Graph indicators of vectorial functions and bounds on the algebraic degree of composite functions

Claude Carlet*,
University of Bergen, Norway.

E-mail: claude.carlet@gmail.com

Abstract

Given a vectorial function $F : \mathbb{F}_2^n \mapsto \mathbb{F}_2^m$, the indicator $1_{\mathcal{G}_F}$ of its graph $\mathcal{G}_F = \{(x, F(x)); x \in \mathbb{F}_2^n\}$ allows to express the algebraic degree of F in a simple way. Exploiting the formula, obtained in a previous paper, for the graph indicator of a composite function $G \circ F$, that involves only a sum of products of $1_{\mathcal{G}_F}$ and $1_{\mathcal{G}_G}$, we deduce bounds on the algebraic degree of $G \circ F$, whose efficiency comes from the fact that the algebraic degree of the product of two Boolean functions is bounded above by the sum of their algebraic degrees, while for a composition, it is bounded above by their product. One of these bounds, that depends on the algebraic degrees of G and $1_{\mathcal{G}_F}$, is tight, general, simple, and most often efficient (for the case where it is not efficient, we give an improved bound, that is a little more complex). As far as we know, it is the first efficient upper bound ever found, that works without any condition on the vectorial functions. It provides a new criterion for the choice of S-boxes in block ciphers. It implies as a corollary a known bound assuming the divisibility of the Walsh transform values by a power of 2. It gives a better view why this latter bound works. Our results nicely generalize to more than two functions.
When F is a permutation, our expression of the algebraic degree of $G \circ F$ simplifies into a formula involving the algebraic degrees of the products of a coordinate function of G and coordinate functions of F^{-1}. This implies and improves another known bound showing that the algebraic degree of F^{-1} has more impact on that of $G \circ F$ than that of F itself, and providing a criterion for the choice of S-boxes in block ciphers when they are permutations: both algebraic degrees of F and F^{-1} should be as large a possible. Our approach by graph indicators gives an explanation to this interesting fact. Our results include all the known efficient bounds as particular cases, and clarify the reasons why they work. We also deduce the exact expression of the algebraic degree of the composition of three functions, leading to a bound that is much more efficient than what we

*The research of the author is partly supported by the Trond Mohn Foundation.

obtain by applying the known bound two times. We also obtain two bounds on the algebraic degree of $G \circ F$, given the divisibility by powers of 2 of coefficients in the numerical normal forms of component functions of F^{-1}, and their sums with a coordinate function of G. We study their consequences and generalizations.

Index Terms: vectorial Boolean function, composition, algebraic degree.

1 Introduction

Vectorial functions (that is, for some positive integers n and m, mappings from $\mathbb{F}_2^n$ to $\mathbb{F}_2^m$, that we shall also call (n,m)-functions) play a central role in stream ciphers (as filter functions) and in block ciphers (as substitution boxes; in brief, S-boxes). Their role is to provide *confusion*, see [14]. They also play a role in coding theory, see [13].

When $m = 1$, we speak of Boolean functions. In the model of block cipher called substitution-permutation network (SPN), we have $m = n$ and these vectorial functions should be permutations (i.e. bijective).

All known block ciphers are the iterations, called *rounds*, of a transformation depending on a round key, acting on blocks of plaintext, and including at least one well chosen nonlinear vectorial function (S-box) in its design. Such iterations make that the output of the i-th round is the composition of vectorial functions, among which at least i are non-affine.

Every Boolean (resp. vectorial) function has a unique representation as a polynomial in $\mathbb{F}_2[x_1,\ldots,x_n]/(x_1^2-x_1,\ldots,x_n^2-x_n)$ (resp. $\mathbb{F}_2^m[x_1,\ldots,x_n]/(x_1^2-x_1,\ldots,x_n^2-x_n)$), called its algebraic normal form (ANF), that we shall define more in detail in the next section. The degree of this unique polynomial in n variables is called the algebraic degree of the function; we shall denote the algebraic degree of a function F by $d_{alg}(F)$. It is important for the designer of a block cipher that the algebraic degree of the output of a series of rounds has an algebraic degree as large as possible, since otherwise, distinguishing attacks may be possible.

Several representations of a vectorial function are possible: the ANF of the function is not the only option, the ANF of the indicator of the graph $\mathcal{G}_F = \{(x,F(x));\, x \in \mathbb{F}_2^n\}$ of F is a second possible one. This latter representation, already studied in [6], will play an important role in the present paper. Note that the graph is at the origin of an important notion of equivalence known for vectorial functions, called the CCZ equivalence [8, 2]: two (n,m)-functions are called CCZ equivalent if their graphs correspond to each other by an affine permutation. This equivalence is the most general known which preserves the two main parameters of vectorial functions quantifying their resistance to the main attacks (the differential attack and the linear attack): their differential uniformity and their nonlinearity (we refer to [5] and to the more recent [7] for more details). The CCZ equivalence does not preserve the algebraic degree (we will say more about that in Subsection 5.2.3).

Every Boolean function has also a unique representation as a polynomial in

$\mathbb{Z}[x_1,\ldots,x_n]/(x_1^2-x_1,\ldots,x_n^2-x_n)$ called its numerical normal form (NNF) [9]. The degree of this unique polynomial in n variables is called the *numerical degree* of the function and is related to the Walsh transform values $W_f(a)=\sum_{x\in\mathbb{F}_2^n}(-1)^{f(x)+a\cdot x}$ (associated to some inner product "$\cdot$" in $\mathbb{F}_2^n$), see more in Section 2.

It has been shown in [1] that, when F is a permutation, the algebraic degree of the compositional inverse F^{-1} of F plays an important role with respect to the algebraic degree of the composition of F by other functions: for every function G, the algebraic degree $d_{alg}(G\circ F)$ of the composite function $G\circ F$ is bounded above by $n-\left\lfloor\frac{n-1-d_{alg}(G)}{d_{alg}(F^{-1})}\right\rfloor$. This shows that the choice of bijective S-boxes in block ciphers should try to maximize both their algebraic degree and the algebraic degree of their compositional inverse. The proof in [1] of this important result has some technicality and does not provide a quite simple view of the reasons why this happens. Moreover, this bound applied for the composition of more than two functions seems rather weak and there is then room for improvement. It is also shown in [3] that if the Walsh transform values of F are divisible by 2^k, then we have $d_{alg}(G\circ F)\leq n-k+d_{alg}(G)$ for every G. This latter result is rather specific, and its proof gives the impression that some property is hidden behind the stage. It would then be good to find a general bound, which would give a global view on these two particular bounds, and would clarify them, and possibly improve upon them. It would also be important to be able to handle more simply and more efficiently the compositions of more than two functions.

In the present paper, we exploit some of the results of [6], to bound the algebraic degree of composite vectorial functions. We consider the ANF of the indicator (i.e. the characteristic function) $1_{\mathcal{G}_F}$ of the graph $\mathcal{G}_F=\{(x,F(x));\ x\in\mathbb{F}_2^n\}$ of any vectorial function F, and we relate the algebraic degree of F to the degree of a part of the ANF of this graph indicator. This provides an upper bound on $d_{alg}(F)$ by means of $d_{alg}(1_{\mathcal{G}_F})$, which is often weak but not always since it is tight. We observed in [6] that the graph indicator of a composite function $G\circ F$ can be expressed by means of $1_{\mathcal{G}_F}$ and $1_{\mathcal{G}_G}$, with a formula involving additions and multiplications only (no composition). This provides relations between the coordinate functions of $G\circ F$ and those of F and G, and it leads to a bound on $d_{alg}(G\circ F)$ which is most often better than the naive (or trivial) bound $d_{alg}(G\circ F)\leq(d_{alg}(F))\ (d_{alg}(G))$ (which comes from the observation that, when calculating the ANF of $G\circ F$ by substituting the coordinate functions of F for the input coordinates to G, at most $d_{alg}(G)$ coordinate functions of F are multiplied). This second bound is tight as well, but still often weak. We derive then an expression of $d_{alg}(G\circ F)$ leading to a third and much better upper bound (see Theorem 1) by the algebraic degree of $1_{\mathcal{G}_F}$ added with the algebraic degree of G, minus the number of variables of G (i.e. of output bits of F). This latter bound is general, tight and simple, and it is efficient (we show that it can be much stronger than the bound of [1], in some cases). In fact, it seems to be the first time that an efficient upper bound on the algebraic degree of composite functions is found with no condition on the functions (the naive bound has also this latter property, but it is most often inefficient). It has also the interest of

involving parameters of F and G separately and provides a new criterion for the choice of S-boxes: the algebraic degree of the graph indicator minus the number of output bits should be large enough. Our bound directly implies the bound of [3] on those vectorial functions whose Walsh transform values are divisible by some power of 2 (which is then a particular case of a much more general bound), thanks to the well-know result that if all the Fourier-Hadamard transform values of a Boolean function are divisible by 2^k then its algebraic degree is bounded above by the number of its variables minus k. Moreover, we generalize this bound to the composition of more than two functions.
It is shown in [6] that, when F is a permutation, the expression of $1_{\mathcal{G}_{G\circ F}}$ simplifies. We show that this leads to an exact expression of $d_{alg}(G\circ F)$ by means of the algebraic degrees of the products of one coordinate function of G and of coordinate functions of F^{-1}. This exact value leads to an upper bound on $d_{alg}(G\circ F)$ which is a slightly stronger bound than the one obtained in [1], that we recalled above, and it provides an alternative and more enlightening proof. All this clarifies the reasons why the algebraic degree of F^{-1} plays a stronger role than that of F itself, and it strengthens the observations of [1]. We generalize this bound to the composition of more than two functions.
We also obtain an upper bound on $d_{alg}(G\circ F)$ by means of the divisibility by a power of 2 of the coefficients of the highest degree term in the numerical normal forms of component functions of F^{-1}, and their sums with one coordinate function of G. The coefficients of numerical normal forms being related to values of Walsh transforms, this bound and the bound of [3] have some similarity; we show that they are neither comparable from the viewpoint of their hypotheses, nor from the viewpoint of the limits they impose to the algebraic degree; they are then complementary. We prove a second bound, dealing with the coefficients of the numerical normal forms of component functions of F^{-1}. This latter bound proves again the bound of [3]. Its hypothesis includes that F is a permutation but, except for this restriction, is lighter than the hypothesis of the bound of [3]. It is then complementary to the bound of [3] as well. We finally study the extensions of these bounds to the composition of three functions.

