there is at least one subset $I' \subseteq [m]$ such that the equation is not satisfied, which gives the contradiction. We can conclude $\mathsf{AI}(\mathsf{DS}(f,g)) > \mathsf{AI}(f) + k$.

□

Remark 2. For $k = 0$, Lemma 8 is equivalent to: $g \in \mathcal{B}_m$ not constant and $f \in \mathcal{B}_n$ such that $\mathsf{AN}(f) \neq \mathsf{AN}(f+1)$ implies $\mathsf{AI}(\mathsf{DS}(f,g)) > \mathsf{AI}(f)$, which is the result of Lemma 11 in [CM20].

Lemma 8 directly gives a more precise lower bound than Lemma 1.

Corollary 3. *Let $n, m \in \mathbb{N}^*$, $f \in \mathcal{B}_n$, $g \in \mathcal{B}_m$, and $\psi = \mathsf{DS}(f,g)$, the following bound holds on its algebraic immunity:*

$$\mathsf{AI}(\psi) \geq \max\left(\mathsf{AI}(f) + \min\{\Delta_{\mathsf{AN}}(f), \mathsf{AI}(g)\}, \mathsf{AI}(g) + \min\{\Delta_{\mathsf{AN}}(g), \mathsf{AI}(f)\}\right).$$

Note that if $\mathsf{AI}(g) \leq \Delta_{\mathsf{AN}}(f)$ of $\mathsf{AI}(f) \leq \Delta_{\mathsf{AN}}(g)$ the lower bound of Corollary 3 reaches the upper bound of Lemma 1, giving exactly the algebraic immunity of ψ.

4.1 Algebraic immunity of direct sums of threshold functions

We study direct sums of functions from the family of threshold functions to illustrate cases where the results of this section improve upon Lemma 1. We use that for all positive integer t we can find values of d and n such that $\Delta_{\mathsf{AN}}(\mathsf{T}_{d,n}) = t$ using Lemma 4. Threshold functions are typical examples of functions with potentially high $\Delta_{\mathsf{AN}}()$, hence allowing to build functions with prescribed AI using Corollary 3.

Proposition 3. *Let $t \in \mathbb{N}^*$, the families of functions indexed by $d \in \mathbb{N}^*$ defined as $f_d = \mathsf{T}_{d,2d+t-1}$ and $g_d = \mathsf{DS}(f_d, \mathsf{T}_{t,2t-1})$ are such that for all d $\Delta_{\mathsf{AN}}(f_d) = t$ and $\mathsf{AI}(g_d) = d + t$.*

Proof. $\Delta_{\mathsf{AN}}(f_d) = t$ comes directly from the $\Delta_{\mathsf{AN}}()$ of threshold functions (Lemma 4). Using the same lemma, $\mathsf{AI}(\mathsf{T}_{t,2t-1}) = t$, hence applying Corollary 3 on f_d and $\mathsf{T}_{t,2t-1}$ gives $\mathsf{AI}(g_d) \geq d + t$. Since using Lemma 4 $\mathsf{AN}(1 + f_d) = d$ and $\mathsf{AN}(1 + \mathsf{T}_{t,2t-1}) = t$, it allows to determine non-null annihilators of their direct sum of degree $d + t$ hence $\mathsf{AN}(g_d) \leq d + t$, allowing to conclude.

□

The direct sum of 2 threshold functions with the same threshold d (lower than half) has its algebraic immunity between d and $2d$. We show sufficient conditions to reach this maximum:

Proposition 4. *Let $d \in \mathbb{N}^*$, and $t_1, t_2, n_1, n_2 \in \mathbb{N}$ such that $n_1 = 2d - 1 + t_1$, and $n_2 = 2d - 1 + t_2$. If $\max(t_1, t_2) \geq d$ then $\mathsf{AI}(\mathsf{DS}(\mathsf{T}_{d,n_1}, \mathsf{T}_{d,n_2})) = 2d$.*