2 Preliminaries

In this paper, some representations of Boolean functions will involve sums in $\mathbb{F}_2$, and some others will involve sums in $\mathbb{Z}$ (i.e. not modulo 2). We shall then need to distinguish between the sums in $\mathbb{Z}$, that we shall denote by $+$, and the sums modulo 2, that we shall denote by $\oplus$. However, instead of denoting the addition in $\mathbb{F}_2^n$ by $\oplus$, we shall denote it by $+$, because $\mathbb{F}_2^n$ will sometimes be identified with the field $\mathbb{F}_{2^n}$, in which the addition is traditionally denoted by $+$. This will create no problem of confusion in the reading. We denote by $w_H(u)$ the Hamming weight of an element u of $\mathbb{F}_2^n$. The functions from $\mathbb{F}_2^n$ to $\mathbb{F}_2^m$ are called (n,m)-*function*s. Such function F being given, the n-*variable* Boolean functions $f_1,\ldots,f_m$ defined at every $x\in\mathbb{F}_2^n$ by $F(x)=(f_1(x),\ldots,f_m(x))$, are called the *coordinate functions* of F. When the numbers m and n are not specified, (n,m)-

functions are called *vectorial Boolean functions* or simply *vectorial functions*.
Vectorial functions (in practice, (n,m)-functions where n and m are most often even, for reasons of efficiency, and are between 4 and 8 for the same reason) play a central role in the security of block ciphers, where they provide the necessary *confusion* (see [14]) and are called in such framework *substitution boxes* (*S-boxes*). An S-box needs to allow resistance to the two main known attacks and their variants: the differential attack (for which the S-boxes need to have low differential uniformity, see below) and the linear attack (for which they need to have large nonlinearity, see below as well). Such S-boxes are in general concatenated in substitution layers. The linear combinations over $\mathbb{F}_2$, with non-all-zero coefficients, of the coordinate functions of a vectorial function are called its *component functions* and play a major role in the security of the block cipher in which it is involved; the nonlinearity of a function is the minimum nonlinearity of its components (see below).
The whole cipher, made of the iteration of rounds that are the combinations of S-boxes, diffusion layers (whose role is to spread the influence of every input bit) and round-key additions, must ensure sufficient complexity after several rounds. In particular, the *algebraic degree* (see the definition below) of the global vectorial function whose input is the plaintext (or the private key, or both) and whose output is given by the r-th round, must be large enough as soon as r is large enough.
For $m = 1$, we call support of a Boolean function[1] f the set $supp(f) = \{x \in \mathbb{F}_2^n;\ f(x) = 1\}$ (and Hamming weight the size of this support), while the support of a vector $x \in \mathbb{F}_2^n$ equals $\{i \in \{1,\ldots,n\}; x_i = 1\}$. We summarize below the definitions and properties that shall be needed in the proofs of this paper, and we refer to [4, 5, 7] for more details.

The *truth-table* of a Boolean function and the *look-up table* of a vectorial function, that is, the table of all pairs of an element of $\mathbb{F}_2^n$ (on which an ordering is chosen) and of the value of the function at this input, gives some information on their cryptographic properties (like the Hamming weight in the former case and the balancedness in the latter) but not enough. The already mentioned *algebraic normal form* (*ANF*), which is the unique n-variable multivariate polynomial representation in $\mathbb{F}_2[x_1,\ldots,x_n]/(x_1^2+x_1,\ldots,x_n^2+x_n)$ in the case of Boolean functions and in $\mathbb{F}_2^m[x_1,\ldots,x_n]/(x_1^2+x_1,\ldots,x_n^2+x_n)$ in the case of (n,m)-functions, allows defining an important parameter, the algebraic degree. The ANF writes:

$$f(x) = \bigoplus_{I\subseteq\{1,\ldots,n\}} a_I \left(\prod_{i\in I} x_i\right) = \bigoplus_{I\subseteq\{1,\ldots,n\}} a_I\, x^I; \quad a_I \in \mathbb{F}_2, \tag{1}$$

$$F(x) = \sum_{I\subseteq\{1,\ldots,n\}} a_I \left(\prod_{i\in I} x_i\right) = \sum_{I\subseteq\{1,\ldots,n\}} a_I\, x^I; \quad a_I \in \mathbb{F}_2^m, \tag{2}$$

[1] We shall use lowercase letters for denoting Boolean functions and capital letters to denote multi-output vectorial functions.

where $x^I = \prod_{i \in I} x_i$ is called a monomial. The values of the function are given by the binary Möbius transform of the coefficients of the ANF:

$$f(x) = \bigoplus_{I \subseteq supp(x)} a_I, \tag{3}$$

where $supp(x)$ denotes the support of x. Conversely, the coefficients of the ANF are given by the binary Möbius transform of the values of the function:

$$\forall I \subseteq \{1, \ldots, n\},\ a_I = \bigoplus_{x \in \mathbb{F}_2^n;\ supp(x) \subseteq I} f(x). \tag{4}$$

The same formulas are valid for vectorial functions.
The degree of the ANF shall be denoted by $d_{alg}(f)$ (resp. $d_{alg}(F)$) and is called the *algebraic degree* of the function: $d_{alg}(f) = \max\{|I|;\ a_I \neq 0\}$, $d_{alg}(F) = \max\{|I|;\ a_I \neq (0, \ldots, 0)\}$, where $|I|$ denotes the size of I (with the convention that the zero function has algebraic degree 0, so that the algebraic degree can be invariant under translation of the output). This makes sense thanks to the existence and uniqueness of the ANF.
Note that the *algebraic degree* of an (n, m)-function equals the maximal algebraic degree of the coordinate functions of F and also equals the maximal algebraic degree of the component functions. It is an affine invariant (that is, its value does not change when we compose F, on the right or on the left, by an affine automorphism). According to Relation (4), we have:

Proposition 1 *For every n-variable Boolean function f, we have $d_{alg}(f) = n$ if and only if $w_H(f)$ is odd. More generally, for every (n, n)-function F, we have $d_{alg}(F) = n$ if and only if $\sum_{x \in \mathbb{F}_2^n} F(x) \neq (0, \ldots, 0)$.*

There is a representation with uniqueness of Boolean functions similar to the ANF but over $\mathbb{Z}$ instead of being over $\mathbb{F}_2$, which is called the *numerical normal form* (NNF) and will be a useful tool for proving some bounds below. It represents Boolean functions by elements of the quotient ring $\mathbb{Z}[x_1, \ldots, x_n]/(x_1^2 - x_1, \ldots, x_n^2 - x_n)$:

$$f(x) = \sum_{I \subseteq \{1,\ldots,n\}} \lambda_I \left(\prod_{i \in I} x_i \right) = \sum_{I \subseteq \{1,\ldots,n\}} \lambda_I\, x^I; \quad \lambda_I \in \mathbb{Z}, \tag{5}$$

where the addition is in $\mathbb{Z}$. The NNF of f can be directly deduced from its ANF since we have:

$$\begin{aligned} f(x) = \bigoplus_{I \subseteq \{1,\ldots,n\}} a_I\, x^I \quad &\Longleftrightarrow \quad (-1)^{f(x)} = \prod_{I \subseteq \{1,\ldots,n\}} (-1)^{a_I\, x^I} \\ &\Longleftrightarrow \quad 1 - 2\ f(x) = \prod_{I \subseteq \{1,\ldots,n\}} (1 - 2\ a_I\, x^I) \end{aligned} \tag{6}$$

and expanding (6) gives the NNF of $f(x)$.
We call the degree of the NNF of a Boolean function f its *numerical degree*

and denote it by $d_{num}(f)$. Since the ANF of a Boolean function is the mod 2 version of its NNF, the numerical degree is always bounded below by the algebraic degree (and determining all the Boolean functions for which these two degrees are equal is an open problem).
Applying Relation (6) when $f(x)$ is a linear function and with Boolean functions $f_1, \ldots, f_k$ instead of variables, we deduce the formula:

$$\bigoplus_{i=1}^{k} f_i = \sum_{\emptyset \neq I \subseteq \{1,\ldots,k\}} (-2)^{|I|-1} \prod_{i \in I} f_i. \tag{7}$$

Inverting Relation (7) is obtained by applying it to every $J \subseteq \{1, \ldots, k\}$ in the place of $\{1, \ldots, k\}$ and making linear combinations of the resulting equalities so as to eliminate all products of less than k functions. This provides an expression of the product of the f_i's by means of their linear combinations in $\mathcal{BF}_n$, which will play a role in the sequel:

$$\prod_{i=1}^{l} f_i = \frac{1}{2^{l-1}} \sum_{\emptyset \neq J \subseteq \{1,\ldots,l\}} (-1)^{|J|-1} \left(\bigoplus_{i \in J} f_i \right). \tag{8}$$

Note that this relation can be easily checked by starting from the right-hand side of (8), applying (7) to $\bigoplus_{i \in J} f_i$ (instead of $\bigoplus_{i=1}^{k} f_i$), and observing that, for every $\emptyset \neq I \subseteq \{1, \ldots, k\}$, $\sum_{J;\, I \subseteq J \subseteq \{1,\ldots,l\}} (-1)^{|J|-1}$ equals $(-1)^{l-1}$ if $I = \{1, \ldots, l\}$ and is null otherwise.
Relation (8) has been originally obtained in [3], but it was proved in a more complex (and purely calculative) way.
Recall that $\mathbb{F}_2^n$ can be endowed with the structure of the finite field $\mathbb{F}_{2^n}$, since the latter is an n-dimensional vector space over $\mathbb{F}_2$. Any (n, n)-function, now viewed as a function from $\mathbb{F}_{2^n}$ to itself, admits a unique representation as a *univariate polynomial* over $\mathbb{F}_{2^n}$ in one variable and of (univariate) degree at most $2^n - 1$:

$$F(x) = \sum_{i=0}^{2^n-1} \delta_i x^i; \quad \delta_i \in \mathbb{F}_{2^n}. \tag{9}$$

Note that this works more generally for any (n, m)-function where m divides n, since such function is a particular case of an (n, n)-function, because $\mathbb{F}_{2^m}$ is a subfield of $\mathbb{F}_{2^n}$ (note that this includes the case $m = 1$). The algebraic degree of F can be directly read on this representation as well; it equals $\max_{j=0,\ldots,2^n-1;\, \delta_j \neq 0} w_2(j)$, where $w_2(j)$ is the Hamming weight of the binary expansion of j (see *e.g.* [5, 7]).
For n even, an $(n, n/2)$-function (or more generally an (n, m)-function where m divides $n/2$) can also be viewed as a function from $\mathbb{F}_{2^{n/2}}^2$ to $\mathbb{F}_{2^{n/2}}$ and represented in *bivariate form*: $\sum_{0 \leq i,j \leq 2^{n/2}-1} a_{i,j} x^i y^j$, where $a_{i,j} \in \mathbb{F}_{2^{n/2}}$.
The *Fourier-Hadamard transform* of the functions φ from $\mathbb{F}_2^n$ to $\mathbb{R}$ (called *pseudo-Boolean functions*) is the $\mathbb{R}$-linear mapping which maps φ to the func-

tion $\widehat{\varphi}$ defined on $\mathbb{F}_2^n$ by:

$$\widehat{\varphi}(u) = \sum_{x \in \mathbb{F}_2^n} \varphi(x) (-1)^{u \cdot x}, \tag{10}$$

where "$\cdot$" is some chosen inner product in $\mathbb{F}_2^n$. Given an n-variable Boolean function f, we have two associated transforms: the Fourier-Hadamard transform of f where f is then viewed as a function from $\mathbb{F}_2^n$ to $\{0,1\}$ and the *Walsh transform* of f which is the Fourier-Hadamard transform of the sign function $(-1)^f$:

$$W_f(u) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x) \oplus u \cdot x}.$$

We have:

$$W_f = 2^n \, \delta_0 - 2\widehat{f}, \tag{11}$$

where δ_0 denotes the *Dirac (or Kronecker) symbol*, *i.e.* the indicator of the singleton $\{(0, \ldots, 0)\}$, defined by $\delta_0(u) = 1$ if u is the null vector and $\delta_0(u) = 0$ otherwise.
The following result (see e.g. [4, 7]) will play a role in the sequel:

Proposition 2 *Let f be an n-variable Boolean function ($n \geq 2$), and let $1 \leq l \leq n$. Assume that the Walsh transform values of f are all divisible by 2^l (*i.e.*, according to Relation (11), that its Fourier-Hadamard transform takes values divisible by 2^{l-1}). Then f has algebraic degree at most $n - l + 1$.*

There is (see *e.g.* [4, 7] as well), a direct relationship between the values of the Walsh transform of a Boolean function and the coefficients of its NNF: for every $u \neq (0, \ldots, 0)$, we have:

$$W_f(u) = 2(-1)^{w_H(u)+1} \sum_{I \subseteq \{1,\ldots,n\};\ supp(u) \subseteq I} 2^{n-|I|} \lambda_I, \tag{12}$$

and, for $I \neq \emptyset$, we have:

$$\lambda_I = 2^{-n} (-2)^{|I|-1} \sum_{u \in \mathbb{F}_2^n;\ I \subseteq supp(u)} W_f(u). \tag{13}$$

This implies that f has numerical degree at most d if and only if $W_f(u) = 0$ for every vector u of Hamming weight strictly larger than d.

For vectorial functions, we define the Walsh transform as follows:

$$W_F(u, v) = W_{v \cdot F}(u) = \sum_{x \in \mathbb{F}_2^n} (-1)^{v \cdot F(x) \oplus u \cdot x};\ u \in \mathbb{F}_2^n, v \in \mathbb{F}_2^m.$$

In fact, the Walsh transform of F equals the Fourier-Hadamard transform of $1_{\mathcal{G}_F}$. Applying Proposition 2 to $1_{\mathcal{G}_F}$, we have then that, if the Walsh transform values of F are all divisible by 2^l, then $1_{\mathcal{G}_F}$ has algebraic degree at most $n+m-l$. *This latter property is stronger than Proposition 2 applied to the coordinate functions of F (or to its component functions), thanks to a bound that we shall obtain below in (17).*

3 State of the art on the algebraic degree of composite functions

There does not exist a general upper bound, except the naive (or trivial) bound seen in introduction (which is inefficient, except in some particular cases, such as when one of the functions is affine).
Two bounds exist in restricted frameworks.
The first bound, shown in [3], needs a very strong hypothesis for being efficient: if the Walsh transform W_F has all its values divisible by 2^l, then $d_{alg}(G \circ F) \leq n - l + d_{alg}(G)$. We shall call this bound the Canteaut-Videau bound. Because of its rather restrictive assumption, it is more useful as an indication, for cryptographers, on those S-boxes which should be avoided in a block cipher (like almost bent functions, see e.g. [5, 7]).
The second bound, shown in [1], does not apply to all functions either, but its assumption is always satisfied when dealing with the model of block ciphers called Substitution-Permutation networks: for every (n,n)-permutation F and any (n,r)-function G, we have $d_{alg}(G \circ F) \leq n - \left\lfloor \frac{n-1-d_{alg}(G)}{d_{alg}(F^{-1})} \right\rfloor$. This upper bound, that we shall call the Boura-Canteaut bound, shows that the algebraic degree of the computational inverse of a permutation plays a role in the algebraic degree of the iterated rounds implementing it. This is a precious indication for the designer of a block cipher. Let us then recall how it is proved in [1]. Firstly, the authors show by calculation that, for every (n,n)-permutation F and every integers k, l, the maximal algebraic degree of the product of at most k coordinate functions of F, that we shall denote by $d_{alg}^{[k]}(F)$, satisfies: $d_{alg}^{[k]}(F) < n - l \iff d_{alg}^{[l]}(F^{-1}) < n - k$ (we shall see that this can be obtained as a direct consequence of Relation (21) below, which will illustrate how graph indicators allow to simply prove and explain properties, that seem obscure without them). Secondly, the basic bound $d_{alg}(G \circ F) \leq d_{alg}^{[d_{alg}(G)]}(F)$ implies that $d_{alg}(G \circ F) \leq n - \left\lfloor \frac{n-1-d_{alg}(G)}{d_{alg}(F^{-1})} \right\rfloor$, by application of the equivalence $d_{alg}^{[d_{alg}(G)]}(F) < n-l \iff d_{alg}^{[l]}(F^{-1}) < n-d_{alg}(G)$ with $l = \left\lfloor \frac{n-1-d_{alg}(G)}{d_{alg}(F^{-1})} \right\rfloor$, using that $d_{alg}^{[l]}(F^{-1}) \leq l\, d_{alg}(F^{-1})$.
This proof does not give much insight on the reasons why this bound works (except for the important fact that it proves it, of course). Moreover, the bound is not very efficient when applying it iteratively, even with just one iteration (as we shall show in Subsection 5.2.4) and this is a limitation to its practical impact. It would then be useful to find an approach which would also apply efficiently to the composition of more than two functions, and this requires to understand what structure and properties are behind the bound.

4 ANF and bivariate representation of the graph indicator of a vectorial function

Denoting by $1_{\mathcal{G}_F}(x,y)$ the indicator (i.e. the characteristic function) of the *graph* $\mathcal{G}_F = \{(x, F(x));\ x \in \mathbb{F}_2^n\}$ of an (n,m)-function F, whose value is 1 if $F(x) = y$ and 0 otherwise, we have:

Proposition 3 *[6] Let F be any (n,m)-function and let $f_1, \ldots, f_m$ be its coordinate functions. Denoting $\{1, \ldots, m\} \setminus J$ by J^c, we have:*

$$1_{\mathcal{G}_F}(x,y) = \prod_{j=1}^{m}(y_j \oplus f_j(x) \oplus 1) = \bigoplus_{J \subseteq \{1,\ldots,m\}} \varphi_{F,J}(x) y^J,$$

where

$$\varphi_{F,J}(x) = \prod_{j \in J^c} (f_j(x) \oplus 1). \tag{14}$$

This is easily proved by observing that, for every $y, y' \in \mathbb{F}_2^m$, we have $y = y'$ if and only if $\prod_{j=1}^m (y_j \oplus y'_j \oplus 1) = 1$, and applying it with $y' = F(x)$. We deduce then:

Corollary 1 *Let F be any (n,m)-function, then with the notation of Proposition 3, we have:*

$$d_{alg}(1_{\mathcal{G}_F}) = \max_{J \subseteq \{1,\ldots,m\}} \Big(d_{alg}\big(\varphi_{F,J}(x)\big) + |J| \Big), \tag{15}$$

$$\begin{aligned} d_{alg}(F) &= \max_{|J|=m-1} d_{alg}(\varphi_{F,J}(x)) && (16)\\ &\leq d_{alg}(1_{\mathcal{G}_F}) - (m-1). && (17) \end{aligned}$$

Relations (15), (16) and (17) are valid for every vectorial function F thanks to our convention that the zero Boolean function has same algebraic degree 0 as the constant function 1.
The upper bound in (17) is tight. For instance, it is an equality when F is affine, since the graph indicator has then algebraic degree m. It is also achieved with equality when F is the multiplicative inverse function, that is, has univariate form $F(x) = x^{2^n-2}$, $x \in \mathbb{F}_{2^n}$, since $d_{alg}(F)$ equals then $n-1$, and we know from [6] that $1_{\mathcal{G}_F}(x,y)$ equals then $x^{2^n-1} + y^{2^n-1} + \sum_{j=0}^{2^n-2}(xy)^j$, and has algebraic degree $2\,n-2$.