Proof. Using Lemma 4: for $i \in \{1,2\}$: $\mathsf{AN}(\mathsf{T}_{d,n_i}) = d + t_i$, $\mathsf{AN}(1 + \mathsf{T}_{d,n_1}) = d$ and $\Delta_{\mathsf{AN}}(\mathsf{T}_{d,n_i}) = t_i$. Corollary 3 gives $\mathsf{AI}(\mathsf{DS}(\mathsf{T}_{d,n_1}, \mathsf{T}_{d,n_2})) \geq \max(d + \min(d, t_1), d + \min(d, t_2)) \geq 2d$. Since $\mathsf{AI}(\mathsf{T}_{d,n_i}) = d$ for $i \in \{1,2\}$ Lemma 1 gives the upper bound of $2d$. □

Using a property on the ANF of threshold functions studied in [Méa19], we show sufficient conditions preventing to reach the maximum:

Lemma 9. *(Adapted from [Méa19], Lemma 5 and Proposition 3) Let $d, n, D \in \mathbb{N}^*$ such that $d \in [n]$ and $D = 2^{\lceil \log d \rceil}$, let a_I for $I \subseteq [n]$ be the ANF coefficients of $\mathsf{T}_{d,n}$. The following holds: if $a_I = 1$ then $|I| \in \cup_{k \in \mathbb{N}} [kD + d, kD + D]$.*

Proposition 5. *Let $d, D \in \mathbb{N}^*$ such that $D = 2^{\lceil \log d \rceil}$ and $t_1, t_2, n_1, n_2 \in \mathbb{N}$ such that $n_1 = 2d - 1 + t_1$, and $n_2 = 2d - 1 + t_2$. If $\max(t_1, t_2) \leq D - d$ then $\mathsf{AI}(\mathsf{DS}(\mathsf{T}_{d,n_1}, \mathsf{T}_{d,n_2})) < 2d$.*

Proof. For readability denote $\psi = \mathsf{DS}(\mathsf{T}_{d,n_1}, \mathsf{T}_{d,n_2})$. The condition $\max(t_1, t_2) \leq D - d$ implies $\max(n_1, n_2) \leq D + d - 1$. Hence, Lemma 9 gives an upper bound on the degree: $\mathsf{deg}(\psi) \leq D$. Since $D < 2d$, $1 + \psi$ is a non-null annihilator of ψ of degree lower than $2d$.

□

We conclude this part with the particular case of the direct sum of twice the same threshold function.

Proposition 6. *Let $d, n \in \mathbb{N}^*$ such that $d \in [n]$. If $1 \leq d \leq (n-1)/3$ or $2(n+1)/3 \leq d < n$ then $\mathsf{AI}(\mathsf{DS}^2(\mathsf{T}_{d,n})) = 2\mathsf{AI}(\mathsf{T}_{d,n})$.*

Proof. The case $d \leq (n-1)/3$ corresponds to the sub-case of Proposition 4 where $n_1 = n_2$. For the case $d \geq 2(n+1)/3$, Lemma 4 gives $\mathsf{AN}(\mathsf{T}_{d,n}) = n - d + 1 \leq (n-1)/3$, $\mathsf{AN}(1 + \mathsf{T}_{d,n}) = d \geq 2(n+1)/3$, $\mathsf{AI}(\mathsf{T}_{d,n}) = n - d + 1$ and $\varDelta_{\mathsf{AN}}(\mathsf{T}_{d,n}) = 2d - n - 1 \geq (n-1)/3$. Hence, applying Corollary 3 we obtain $\mathsf{AI}(\mathsf{DS}^2(\mathsf{T}_{d,n})) \geq 2(n-d+1)$, and since in this case $2(n-d+1) = \mathsf{AI}(\mathsf{T}_{d,n}) + \mathsf{AI}(\mathsf{T}_{d,n})$ Lemma 1 allows to conclude.

□

5 Mixed approach: combining $\varDelta_{\mathsf{AN}}(f)$ and AI increasing property.

In this part we combine the two approaches of Section 3 and Section 4 to improve the lower bound of Lemma 1. We extend the approach of Section 3 on the algebraic degree, showing that under some conditions the higher degree part of the ANF coefficients of a function are sufficient to guarantee a lower bound on its AI. We refer to functions satisfying this property as *AI increasing functions* and show how they can guarantee an higher AI than the lower bound of Lemma 1 when they are used as component of a direct sum. In the main theorem we combine the $\varDelta_{\mathsf{AN}}(f)$ and the AI increasing property of the function g to obtain our best lower bound on $\mathsf{AI}(\mathsf{DS}(f, g))$, encompassing the improvements from both approaches. Then in Section 5.1 we give examples of AI increasing functions and cases where the theorem allows to determine exactly the algebraic immunity.

Definition 9. *Let $n \in \mathbb{N}^*$, $d, t \in \mathbb{N}$ such that $t \leq n$ and $d \leq n$, and $f \in \mathcal{B}_n$ with ANF coefficients $(f_I)_{I \subseteq [n]}$. We denote $T = \sum_{i=t}^{n} \binom{n}{i}$ and $D = \sum_{i=0}^{d} \binom{n}{i}$. We call $\mathbf{A}$-matrix of f with parameter t, d the binary matrix $\mathbf{A}_{t,d}(f)$ where:*

- *The T rows are indexed by the sets $I \subseteq [n]$ such that $|I| \geq t$,*
- *the D columns are indexed by the sets $J \subseteq [n]$ such that $|J| \leq d$,*
- *the entry at row I and column J is 0 if $J \not\subseteq I$ and $\sum_{K \subseteq J} f_{K \cup \{I \setminus J\}}$ otherwise.*

Note that the $\mathbf{A}$-matrices represent sub-systems of the equations appearing when we consider the ANF coefficients of an annihilator of f as in Lemma 7 or the PANF coefficients of an annihilator of a direct sum as in Lemma 2. For example, $\mathsf{AI}(f) > d$ is equivalent to $\mathsf{rank}(\mathbf{A}_{0,d}(f)) = D$ and $\mathsf{rank}(\mathbf{A}_{0,d}(f+1)) = D$. We give a stronger property on these matrices allowing to bound the algebraic immunity of f and then we use this matrix formalism to improve upon Lemma 8.

Property 2. *Let $n \in \mathbb{N}^*$, $d, t \in \mathbb{N}$ such that $d \leq n$ and $t \leq n$, we denote $D = \sum_{i=0}^{d} \binom{n}{i}$. If $t > d$ and $\mathsf{rank}(\mathbf{A}_{t,d}(f)) = D$ then $\mathsf{AI}(f) > d$.*

Proof. First, note that the ANF coefficient $f_\emptyset$ can be involved only when $\{I \setminus J\} = \emptyset$, which does not happen for $t > d$. Hence we obtain $\mathbf{A}_{t,d}(f) = \mathbf{A}_{t,d}(f+1)$. Then, by definition $\mathbf{A}_{t,d}(f)$ is a sub-matrix of $\mathbf{A}_{0,d}(f)$, therefore $\mathsf{rank}(\mathbf{A}_{t,d}(f)) = D$ implies $\mathsf{rank}(\mathbf{A}_{0,d}(f)) = D$. Since the coefficients of $\mathbf{A}_{0,d}(f)$ correspond to the system of Lemma 7, $\mathsf{rank}(\mathbf{A}_{0,d}(f)) = D$ means that for all not trivial combination of the columns the result for at least one of the row indexed by $I' \in [n]$ is equal to one, which is equivalent to $\mathsf{AN}(f) > d$ due to Lemma 7. The same reasoning applying on $f+1$, we can conclude $\mathsf{AI}(f) > d$.