The following result, that is a straightforward consequence of Relation (15), is not very strong, but it may be useful in some particular cases:

Corollary 2 *For every (n,m)-function F such that, for each $j = 1, \ldots, m$, the j-th coordinate function f_j of F has algebraic degree at least 1, we have $d_{alg}(1_{\mathcal{G}_F}) \leq \sum_{j=1}^m d_{alg}(f_j)$.*

If $F(x)$ is given in univariate representation, then we have, for every $x, y \in \mathbb{F}_{2^n}$, that

$$1_{\mathcal{G}_F}(x,y) = 1 + (y + F(x))^{2^n-1} = 1 + \sum_{j=0}^{2^n-1} y^{2^n-1-j}(F(x))^j, \tag{18}$$

and then we have:

$$d_{alg}(1_{\mathcal{G}_F}) = \max_{0 \le j \le 2^n-1} \left[d_{alg}\left((F(x))^j\right) + n - w_2(j) \right].$$

4.1 Case where F is bijective

4.1.1 Representation by the ANF

If F is a *permutation* (assuming $m = n$), then we have $1_{\mathcal{G}_F}(x,y) = 1_{\mathcal{G}_{F^{-1}}}(y,x)$, where F^{-1} is the compositional inverse of F, and thus, as observed in [6], if we use the alternative decomposition:

$$1_{\mathcal{G}_F}(x,y) = \bigoplus_{I \subseteq \{1,\ldots,n\}} \psi_{F,I}(y)\, x^I, \tag{19}$$

(that is valid for every function) we have, when F is bijective:

$$\psi_{F,I}(y) = \prod_{i \in \{1,\ldots,n\} \setminus I} (f'_i(y) \oplus 1), \tag{20}$$

where f'_i is the i-th coordinate function of F^{-1}.
We have then the following relation between the coordinate functions of F and F^{-1}:

$$\bigoplus_{J \subseteq \{1,\ldots,m\}} y^J \prod_{j \in \{1,\ldots,m\} \setminus J} (f_j(x) \oplus 1) = \bigoplus_{I \subseteq \{1,\ldots,n\}} x^I \prod_{i \in \{1,\ldots,n\} \setminus I} (f'_i(y) \oplus 1). \tag{21}$$

This directly implies that $d_{alg}^{[k]}(F) < n-l \iff d_{alg}^{[l]}(F^{-1}) < n-k$, by considering I such that $|I| = n-l$ and J such that $|J| = n-k$, since $\prod_{j \in \{1,\ldots,n\} \setminus J}(f_j(x) \oplus 1) = \bigoplus_{J' \subseteq \{1,\ldots,n\} \setminus J} \prod_{j \in J'} f_j(x)$ has algebraic degree at most $d_{alg}^{[n-|J|]}(F)$.

4.1.2 Representation in bivariate form

As already observed in [6], we have:

$$1_{\mathcal{G}_F}(x,y) = 1_{\mathcal{G}_{F^{-1}}}(y,x) = 1 + \sum_{j=0}^{2^n-1} x^{2^n-1-j}(F^{-1}(y))^j. \tag{22}$$

5 Related bounds on the algebraic degree of composite functions

In this section, we shall first recall what is known on the graph indicators of composite functions $G \circ F$. Then in Subsection 5.1, we shall derive simple bounds on the algebraic degree of $G \circ F$ for general functions, by means of the algebraic degrees of the graph indicators of the functions and of those of the functions themselves. One of these bounds is very efficient. It is deduced from an exact expression of $d_{alg}(G \circ F)$ that we shall derive and which has its own interest. It implies the Canteaut-Videau bound as a simple corollary. We shall generalize these results to three functions. Subsequently, in Subsection 5.2, we shall study the case where F is a permutation and prove a bound which slightly improves upon the Boura-Canteaut bound. Our approach will give insight on the reasons why this bound is true. We will generalize the approach to the compositions of three functions and obtain bounds more efficient than the iteration of the Boura-Canteaut bound and its slight improvement.
We start with the following formula from [6], expressing the graph indicator of $G \circ F$ by means of those over F and G, and involving only additions and multiplications: for every (n, m)-function F and every (m, r)-function G, we have:

$$1_{\mathcal{G}_{G\circ F}}(x, z) = \sum_{y \in \mathbb{F}_2^m} 1_{\mathcal{G}_F}(x, y) 1_{\mathcal{G}_G}(y, z) = \bigoplus_{y \in \mathbb{F}_2^m} 1_{\mathcal{G}_F}(x, y) 1_{\mathcal{G}_G}(y, z). \tag{23}$$

According to Relation (23) and to Proposition 3, and using Relation (19) and Proposition 1, we have then:

$$\begin{aligned} 1_{\mathcal{G}_{G\circ F}}(x, z) &= \bigoplus_{\substack{I \subseteq \{1,\ldots,n\} \\ K \subseteq \{1,\ldots,r\}}} x^I z^K \left(\bigoplus_{y \in \mathbb{F}_2^n} \left(\psi_{F,I}(y) \varphi_{G,K}(y) \right) \right) \\ &= \bigoplus_{\substack{I \subseteq \{1,\ldots,n\}, K \subseteq \{1,\ldots,r\}; \\ d_{alg}\left(\psi_{F,I}(y) \prod_{k \in K^c} (g_k \oplus 1)(y) \right) = m}} x^I z^K, \end{aligned} \tag{24}$$

where $K^c = \{1, \ldots, r\} \setminus K$ and the g_k's are the coordinate functions of G. This implies:

$$d_{alg}\left(1_{\mathcal{G}_{G\circ F}}\right) =$$

$$\max \left\{ |I| + |K|; \begin{array}{l} I \subseteq \{1, \ldots, n\} \\ K \subseteq \{1, \ldots, r\} \end{array} ; d_{alg}\left(\psi_{F,I}(y) \prod_{k \in K^c} (g_k \oplus 1) \right) = m \right\}. \tag{25}$$

Relation (25) provides an efficient information for deriving bounds on the algebraic degree, because it deals with a multiplication instead of a composition, and the algebraic degree of the product of two Boolean functions is bounded above by the sum of their algebraic degrees while the algebraic degree of the composition of two vectorial functions is bounded above by the product of their

algebraic degrees.
For being able to exploit Relation (25), we need to have an expression of $\psi_{F,I}(y)$. A first observation is that, according to Relation (4) applied to $1_{\mathcal{G}_F}(x,y) = \bigoplus_{I\subseteq\{1,\ldots,n\}} \psi_{F,I}(y)\, x^I$, we have:

$$\psi_{F,I}(y) = \bigoplus_{x\in\mathbb{F}_2^n;\, supp(x)\subseteq I} 1_{\mathcal{G}_F}(x,y) = |F^{-1}(y)\cap E_I| \text{ [mod 2]}, \tag{26}$$

where $E_I = \{x \in \mathbb{F}_2^n;\, supp(x) \subseteq I\}$. This relation has the interest of being completely general. But when F is bijective, we shall derive a more exploitable one. This is why we shall treat the case where F is a permutation apart.

5.1 Algebraic degree of composite functions in general

5.1.1 Case of two functions

We observe first that Relation (23) implies that:

$$d_{alg}(1_{\mathcal{G}_{G\circ F}}) \leq d_{alg}(1_{\mathcal{G}_F}) + d_{alg}(1_{\mathcal{G}_G}).$$

Note the similarity with the bound $d_{alg}(fg) \leq d_{alg}(f) + d_{alg}(g)$ on general Boolean functions. But this latter bound is tight while the former is not, except in some extreme cases, and we have $d_{alg}(1_{\mathcal{G}_F})+d_{alg}(1_{\mathcal{G}_G}) \geq m+r$, which implies that m needs to be small enough, for the bound to be useful. This comes from the fact that the Hamming weights of the exponents of y are taken into account when calculating $d_{alg}(1_{\mathcal{G}_F})$ and $d_{alg}(1_{\mathcal{G}_G})$, while they are not when calculating the algebraic degree of the function $(x,z) \mapsto \bigoplus_{y\in\mathbb{F}_2^m} 1_{\mathcal{G}_F}(x,y)1_{\mathcal{G}_G}(y,z)$.
We deduce by using (17) that, for every (n,m)-function F and every (m,r)-function G, we have:

$$d_{alg}(G\circ F) \leq d_{alg}(1_{\mathcal{G}_F}) + d_{alg}(1_{\mathcal{G}_G}) - (r-1). \tag{27}$$

Relation (27) is often better than the naive (or trivial) bound $d_{alg}(G\circ F) \leq (d_{alg}(F))\,(d_{alg}(G))$, but it is also most often too weak for giving any information (when its right hand side term is larger than n).