□

Theorem 1. *Let $n, m \in \mathbb{N}^*$, $d, D, k \in \mathbb{N}$ such that $k \geq d$ and $D = \sum_{i=0}^{d} \binom{m}{i}$. Let $f \in \mathcal{B}_n$ and $g \in \mathcal{B}_m$, if $\mathsf{rank}(\mathbf{A}_{\mathsf{AI}(f)+k+1,d}(g)) = D$, $\mathsf{AI}(g) > k$, and $\Delta_{\mathsf{AN}}(f) > k-d-1$ then $\mathsf{AI}(\mathsf{DS}(f,g)) > \mathsf{AI}(f) + k$.*

Proof. The proof structure is similar to the one of Lemma 8. Without loss of generality we take f such that $\mathsf{AN}(f+1) \geq \mathsf{AN}(f)$. We do the proof by contradiction, assuming there exists $h \neq 0$ in $\mathcal{B}_{n+m}$ of degree at most $\mathsf{AI}(f)+k$ such that $h \cdot \mathsf{DS}(f,g) = 0$. Since $\mathsf{AI}(g+1) = \mathsf{AI}(g)$ and $\mathbf{A}_{\mathsf{AI}(f)+k+1,d}(g+1) = \mathbf{A}_{\mathsf{AI}(f)+k+1,d}(g)$ since $k \geq d$ (see the beginning of the proof of Property 2), the reasoning on $\mathsf{DS}(f,g)$ also applies on $\mathsf{DS}(f,g+1)$. Hence the contradiction on h is sufficient to prove the bound on the algebraic immunity of $\mathsf{DS}(f,g)$.

h being an annihilator of $\mathsf{DS}(f,g)$, Lemma 2 gives:

$$\forall I \subseteq [m], \quad h_I(x) f = \sum_{J \subseteq I} h_J(x) \sum_{K \subseteq J} g_{K \cup \{I \setminus J\}}. \tag{7}$$

The right hand side is always an annihilator of $f+1$, hence the null function or a function of algebraic degree at least $\mathsf{AN}(f+1)$. For all the equations such that $|I| > \mathsf{AI}(f)+k$ the PANF coefficient $h_I(x)$ is the null function (otherwise h as degree greater than $\mathsf{AI}(f)+k$), forcing the right hand side to be the null function. Since $\mathsf{rank}(\mathbf{A}_{\mathsf{AI}(f)+k+1,d}(g)) = D$ it means that each one of the D PANF coefficients $h_J(x)$ for $|J| \leq d$ can be isolated from the others by summing equations over subsets $I \subseteq [m]$. Without loss of generality, isolating the PANF coefficient relative to J', we obtain the following equation:

$$0 = h_{J'}(x) + \sum_{\substack{K \subseteq [m] \\ |K| > d}} h_K(x) b_K, \tag{8}$$

where b_K are binary coefficients.

The constraint on the degree: $\forall I \subseteq [m], \mathsf{deg}(h_I(x)) \leq \mathsf{AN}(f)+k-|I|$, forces the sum in Equation 8 to have degree at most $\mathsf{AN}(f)+k-d-1$. Hence, $\mathsf{deg}(h_{J'}) \leq \mathsf{AN}(f)+k-d-1$ to satisfy Equation 8. Thus, all PANF coefficients h_I such that $|I| \leq d$ have degree at most $\mathsf{AN}(f)+k-d-1$. For all subsets $I \subseteq [m]$ such that $|I| > d$ the degree constraint leads to $\mathsf{deg}(h_I) \leq \mathsf{AN}(f)+k-d-1$. Hence, for all subsets of $[m]$ the PANF coefficient has degree at most $\mathsf{AN}(f)+k-d-1$, and therefore the right hand side of Equation 7 is always the null function since $\mathsf{AN}(f+1) > \mathsf{AN}(f)+k-d-1$. From there we use the same reasoning as for Lemma 8.

For $|I| > k$ all $h_I(x)$ are null due to the degree constraint, therefore all the right hand sides for the 2^m equations are combinations of the PANF coefficients related to subset of cardinal at most k. Hence we can rewrite Equation 7 as:

$$\forall I \subseteq [m], \quad h_I(x) f = \sum_{\substack{J \subseteq I \\ 0 \leq |J| \leq k}} h_J(x) \sum_{K \subseteq J} g_{K \cup \{I \setminus J\}} = 0. \tag{9}$$