We shall show now that we can obtain a much better bound, after deriving an exact expression of $d_{alg}(G\circ F)$. Relation (24), and Relation (16) applied to $G\circ F$ instead of F, imply:

$$d_{alg}(G\circ F) =$$
$$\max_{k\in\{1,\ldots,r\}} \left(\max\left\{|I|;\, I\subseteq\{1,\ldots,n\}; d_{alg}\left((g_k\oplus 1)\psi_{F,I}\right) = m\right\}\right). \tag{28}$$

Remark. According to Proposition 1 and to Relation (26), and because of the equality $\bigcup_{y\in g_k^{-1}(0)} F^{-1}(y) = F^{-1}(g_k^{-1}(0))$, we have then:

$$d_{alg}(G\circ F) = \max_{k\in\{1,\ldots,r\}} \left(\max\left\{|I|;\, I\subseteq\{1,\ldots,n\}; |F^{-1}(g_k^{-1}(0))\cap E_I| \text{ odd}\right\}\right),$$

where E_I is defined after Relation (26), and this is also what gives Relation (4) applied to $G \circ F$. □

Relation (28) leads to the following bound:

Theorem 1 *For every (n,m)-function F and every (m,r)-function G, we have:*

$$d_{alg}(G \circ F) \leq d_{alg}(1_{\mathcal{G}_F}) + d_{alg}(G) - m.$$

Proof. The equality $1_{\mathcal{G}_F}(x,y) = \bigoplus_{I \subseteq \{1,\ldots,n\}} \psi_{F,I}(y)\, x^I$ implies that, for every $I \subseteq \{1,\ldots,n\}$, we have $d_{alg}(\psi_{F,I}) \leq d_{alg}(1_{\mathcal{G}_F}) - |I|$. The condition $d_{alg}\left((g_k \oplus 1)\psi_{F,I}\right) = m$ in Relation (28) implies then that $m \leq d_{alg}(1_{\mathcal{G}_F}) - |I| + d_{alg}(g_k)$, that is, $|I| \leq d_{alg}(1_{\mathcal{G}_F}) + d_{alg}(g_k) - m$. Relation (28) completes the proof. □

Strengths and weaknesses of the bound of Theorem 1:

1. The bound is simple and general. Thanks to (17) applied to G, with r instead of m, it is always much better than the bound in (27). Let us study its tightness. It is easily seen that it is achieved with equality when F is an affine automorphism, since $d_{alg}(1_{\mathcal{G}_F})$ equals then n. It is also achieved with equality when F is the multiplicative inverse function and G is an affine automorphism, since we have seen that $d_{alg}(1_{\mathcal{G}_F})$ equals then $2\,n-2$ and since $d_{alg}(F)$ equals $n-1$. Let us give now an example where none of the functions has algebraic degree 1: we have seen in [6], when studying the so-called switching method, that taking an (n,m)-function $F = (f_1,\ldots,f_m)$ and an n-variable Boolean function f, and denoting $F_f = (f_1,\ldots,f_{m-1},f_m \oplus f)$, $F' = (f_1,\ldots,f_{m-1})$ and $y' = (y_1,\ldots,y_{m-1})$, we have $1_{\mathcal{G}_{F_f}}(x,y) = 1_{\mathcal{G}_F}(x,y) + f(x)\,1_{\mathcal{G}_{F'}}(x,y')$. If F is affine and f is non-affine then $d_{alg}(1_{\mathcal{G}_{F_f}}) = d_{alg}(f)+m-1$. Note that $G\circ F_f(x) = G\circ F(x) + f(x)D_{(0,\ldots,0,1)}G(x)$. If $d_{alg}(D_{(0,\ldots,0,1)}G) = d_{alg}(G)-1$, then for every f such that $d_{alg}(f(x)D_{(0,\ldots,0,1)}G(x)) = d_{alg}(f) + d_{alg}(D_{(0,\ldots,0,1)}G)$, the bound of Theorem 1 for $G \circ F_f$ is an equality.

2. The bound, that has also the interest of depending on F and G separately, provides an important tool for designers when they choose S-boxes supposed to increase the algebraic degree of functions by composition:

Function F must have a graph indicator of algebraic degree significantly larger than m (more precisely, as close to $n+m-1$ as possible), for allowing that $G \circ F$ has algebraic degree significantly larger than G.

Remark. It is interesting to see that the multiplicative inverse function, which has been chosen as S-box in the AES (see [10]), fulfills this criterion, since we have seen that its graph indicator has algebraic degree $2\,n-2$. We need however to moderate this observation: it is easily seen with Relation (17) that, for every (n,m)-function F, we have $d_{alg}(F)+m-1 \leq d_{alg}(1_{\mathcal{G}_F}) \leq n+m-1$ (the right hand-side inequality coming from the fact that $1_{\mathcal{G}_F}$ has even Hamming weight), and for every (n,m)-function F of algebraic degree at most $n-1$, we have $d_{alg}(F)+m-1 \leq d_{alg}(1_{\mathcal{G}_F}) \leq n+m-2$ (since the coefficients of $x^{\{1,\ldots,n\}}y^J$

and $x^I y^{\{1,\ldots,m\}}$ equal 0, for every I, J such that $|I| \geq n-1$ and $|J| \geq m-1$), and any (n,m)-function of algebraic degree $n-1$, such as the multiplicative inverse function, satisfies then $d_{alg}(1_{\mathcal{G}_F}) = n+m-2$. □

3. The bound has also the advantage of being valid without the assumption that $n = m$. In the case $n = m$, we shall derive, in the particular case where F is bijective, a bound that does not involve $d_{alg}(1_{\mathcal{G}_F})$, that may be hard to calculate in some cases, but only $d_{alg}(F^{-1})$ and $d_{alg}(G)$, and that is slightly stronger than the Boura-Canteaut bound. Before that, let us illustrate with an example how the two bounds are evaluated, and compare their values. Let us take for instance $n = m$ odd, $x \in \mathbb{F}_{2^n}$, $F(x) = x^{2^i+1}$, with $i < n/2$ and $\gcd(i,n) = 1$ (this power function is called a Gold function; under these conditions, it is a permutation). The graph indicator of F equals $1_{\mathcal{G}_F}(x,y) = 1+(y+x^{2^i+1})^{2^n-1} = 1+\sum_{j=0}^{2^n-1} y^{2^n-1-j}\, x^{j(2^i+1)}$ and we have then $d_{alg}(1_{\mathcal{G}_F}) = \max_{0\leq j\leq 2^n-1}(w_2(2^n-1-j)+w_2(j(2^i+1))) = \max_{0\leq j\leq 2^n-1}(n-w_2(j)+w_2(j(2^i+1)))$. The bound of Theorem 1 gives then $d_{alg}(G\circ F) \leq \max_{0\leq j\leq 2^n-1}(w_2(j(2^i+1))-w_2(j))+d_{alg}(G)$ and the Boura-Canteaut bound gives (since the algebraic degree of F^{-1} equals $\frac{n+1}{2}$ as shown in [11]) $d_{alg}(G\circ F) \leq n - \left\lfloor \frac{n-1-d_{alg}(G)}{d_{alg}(F^{-1})} \right\rfloor = n - \left\lfloor \frac{2(n-1-d_{alg}(G))}{n+1} \right\rfloor$, which equals n or $n-1$, while the bound of Theorem 1 is much better. For instance, taking $r = n$ and $G(x) = x^{2^{i'}+1}$, with $i \neq i' < n/2$, we have checked for small values of n that, in average, the value of the bound of Theorem 1 is about half the value of the Boura-Canteaut bound.

4. However, there are some cases where the bound of Theorem 1 is weak (which is probably inevitable for a bound valid without any constraint). Take for instance for F the multiplicative inverse function, and for G a power function $G(x) = x^d$. Then $G\circ F(x) = x^{2^n-1-d}$ has algebraic degree $n - w_2(d) = n - d_{alg}(G)$, while the bound gives $d_{alg}(G\circ F) \leq n-2+d_{alg}(G)$. The bound is then efficient only when G has low algebraic degree. There are even some extreme cases where the bound is weaker than the naive bound. Take for instance G affine. Then the naive bound writes $d_{alg}(G\circ F) \leq d_{alg}(F)$ (and if G is a permutation, there is equality) but our bound may be far from this value, because $1_{\mathcal{G}_F}$ may have algebraic degree much larger than $m + d_{alg}(F) - 1$.

Improved bound, also efficient when G is affine (and exact when G is an affine permutation):
The next result shows that, in the bound of Theorem 1, the algebraic degree of $1_{\mathcal{G}_F}$ may be replaced by that of a function whose ANF is a part of that of $1_{\mathcal{G}_F}(x,y)$.

Theorem 2 *For every (n,m)-function F and every (m,r)-function G, let the functions $\varphi_{F,J}$ be defined by Relation (14) and let*

$$h_{F,G}(x,y) = \bigoplus_{\substack{J\subseteq\{1,\ldots,m\};\\ |J|\geq m-d_{alg}(G)}} \varphi_{F,J}(x)\, y^J. \tag{29}$$

We have:

$$d_{alg}(G\circ F) \leq d_{alg}(h_{F,G}) + d_{alg}(G) - m.$$

Proof. For every $I \subseteq \{1,\ldots,n\}$, the only terms in $\psi_{F,I}(y)$ which can ensure, for some k, the condition $d_{alg}\left((g_k \oplus 1)\psi_{F,I}\right) = m$ of Relation (28), have degree at least $m - d_{alg}(G)$ and then, when multiplied by x^I, are terms of $h_{F,G}$. Writing Relation (29) in the form $h_{F,G}(x,y) = \bigoplus_{I\subseteq\{1,\ldots,n\}} \psi_{F,G,I}(y)\, x^I$, the rest of the proof is similar to the proof of Theorem 1: we have then $d_{alg}(G\circ F) = \max_{k\in\{1,\ldots,r\}} \left(\max\left\{|I|;\, I\subseteq\{1,\ldots,n\}; d_{alg}\left((g_k\oplus 1)\psi_{F,G,I}\right) = m\right\}\right)$; the relation $d_{alg}\left((g_k\oplus 1)\psi_{F,G,I}\right) = m$ implies that $m \leq d_{alg}(h_{F,G}) - |I| + d_{alg}(g_k)$, that is, $|I| \leq d_{alg}(g_k) + d_{alg}(h_{F,G}) - m$. □

If G is affine, then $h_{F,G}(x,y) = \bigoplus_{\substack{J\subseteq\{1,\ldots,m\};\\ |J|\geq m-1}} \varphi_{F,J}(x)\, y^J$. According to relation (14), we have $d_{alg}(\varphi_{F,J}) \leq d_{alg}(F)$ for every $J \subseteq \{1,\ldots,m\}$ such that $|J| = m-1$ and $d_{alg}(\varphi_{F,\{1,\ldots,m\}}) = 0$. Hence, if F is not constant, we have $d_{alg}(h_{F,G}) \leq d_{alg}(F) + m - 1$, and the bound of Theorem 2 gives then $d_{alg}(G\circ F) \leq d_{alg}(F)$ and the bound is tight when G is affine.