Since h is non-null there exists at least one $z = (x', y') \in \mathbb{F}_2^{n+m}$ such that $h(z) = 1$ and therefore at least one $J \subseteq [m]$ such that $h_J(x') = 1$. Hence for this particular $x' \in \mathbb{F}_2^n$ Equation 9 leads to:

$$\forall I \subseteq [m], \quad \sum_{\substack{J \subseteq I \\ 0 \leq |J| \leq k}} h_J(x') \sum_{K \subseteq J} g_{K \cup \{I \setminus J\}} = 0,$$

where the binary values $h_J(x')$ are not all null. Since $\mathsf{AI}(g) > k$ applying Lemma 7 we know that there is at least one subset $I' \subseteq [m]$ such that the equation is not satisfied, which gives the contradiction. We can conclude $\mathsf{AI}(\mathsf{DS}(f,g)) > \mathsf{AI}(f) + k$.

□

Remark 3. The case $k = 0$ corresponds to: if $\mathsf{rank}\mathbf{A}_{\mathsf{AI}(f)+1,0}(g) = 1$, $\mathsf{AI}(g) > 0$, and $\Delta_{\mathsf{AN}}(f) > -1$ then $\mathsf{AI}(\mathsf{DS}(f,g)) > \mathsf{AI}(f)$. Note that $\mathsf{rank}\mathbf{A}_{\mathsf{AI}(f)+1,0}(g) = 1$ is equivalent to $\mathsf{deg}(g) > \mathsf{AI}(f)$, $\mathsf{AI}(g) > 0$ means that g is non constant and $\Delta_{\mathsf{AN}}(f) > -1$ is true for all function. Hence the case $k = 0$ gives the result of Lemma 5.

5.1 AI increasing functions

Functions satisfying Property 2 can be used to produce direct sums with algebraic immunity higher than the two components. In this part we show such functions exist for appropriate choices of n, d and t, and give an example of application of Theorem 1.

Definition 10 (AI increasing functions). *Let $n \in \mathbb{N}^*$, $d, e \in \mathbb{N}$ such that $d \leq (n-1)/2$, and $e \geq d$ we denote $D = \sum_{i=0}^{d} \binom{n}{i}$. We denote $\mathcal{C}(n,d)$ the functions $f \in \mathcal{B}_n$ such that $\mathsf{rank}(\mathbf{A}_{n-d,d}(f)) = D$ and $\mathcal{C}(n,d,e)$ such function with algebraic immunity greater than e.*

Since these functions are defined to satisfy Property 2 (with $t = n - d$) their AI is greater than d. Moreover they allow to find more functions with the same bound on the AI:

Proposition 7. *Let $n \in \mathbb{N}^*$, $d \in \mathbb{N}$ such that $d \leq (n-1)/2$. If $f \in \mathcal{C}(n,d)$ then $\forall g$ $\mathcal{B}_n$ such that $\mathsf{deg}(g) < n - 2d$ then the function[1] $f + g$ belongs to $\mathcal{C}(n,d)$.*

Proof. Since $f \in \mathcal{C}(d,n)$ by definition $\mathsf{rank}(\mathbf{A}_{n-d,d}(f)) = D$. Then, by definition of the matrix $\mathbf{A}_{n-d,d}(f)$ (Definition 9) for each row I the elements are obtained with equations depending on the ANF coefficients f_U where $|I| \geq |U| \geq |I| - |J|$. Since in $\mathbf{A}_{n-d,d}(f)$ we have $|I| \geq n - d$ and $|J| \leq d$ the matrix is independent of the value of the ANF coefficients f_U such that $0 \leq |U| < n - 2d$. Thereafter, for all $g \in \mathcal{B}_n$ such that $\mathsf{deg}(g) < n - 2d$, $\mathbf{A}_{n-d,d}(f+g) = \mathbf{A}_{n-d,d}(f)$ hence $f + g \in \mathcal{C}(d,n)$.