We have seen in Section 2 that, if the Walsh transform values of an (n,m)-function F are all divisible by 2^l, then $1_{\mathcal{G}_F}$ has algebraic degree at most $n+m-l$. We immediately deduce then from Theorem 1:

Corollary 3 *Given three positive integers n, m, l such that $1 \leq l \leq n$, for every (n,m)-function F whose Walsh transform values are all divisible by 2^l and every (m,r)-function G, we have:*

$$d_{alg}(G\circ F) \leq n - l + d_{alg}(G).$$

This is exactly the Canteaut-Videau bound, which happens then to be a direct consequence of Theorem 1 and whose explanation becomes crystal clear. Note that using Theorem 2 instead of Theorem 1 would give the same bound under a different hypothesis.

5.1.2 Case of more than two functions

Another advantage of our approach is that Theorems 1 and 2 can be generalized to several functions. Let us consider the case of three functions. For every (n,m)-function F, every (m,r)-function G and every (r,s)-function H, we have, according to Relations (16) and (23):

$$d_{alg}(H\circ G\circ F) =$$

$$\max_{k\in\{1,\ldots,s\}} \left(\max\left\{|I|;\, I\subseteq\{1,\ldots,n\}; d_{alg}\left(\psi_{F,I}(y)1_{\mathcal{G}_G}(y,z)(h_k(z)\oplus 1)\right) = m+r\right\}\right),$$

and $d_{alg}\left(\psi_{F,I}(y)1_{\mathcal{G}_G}(y,z)(h_k(z)\oplus 1)\right) = m+r$ implies that $m+r \leq d_{alg}\left(\psi_{F,I}\right) + d_{alg}\left(1_{\mathcal{G}_G}\right) + d_{alg}(h_k(z)\oplus 1) \leq d_{alg}\left(1_{\mathcal{G}_F}\right) + d_{alg}\left(1_{\mathcal{G}_G}\right) + d_{alg}(h_k(z)\oplus 1) - |I|$. Moreover, with the same arguments as in the proof of Theorem 2, we can replace in this reasoning $1_{\mathcal{G}_F}(x,y)$ by $h_{F,G,H}(x,y) = \bigoplus_{\substack{J\subseteq\{1,\ldots,m\};\\ |J|\geq m+r-d_{alg}(1_{\mathcal{G}_G})-d_{alg}(H)}} \varphi_{F,J}(x)\, y^J$. Therefore:

Theorem 3 *For every (n,m)-function F, every (m,r)-function G and every (r,s)-function H, we have:*

$$d_{alg}(H \circ G \circ F) \leq d_{alg}\left(1_{\mathcal{G}_F}\right) + d_{alg}\left(1_{\mathcal{G}_G}\right) + d_{alg}(H) - m - r. \tag{30}$$

Moreover, we can replace $d_{alg}(1_{\mathcal{G}_F})$ by $d_{alg}(h_{F,G,H})$ in this evaluation, where

$$h_{F,G,H}(x,y) = \bigoplus_{\substack{J \subseteq \{1,\ldots,m\}\,; \\ |J| \geq m+r-d_{alg}(1_{\mathcal{G}_G})-d_{alg}(H)}} \varphi_{F,J}(x)\, y^J.$$

Note that the bound of Relation (30) is much better than what we get when applying Theorem 1 to the functions $G \circ F$ and H, using Relation (23) and bounding the algebraic degree of a product by the sum of the algebraic degrees (this gives a gap of m). The bound of Theorem 3 is tight (take F and G affine).

We can now address the case of the composition of more than three functions. By iterating the relation $1_{\mathcal{G}_{G\circ F}}(x,z) = \sum_{y \in \mathbb{F}_2^m} 1_{\mathcal{G}_F}(x,y) 1_{\mathcal{G}_G}(y,z)$, we obtain that when composing r functions $G_1, \ldots, G_r$ where G_t is from $\mathbb{F}_2^{m_{t-1}}$ to $\mathbb{F}_2^{m_t}$, we have:

$$\forall x \in \mathbb{F}_2^{m_0}, \forall z \in \mathbb{F}_2^{m_r},\ 1_{\mathcal{G}_{G_r \circ \cdots \circ G_1}}(x,z) =$$
$$\sum_{\substack{(y_1,\ldots,y_{r-1}) \in \\ \mathbb{F}_2^{m_1} \times \cdots \times \mathbb{F}_2^{m_{r-1}}}} \left(1_{\mathcal{G}_{G_1}}(x,y_1) \left(\prod_{t=2}^{r-1} 1_{\mathcal{G}_{G_t}}(y_{t-1},y_t) \right) 1_{\mathcal{G}_{G_r}}(y_{r-1},z) \right). \tag{31}$$

Relation (31) implies, according to the fact recalled above that a Boolean function has odd Hamming weight if and only if its algebraic degree equals its number of variables:

$$d_{alg}(1_{\mathcal{G}_{G_r \circ \cdots \circ G_1}}) = \max \Big\{ |I|+|J|;\ I \subseteq \{1,\ldots,m_0\}, J \subseteq \{1,\ldots,m_r\};$$
$$d_{alg}\Big(\psi_{G_1,I}(y_1) \Big(\prod_{t=2}^{r-1} 1_{\mathcal{G}_{G_t}}(y_{t-1},y_t) \Big) \varphi_{G_r,J}(y_{r-1}) \Big) = \sum_{t=1}^{r-1} m_t \Big\}. \tag{32}$$

And Relation (16) applied to $F = G_r \circ \cdots \circ G_1$ and $m = m_r$, implies:

$$d_{alg}(G_r \circ \cdots \circ G_1) = \max_{k \in \{1,\ldots,m_r\}} \Big(\max \Big\{ |I|;\ I \subseteq \{1,\ldots,m_0\};$$
$$d_{alg}\Big(\psi_{G_1,I}(y_1) \Big(\prod_{t=2}^{r-1} 1_{\mathcal{G}_{G_t}}(y_{t-1},y_t) \Big) \big((G_r)_k(y_{r-1}) + 1 \big) \Big) = \sum_{t=1}^{r-1} m_t \Big\} \Big), \tag{33}$$

where $(G_r)_k$ denotes the kth coordinate function of G_r.
Condition (33) implies that $\sum_{t=1}^{r-1} m_t \leq d_{alg}\left(\psi_{G_1,I}\right) + \sum_{t=2}^{r-1} d_{alg}(1_{\mathcal{G}_{G_t}}) + d_{alg}(G_r)$. Note that the relation $d_{alg}(1_{\mathcal{G}_{G_1}}) = \max_{I \subseteq \{1,\ldots,m_0\}} \Big(d_{alg}(\psi_{G_1,I}) + |I| \Big)$ implies $d_{alg}(\psi_{G_1,I}) \leq d_{alg}(1_{\mathcal{G}_{G_1}}) - |I|$, for every I. Moreover, for every $I \subseteq \{1,\ldots,m_1\}$, the only terms in $\psi_{G_1,I}(y_1)$ which can ensure Condition (33) for some k have degree at least $\sum_{t=1}^{r-1} m_t - \sum_{t=2}^{r-1} d_{alg}(1_{\mathcal{G}_{G_t}}) - d_{alg}(G_r)$. We deduce:

Theorem 4 *For every functions $G_1, \ldots, G_r$ where G_t is from $\mathbb{F}_2^{m_{t-1}}$ to $\mathbb{F}_2^{m_t}$, we have:*

$$d_{alg}(G_r \circ \cdots \circ G_1) \leq \sum_{t=1}^{r-1} d_{alg}(1_{\mathcal{G}_{G_t}}) + d_{alg}(G_r) - \sum_{t=1}^{r-1} m_t. \tag{34}$$

Moreover, in Relation (34), the first term $d_{alg}(1_{\mathcal{G}_{G_1}})$ of the first sum may be replaced, if $F(x)$ is given by its ANF, by the algebraic degree of the function

$$\sum_{\substack{J \subseteq \{1,\ldots,m_1\} \\ |J| \geq \sum_{t=1}^{r-1} m_t - \sum_{t=2}^{r-1} d_{alg}(1_{\mathcal{G}_{G_t}}) - d_{alg}(G_r)}} \varphi_{G_1,J}(x) y_1^J,$$

and if $F(x)$ is given in univariate form, by the algebraic degree of the function

$$\sum_{\substack{j \in \{0,\ldots,2^{m_1}-1\} \\ w_2(j) \geq \sum_{t=1}^{r-1} m_t - \sum_{t=2}^{r-1} d_{alg}(1_{\mathcal{G}_{G_t}}) - d_{alg}(G_r)}} \varphi_{G_1,j}(x) y_1^j,$$

where $\varphi_{G_1,j}$ is defined by $1_{\mathcal{G}_{G_1}}(x,y) = \sum_{j=0}^{2^{m_1}-1} \varphi_{G_1,j}(x) y^j$, that is, $\varphi_{G_1,j}(x) = \delta_0(j) + (F(x))^{2^{m_1}-1-j}$.

Theorem 4, whose bound (34) can be written in the form $d_{alg}(G_r \circ \cdots \circ G_1) \leq \sum_{t=1}^{r-1}(d_{alg}(1_{\mathcal{G}_{G_t}}) - m_t) + d_{alg}(G_r)$, confirms the criterion for each S-box in a block cipher: not only must its algebraic degree be large, but the algebraic degree of its graph indicator should be as far as possible from its number of output bits.
Note that in the rest of the theorem, the functions $\varphi_{G_1,J}$ are Boolean but the functions $\varphi_{G_1,j}$ are in general not. This second part of the theorem improves upon its first part only when the algebraic degree of G_r added with the sum of the algebraic degrees of the graph indicators of $G_1, \ldots, G_{r-1}$, is smaller than $\sum_{t=1}^{r-1} m_t$. If we compose a same (n,n)-function by itself r times for instance, this needs that the algebraic degree of the graph indicator be not much larger than n, and that r be small enough.