□

Then, we show that $\mathcal{C}(n,d)$ is not empty:

Proposition 8. *Let $n \in \mathbb{N}^*$, and $d \in \mathbb{N}$ such that $d \leq (n-1)/2$, then $\mathsf{T}_{n-d,n} \in \mathcal{C}(d,n)$.*

Proof. From Lemma 4 $\mathsf{AI}(\mathsf{T}_{n-d,n}) = d + 1$, hence $\mathsf{AN}(\mathsf{T}_{n-d,n}) > d$ and $\mathsf{AN}(\mathsf{T}_{n-d,n} + 1) > d$ which is equivalent to $\mathbf{A}_{0,d}(\mathsf{T}_{n-d,n})$ and $\mathbf{A}_{0,d}(\mathsf{T}_{n-d,n} + 1)$ being of rank D by Lemma 7. From Lemma 9 only the ANF coefficients relative to the subsets I such that $|I| \geq n - d$ can be equal to 1, hence only the $D \times D$ sub-matrix $\mathbf{A}_{n-d,d}(\mathsf{T}_{n-d,n})$ of $\mathbf{A}_{0,d}(\mathsf{T}_{n-d,n})$ is not null, and then of rank D, allowing to conclude $\mathsf{T}_{n-d,n} \in \mathcal{C}(d,n)$.

□

Using the notation of $\mathcal{C}(d,n)$ we derive a corollary of Theorem 1 and an example of construction.

[1] here $f + g$ is the standard sum of f and g and not the direct sum

Corollary 4. *Let $n, m \in \mathbb{N}^*$, $d, t \in \mathbb{N}$. Let $f \in \mathcal{B}_n$ such that $\varDelta_{\mathsf{AN}}(f) \geq t$ and $m \geq \mathsf{AI}(f) + t + 1 + 2d$. The following holds:*

$$\forall g \in \mathcal{C}(d, m, d+t), \quad \mathsf{AI}(\mathsf{DS}(f, g)) > \mathsf{AI}(f) + d + t,$$

and in particular if $\mathsf{AI}(g) = d + t + 1$ then $\mathsf{AI}(\mathsf{DS}(f, g)) = \mathsf{AI}(f) + d + t + 1$.

Proof. t is replacing $k - d$ in the theorem, since $g \in \mathcal{C}(d, m, d+t)$ the matrix $\mathbf{A}_{m,g}(m-d)$ is full rank and $\mathsf{AI}(f) + t + d + 1 \leq m - d$ by definition of m, then f and g satisfy the conditions of the theorem, allowing to conclude $\mathsf{AI}(\mathsf{DS}(f, g)) > \mathsf{AI}(f) + d + t$. The particular case comes from the upper bound of Lemma 1. □

Proposition 9. *Let $\ell \in \mathbb{N}^*$, $d, m \in \mathbb{N}$ such that $m \geq \ell + 1 + 2d$, the direct sum of the ℓ-th triangular function and the m-variable threshold function of threshold $m - d$ has algebraic immunity:*

$$\mathsf{AI}(\mathsf{DS}(T_\ell, \mathsf{T}_{m-d,m})) = \ell + d + 1.$$

Proof. We apply Corollary 4 with $f = T_\ell$ and $g = \mathsf{T}_{m-d,d}$. The algebraic immunity of T_ℓ is ℓ by Lemma 3 and since T_ℓ and $1 + T_\ell$ are degree $\mathsf{AI}(T_\ell)$ functions annihilating each other $\varDelta_{\mathsf{AN}}(T_\ell) = 0 = t$. Then, $m \geq \mathsf{AI}(T_\ell) + \ell + t + 1$, $\mathsf{T}_{m-d,m} \in \mathcal{C}(d, m)$ by Proposition 8 and $\mathsf{AI}(\mathsf{T}_{m-d,m}) = d + 1$ by Lemma 4. It allows to apply Corollary 4, giving $\mathsf{AI}(\mathsf{DS}(T_\ell, \mathsf{T}_{m-d,m})) = \ell + d + 1$.

□

6 Conclusion

In this article we studied criteria on Boolean functions f and g allowing to improve the usual lower bound on the algebraic immunity of their direct sum: $\mathsf{AI}(\mathsf{DS}(f, g)) \geq \max(\mathsf{AI}(f), \mathsf{AI}(g))$. We showed the degree of one of the functions is sometimes enough to get a better bound: $\mathsf{AI}(\mathsf{DS}(f, g)) > \max(\mathsf{AI}(f), \mathsf{AI}(g))$, which happens to be sufficient to determine the algebraic immunity of iterated direct sums such as DSM or WPRF candidate functions. Extending this property from the degree, we studied sufficient conditions on the higher part of the ANF coefficients to guarantee an AI of at least d, ant it lead to the concept of *AI increasing functions* which allow to improve the lower bound of $\mathsf{AI}(\mathsf{DS}(f, g))$ by d. We proved than the value of the difference of minimal degree of annihilators, $\varDelta_{\mathsf{AN}}(f)$, can also be a sufficient criteria to improve the lower bound. More precisely, a $\varDelta_{\mathsf{AN}}(f)$ of value k can lead to an improvement of k over the maximum of the two AI, and we gave examples of constructions with this property using threshold functions as components. To conclude, we showed both approaches can be combined, giving our best improvement on the lower bound. It allows to use the secondary construction to produce functions with AI strictly higher than its components, and it reduces the gap between the minimum value of AI that can be guaranteed and its maximal value.

In view of these results, a natural direction for further research would be to study if the same properties allow to improve on the upper bound, $\mathsf{AI}(f) + \mathsf{AI}(g)$. Such study could allow to determine the exact AI of some constructions, and eventually lead to a full characterization of the AI of direct sums, by the mean of the AI and few more parameters of the component functions. Such a full characterization for the simplest secondary construction would lead to multiple constructions with known AI in a high number of variables, and the exact AI of functions in $n > 20$ variables for which the exact AI computation is out of reach with the current computational power.

Another natural question from this work is the potential use of AI increasing functions to improve algorithms determining the algebraic immunity of a function, or discarding functions with low algebraic immunity. Various algorithms to compute the algebraic immunity of any Boolean function are given in [ACG$^+$06] and the complexity of such algorithms is discussed in further works such as [Dal13, JZW14].

Instead, the AI increasing property of order d is not a property shared by all functions of AI greater than d, but checking this property requires to check only the coefficients of degree $n-2d$ to n, in order to guarantee an AI of at least $d+1$. Hence, it could lead to algorithms with a lower computational cost used to guarantee a minimal quantity of AI for tested functions.

7 Acknowledgments

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

References

ACG$^+$06. Frederik Armknecht, Claude Carlet, Philippe Gaborit, Simon Künzli, Willi Meier, and Olivier Ruatta. Efficient computation of algebraic immunity for algebraic and fast algebraic attacks. In Serge Vaudenay, editor, *EUROCRYPT 2006*, volume 4004 of *LNCS*. Springer, Heidelberg, May / June 2006.

AL16. Benny Applebaum and Shachar Lovett. Algebraic attacks against random local functions and their countermeasures. In Daniel Wichs and Yishay Mansour, editors, *48th ACM STOC*. ACM Press, June 2016.

BCG$^+$20. Elette Boyle, Geoffroy Couteau, Niv Gilboa, Yuval Ishai, Lisa Kohl, and Peter Scholl. Correlated pseudorandom functions from variable-density lpn. In *61st FOCS*. IEEE Computer Society Press, 2020.

BIP$^+$18. Dan Boneh, Yuval Ishai, Alain Passelègue, Amit Sahai, and David J. Wu. Exploring crypto dark matter: - new simple PRF candidates and their applications. In *Theory of Cryptography - 16th International Conference, TCC 2018, Panaji, India, November 11-14, 2018, Proceedings, Part II*, pages 699–729, 2018.

BP05. An Braeken and Bart Preneel. 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*, pages 35–48, 2005.

Car06. Claude Carlet. On the higher order nonlinearities of algebraic immune functions. In Cynthia Dwork, editor, *CRYPTO 2006*, volume 4117 of *LNCS*, pages 584–601. Springer, Heidelberg, August 2006.

Car21. Claude Carlet. *Boolean Functions for Cryptography and Coding Theory*. Cambridge University Press, 2021.

CM03. Nicolas Courtois and Willi Meier. Algebraic attacks on stream ciphers with linear feedback. In Eli Biham, editor, *EUROCRYPT 2003*, volume 2656 of *LNCS*. Springer, Heidelberg, May 2003.