5.2 Algebraic degree of $G \circ F$ when F is a permutation

If F is a permutation, then we have seen in Subsection 4.1 that $\psi_{F,I}$ introduced in Relation (19) factorizes into $\prod_{i \in I^c}(f_i'(y) \oplus 1)$, where f_i' is the i-th coordinate function of F^{-1}. According to Relation (24), we have:

$$1_{\mathcal{G}_{G \circ F}}(x,z) = \bigoplus_{\substack{I \subseteq \{1,\ldots,n\}, K \subseteq \{1,\ldots,r\}; \\ d_{alg}(\prod_{i \in I^c}(f_i' \oplus 1) \prod_{k \in K^c}(g_k \oplus 1)) = n}} x^I z^K, \tag{35}$$

where $I^c = \{1, \ldots, n\} \setminus I$, $K^c = \{1, \ldots, r\} \setminus K$ and the g_k's are the coordinate functions of G.

Remark. If F is identity (that is, $f'_i(y) = y_i$ for every i), then this formula gives correctly $1_{\mathcal{G}_{G\circ F}}(x,z) = 1_{\mathcal{G}_G}(x,z)$, but this needs a little work to be checked. In a nutshell, given the function $h(y) = \prod_{k\in K^c}(g_k(y)\oplus 1)$, we have $d_{alg}\left(\prod_{i\in I^c}(y_i\oplus 1)h(y)\right) = n$ if and only the intersection of the support of h with the set $\{y\in\mathbb{F}_2^n;\ supp(y)\subseteq I\}$ has odd size, that is, according to Relation (4), if y^I has coefficient 1 in the ANF of $h(y)$; Proposition 3 completes then the argumentation. And if G is identity, the formula gives also $1_{\mathcal{G}_{G\circ F}}(x,z) = 1_{\mathcal{G}_F}(x,z)$, that can be checked similarly. The reason why checking $1_{\mathcal{G}_{G\circ Id}}(x,z) = 1_{\mathcal{G}_G}(x,z)$ is not direct is related to the fact that Relation (35) brings a different viewpoint on the composition of functions, and this will be illustrated by bounds. □

5.2.1 An exact evaluation of the degree of $G\circ F$ when F is bijective

Representation by the ANF Relation (35) implies:

$$d_{alg}\left(1_{\mathcal{G}_{G\circ F}}(x,z)\right) = \max_{\substack{I\subseteq\{1,\ldots,n\},K\subseteq\{1,\ldots,r\};\\ d_{alg}\left(\prod_{i\in I^c}(f'_i\oplus 1)\prod_{k\in K^c}(g_k\oplus 1)\right)=n}} (|I|+|K|), \tag{36}$$

which will be useful in the sequel. According to Relation (16) applied with $G\circ F$ instead of F, we have:

Theorem 5 *For any (n,n)-permutation F and any (n,r)-function G, we have:*

$$d_{alg}(G\circ F) = \max_{k\in\{1,\ldots,r\}}\left(\max\left\{|I|; d_{alg}\left((g_k\oplus 1)\prod_{i\in I^c}(f'_i\oplus 1)\right) = n\right\}\right), \tag{37}$$

where the f'_i's are the coordinate functions of the compositional inverse F^{-1} of F and the g_k's are the coordinate functions of G.

Important remark. In Theorem 5, the "$\oplus 1$"'s play no role. In other words, each coordinate function may be complemented or not in Relation (37). Indeed, complementing some coordinate functions of F^{-1} corresponds to applying a translation to the output of F^{-1}, that is, the input to F, and does not change the algebraic degree of $G\circ F$; and complementing some coordinate functions of G corresponds to translating the output of G, which does not change the algebraic degree either. □

Remark. According to Theorem 5, $G\circ F$ has algebraic degree strictly smaller than t if and only if the product of at most $n-t$ coordinate functions of F^{-1} and of one coordinate function of G never reaches algebraic degree n (i.e. has always even Hamming weight). And the fact that each of these coordinate functions can be complemented or not can be checked by another way: the product of at most $n-t$ functions, some being complemented, equals a sum of functions equal to the products of at most $n-t$ functions. □

Remark. Since F^{-1} is a permutation, the product of less than n of its coordinate functions (complemented or not) has even Hamming weight. Hence, in

Relation (37), it is thanks to the multiplication by $(g_k \oplus 1)$ that the product $(g_k \oplus 1)\prod_{i\in I^c}(f'_i \oplus 1)$ can reach algebraic degree n. Note that Theorem 5 also shows that, for every permutation F and every non-constant function G, there exists I such that $d_{alg}\left((g_k \oplus 1)\prod_{i\in I^c}(f'_i \oplus 1)\right) = n$. □

Representation in bivariate form From (18), (22) and (23) we deduce, as altready observed in [6], that for every (n,n)-permutation F and every (n,n)-function G, we have:

$$1_{\mathcal{G}_{G\circ F}}(x,z) =$$
$$\sum_{i=0}^{2^n-1}\left(\sum_{y\in\mathbb{F}_{2^n}}(F^{-1}(y))^i\right)x^{2^n-1-i} + \sum_{k=0}^{2^n-1}\left(\sum_{y\in\mathbb{F}_{2^n}}(G(y))^k\right)z^{2^n-1-k} +$$
$$\sum_{i,k\in\{0,\ldots,2^n-1\}}\left(\sum_{y\in\mathbb{F}_{2^n}}(F^{-1}(y))^i(G(y))^k\right)x^{2^n-1-i}z^{2^n-1-k}.$$

Gathering the terms, monomial by monomial, we obtain:

$$1_{\mathcal{G}_{G\circ F}}(x,z) =$$
$$\left(\sum_{y\in\mathbb{F}_{2^n}}\left((F^{-1}(y))^{2^n-1} + (G(y))^{2^n-1} + (F^{-1}(y))^{2^n-1}(G(y))^{2^n-1}\right)\right) +$$
$$\sum_{i=0}^{2^n-2}\left(\sum_{y\in\mathbb{F}_{2^n}}(F^{-1}(y))^i\left(1+(G(y))^{2^n-1}\right)\right)x^{2^n-1-i} +$$
$$\sum_{k=0}^{2^n-2}\left(\sum_{y\in\mathbb{F}_{2^n}}(G(y))^k\left(1+(F^{-1}(y))^{2^n-1}\right)\right)z^{2^n-1-k} +$$
$$\sum_{i,j\in\{0,\ldots,2^n-2\}}\left(\sum_{y\in\mathbb{F}_{2^n}}(F^{-1}(y))^i(G(y))^k\right)x^{2^n-1-i}z^{2^n-1-k}.$$

Note that, for $k=0$, we have $\sum_{y\in\mathbb{F}_{2^n}}(G(y))^k\left(1+(F^{-1}(y))^{2^n-1}\right) = 1$, since there is a unique y such that $F^{-1}(y)=0$. We deduce, using Proposition 1 and what is recalled in the introduction about the algebraic degree in polynomial representation:

$$d_{alg}(1_{\mathcal{G}_{G\circ F}}) = \max\left(n, \max_{\substack{i,k\in\{0,\ldots,2^n-2\}\\ d_{alg}((F^{-1}(y))^i(G(y))^k)=n}}\left(2n - w_2(i) - w_2(k)\right)\right).$$

As observed in [6], the univariate representation of a function $F(x)$ is obtained from $1_{\mathcal{G}_F}(x,z)$ as the x-dependent coefficient of z^{2^n-2}. Exploiting this property with $G \circ F$ instead of F, we deduce:

$$d_{alg}(G \circ F) = \max_{\substack{i \in \{0,\ldots,2^n-2\} \\ d_{alg}((F^{-1}(y))^i G(y))=n}} \big(n - w_2(i)\big). \tag{38}$$

Remark. Since the inverse of F plays such important role, let us recall that the inverses of the exponents of known APN (in fact, AB) power permutations (n odd) and the algebraic degrees of the corresponding power functions (which are also APN (AB)) have been determined in [11]. □

5.2.2 Deduced bound on the algebraic degree of $G \circ F$

In Relation (37), the equality $d_{alg}\left((g_k \oplus 1)\prod_{i\in I^c}(f'_i \oplus 1)\right) = n$ implies $n \leq d_{alg}(G) + (n - |I|)\, d_{alg}(F^{-1})$, that is, $|I| \leq n - \frac{n-d_{alg}(G)}{d_{alg}(F^{-1})}$. We deduce:

Corollary 4 *For every (n,n)-permutation F and every (n,r)-function G, we have:*

$$d_{alg}(G \circ F) \leq n - \left\lceil \frac{n - d_{alg}(G)}{d_{alg}(F^{-1})} \right\rceil.$$

The Boura-Canteaut bound and that of Corollary 4 provide a criterion for the choice of S-boxes in block ciphers when they are permutations: *both algebraic degrees of F and F^{-1} should be as large a possible*. Indeed, this S-box may play the role of G or that of F, according to the situations.

For most pairs (F,G) of (n,n)-functions, the bound of Corollary 4 improves by one unit the Boura-Canteaut bound recalled in Section 3, since the inequality $\frac{n-1-d_{alg}(G)}{d_{alg}(F^{-1})} < \frac{n-d_{alg}(G)}{d_{alg}(F^{-1})}$ most often implies $\left\lfloor \frac{n-1-d_{alg}(G)}{d_{alg}(F^{-1})} \right\rfloor = \left\lceil \frac{n-d_{alg}(G)}{d_{alg}(F^{-1})} \right\rceil - 1$. The bound is tight, since it is an equality when F is affine (which is not the case of the Boura-Canteaut bound); another example for which it is an equality is when $G \circ F$ has algebraic degree n, that is, when G has (maximal) algebraic degree r (since F is a permutation, we have then $\sum_{x\in\mathbb{F}_2^n}(G \circ F)(x) = \sum_{y\in\mathbb{F}_2^n} G(y) \neq 0$; note that the Boura-Canteaut bound is also an equality in that case). It would be nice to be able to determine all the pairs of functions for which the bound is an equality.