CM19. Claude Carlet and Pierrick Méaux. Boolean functions for homomorphic-friendly stream ciphers. *Algebra, Codes and Cryptology*, pages 166–182, 11 2019.

CM20. Claude Carlet and Pierrick Méaux. A complete study of two classes of boolean functions for homomorphic-friendly stream ciphers. *IACR Cryptol. ePrint Arch.*, 2020:1562, 2020.

Dal13. Deepak Kumar Dalai. Computing the rank of incidence matrix and the algebraic immunity of Boolean functions. Cryptology ePrint Archive, Report 2013/273, 2013.

Dil76. J. Dillon. *Elementary Hadamard Difference Sets*. PhD thesis, University of Maryland, USA, 1976.

DMS06. Deepak Kumar Dalai, Subhamoy Maitra, and Sumanta Sarkar. Basic theory in construction of boolean functions with maximum possible annihilator immunity. *Designs, Codes and Cryptography*, 2006.

GJLS20. Romain Gay, Aayush Jain, Huijia Lin, and Amit Sahai. Indistinguishability obfuscation from simple-to-state hard problems: New assumptions, new techniques, and simplification. *IACR Cryptol. ePrint Arch.*, 2020:764, 2020.

Gol01. Oded Goldreich. *Foundations of Cryptography: Basic Tools*, volume 1. Cambridge University Press, Cambridge, UK, 2001.

HKM17. Matthias Hamann, Matthias Krause, and Willi Meier. LIZARD - A lightweight stream cipher for power-constrained devices. *IACR Trans. Symmetric Cryptol.*, 2017(1):45–79, 2017.

JZW14. Lin Jiao, Bin Zhang, and M. Wang. Revised algorithms for computing algebraic immunity against algebraic and fast algebraic attacks. In *ISC*, 2014.

MCJS19. Pierrick Méaux, Claude Carlet, Anthony Journault, and François-Xavier Standaert. Improved filter permutators for efficient FHE: better instances and implementations. In Feng Hao, Sushmita Ruj, and Sourav Sen Gupta, editors, *Progress in Cryptology - INDOCRYPT*, volume 11898 of *LNCS*, pages 68–91. Springer, 2019.

Méa19. Pierrick Méaux. On the fast algebraic immunity of majority functions. In Peter Schwabe and Nicolas Thériault, editors, *Progress in Cryptology - LATINCRYPT*, volume 11774 of *LNCS*, pages 86–105. Springer, 2019.

Méa20. Pierrick Méaux. On the fast algebraic immunity of threshold functions. Cryptology ePrint Archive, Report 2020/273, 2020.

Mes08. Sihem Mesnager. Improving the lower bound on the higher order nonlinearity of boolean functions with prescribed algebraic immunity. *IEEE Trans. Inf. Theory*, 54(8):3656–3662, 2008.

MJSC16. Pierrick Méaux, Anthony Journault, François-Xavier Standaert, and Claude Carlet. Towards stream ciphers for efficient FHE with low-noise ciphertexts. In Marc Fischlin and Jean-Sébastien Coron, editors, *EUROCRYPT 2016, Part I*, volume 9665 of *LNCS*, pages 311–343. Springer, Heidelberg, May 2016.

MPC04. Willi Meier, Enes Pasalic, and Claude Carlet. Algebraic attacks and decomposition of boolean functions. In *EUROCRYPT*, volume 3027, pages 474–491. Springer, 2004.

NLV11. Michael Naehrig, Kristin E. Lauter, and Vinod Vaikuntanathan. Can homomorphic encryption be practical? In Christian Cachin and Thomas Ristenpart, editors, *ACM Cloud Computing Security Workshop, CCSW*, pages 113–124. ACM, 2011.

Riz10. Panagiotis Rizomiliotis. Improving the high order nonlinearity lower bound for boolean functions with given algebraic immunity. *Discrete Applied Mathematics*, 158:2049–2055, 11 2010.

Rot76. O.S Rothaus. On bent functions. *Journal of Combinatorial Theory, Series A*, 20(3):300–305, 1976.