Remark. The bound of Corollary 4 and the bound of Theorem 1 both are consequences of Relation (28). The bound of Theorem 1 exploits the fact that $d_{alg}(\psi_{F,I}) \leq d_{alg}(1_{\mathcal{G}_F}) - |I|$ (and is then weaker than the exact expression given by (28)) while Corollary 4 uses this exact expression, which results in (37), but also weakens it. And both bounds use that the algebraic degree of a product of functions is bounded above by the sum of their algebraic degrees, but in different ways, with different impairments. These two bounds and the naive bound are complementary; none is a corollary of one of the others. □

Remark. There are cases where the gap in the bound of Corollary 4 is small, see the remark after Theorem 6, in which we investigate a more general setting. There are also cases where the value of the bound is far from the exact value given by Relation (37). A first example is when $G = F^{-1}$: Relation (37) gives directly that $d_{alg}(G \circ F)$ equals the algebraic degree 1 of the identity function (indeed, for any I of size larger than 1, $(g_k \oplus 1)\prod_{i\in I^c}(f'_i \oplus 1)$ never reaches algebraic degree n, since $(g_k \oplus 1)\prod_{i\in I^c}(f'_i \oplus 1)$ equals the product of less than n functions $(f'_i \oplus 1)$), while the bound of Corollary 4 gives $d_{alg}(G \circ F) \leq n - \left\lceil \frac{n-d_{alg}(F^{-1})}{d_{alg}(F^{-1})} \right\rceil = \left\lfloor n(1 - \frac{1}{d_{alg}(F^{-1})}) \right\rfloor + 1$, and this latter value is arbitrarily larger than 1. □

Let us now apply Corollary 4 to $G \circ F$ in the place of G, and to F^{-1} in the place of F. We obtain $d_{alg}(G) \leq n - \frac{n-d_{alg}(G\circ F)}{d_{alg}(F)}$, that is:

Corollary 5 *For every (n,n)-permutation F and every (n,r)-function G, we have:*

$$d_{alg}(G \circ F) \geq n - (n - d_{alg}(G))\, d_{alg}(F). \tag{39}$$

This lower bound gives information only when $d_{alg}(G)$ is near n and $d_{alg}(F)$ is reasonably small, but it is worth mentioning.

We have also the following bound (that will be useful in relation with Theorem 8 below) that we easily deduce from Relation (36) and the fact that the equality $d_{alg}\left(\prod_{i\in I^c}(f'_i \oplus 1)\prod_{k\in K^c}(g_k \oplus 1)\right) = n$ in this relation implies $n \leq (n-|I|)\, d_{alg}(F^{-1}) + (r-|K|)\, d_{alg}(G) \leq (n+r-|I|-|K|)\max(d_{alg}(F^{-1}), d_{alg}(G))$, which implies:

Corollary 6 *For every (n,n)-permutation F and any (n,r)-function G, we have:*

$$d_{alg}(1_{\mathcal{G}_{G\circ F}}) \leq n + r - \left\lceil \frac{n}{\max(d_{alg}(F^{-1}), d_{alg}(G))} \right\rceil.$$

Note that, applying Relation (16) to $G \circ F$ instead of F, we deduce:

$$d_{alg}(G \circ F) \leq n + 1 - \left\lceil \frac{n}{\max(d_{alg}(F^{-1}), d_{alg}(G))} \right\rceil.$$

But this is never stronger than the bound of of Corollary 4, since, if $d_{alg}(G) \leq d_{alg}(F^{-1})$, then $n + 1 - \left\lceil \frac{n}{\max(d_{alg}(F^{-1}), d_{alg}(G))} \right\rceil = n - \left\lceil \frac{n-d_{alg}(F^{-1})}{d_{alg}(F^{-1})} \right\rceil \geq n - \left\lceil \frac{n-d_{alg}(G)}{d_{alg}(F^{-1})} \right\rceil$, and if $d_{alg}(G) \geq d_{alg}(F^{-1})$ then $n + 1 - \left\lceil \frac{n}{\max(d_{alg}(F^{-1}), d_{alg}(G))} \right\rceil = n - \left\lceil \frac{n-d_{alg}(G)}{d_{alg}(G)} \right\rceil \geq n - \left\lceil \frac{n-d_{alg}(G)}{d_{alg}(F^{-1})} \right\rceil$.

5.2.3 Case of the composition of three functions

To be able to evaluate the algebraic degree of the function whose output is that of the i-th round of a block cipher, it is necessary to address the compositions of more than two functions. We shall see that it is enough to extend Relation (35)

to three functions. Given any (n,n)-permutation F, any (n,m)-function G and any (m,r)-function H, we have, iterating Relation (23), that $1_{\mathcal{G}_{H\circ G\circ F}}(x,t) = \bigoplus_{y\in\mathbb{F}_2^m} 1_{\mathcal{G}_F}(x,y)1_{\mathcal{G}_{H\circ G}}(y,t) = \bigoplus_{y\in\mathbb{F}_2^m, z\in\mathbb{F}_2^r} 1_{\mathcal{G}_F}(x,y)1_{\mathcal{G}_G}(y,z)1_{\mathcal{G}_H}(z,t)$. Using Proposition 3 and Relations (14) and (20), we deduce:

$$\begin{aligned} & 1_{\mathcal{G}_{H\circ G\circ F}}(x,t) \\ = & \bigoplus_{\substack{I\subseteq\{1,\ldots,n\}\\ K\subseteq\{1,\ldots,r\}}} x^I t^K \left(\bigoplus_{\substack{y\in\mathbb{F}_2^n\\ z\in\mathbb{F}_2^m}} \left[\prod_{i\in I^c} (f_i'(y)\oplus 1) \prod_{j=1}^{m} (g_j(y)\oplus z_j\oplus 1) \prod_{k\in K^c} (h_k(z)\oplus 1) \right] \right), \end{aligned}$$

and according to Relation (16):

Theorem 6 *For any (n,n)-permutation F, any (n,m)-function G and any (m,r)-function H, we have:*

$$d_{alg}(H\circ G\circ F) = \max_{k\in\{1,\ldots,r\}} \left(\max\left\{|I|; d_{alg}\left(\theta_{k,I}(y,z)\right) = n+m\right\}\right), \tag{40}$$

where

$$\begin{aligned} \theta_{k,I}(y,z) &= \Big(\prod_{i\in I^c}(f_i'(y)\oplus 1)\Big)\Big(\prod_{j=1}^{m}(g_j(y)\oplus z_j\oplus 1)\Big)\big(h_k(z)\oplus 1\big) \\ &= \Big(\prod_{i\in I^c}(f_i'(y)\oplus 1)\Big)\big(1_{\mathcal{G}_G}(y,z)\big)\big(h_k(z)\oplus 1\big), \end{aligned}$$

the f_i''s are the coordinate functions of the compositional inverse F^{-1} of F, $1_{\mathcal{G}_G}$ is the graph indicator of G and the h_k's are the coordinate functions of H.

Remark. We deduce, taking F and H equal to identity, that, for every (n,m)-function G, we have $d_{alg}(G) =$

$$\max_{1\le k\le r}\left[\max\left\{|I|; d_{alg}\left(\Big(\prod_{i\in I^c}(y_i\oplus 1)\Big)\big(1_{\mathcal{G}_G}(y,z)\big)\big(z_k\oplus 1\big)\right) = n+m\right\}\right]. \tag{41}$$

Remark. In Theorem 6, function $1_{\mathcal{G}_G}$ is involved as a whole in the algebraic degree of $H\circ G\circ F$. Anyway, changing G for a CCZ equivalent function, even if it does not change the algebraic degree of $1_{\mathcal{G}_G}$, changes in general the algebraic degree of $H\circ G\circ F$. If we look for instance again at the case where F and H are both the identity function, we know (see [2]) that CCZ equivalence does not preserve the algebraic degree; this can be verified with Relation (41), since changing $\mathcal{G}_G$ into an affinely equivalent set may change the size of the index set I of maximal size satisfying the condition of Relation (41). □

Deduced bounds on the algebraic degree The equality $d_{alg}\left(\theta_{k,I}(y,z)\right) = n+m$ in Relation (40) implies:

$$\begin{aligned} n+m &\le d_{alg}\left(\prod_{i\in I^c}(f_i'(y)\oplus 1)(1_{\mathcal{G}_G}(y,z))(h_k(z)\oplus 1)\right) \\ &\le d_{alg}(H) + (n-|I|)\, d_{alg}(F^{-1}) + d_{alg}(1_{\mathcal{G}_G}) \end{aligned}$$

that is, $|I| \leq n - \frac{n+m-d_{alg}(1_{\mathcal{G}_G})-d_{alg}(H)}{d_{alg}(F^{-1})}$. We deduce:

Theorem 7 *For every (n,n)-permutation F, every (n,m)-function G and every (m,r)-function H, we have:*

$$d_{alg}(H \circ G \circ F) \leq n - \left\lceil \frac{n+m-d_{alg}(1_{\mathcal{G}_G})-d_{alg}(H)}{d_{alg}(F^{-1})} \right\rceil .$$

Remark. The bound of Theorem 8 is tight. For instance, if we take $n = m$ and F, G affine, then we have $n - \left\lceil \frac{n+m-d_{alg}(1_{\mathcal{G}_G})-d_{alg}(H)}{d_{alg}(F^{-1})} \right\rceil = d_{alg}(H)$, since $1_{\mathcal{G}_G}(y,z) = \prod_{j=1}^{n}(y_i \oplus z_i \oplus 1)$ has algebraic degree n. If $d_{alg}(H \circ G \circ F)$ equals n, then the bound is also clearly tight. There are also cases where the bound is not an equality but the difference is small. If we take $n = m = r$ and $F^{-1}(x) = x^{2^i+1}$, with $i < n/2$, n odd and $\gcd(i,n) = 1$ and $G(x) = x^{2^{i'}+1}$, with $i \neq i' < n/2$, $H(x) = x^{2^{i''}+1}$, with $i, i' \neq i'' < n/2$, the graph indicator of G equals $1_{\mathcal{G}_G}(y,z) = (z + y^{2^{i'}+1})^{2^n-1} + 1 = \sum_{j=0}^{2^n-1} z^{2^n-1-j}\, y^{j(2^{i'}+1)} + 1$ and has then algebraic degree equal to: $\max_{0 \leq j \leq 2^n-1}(n - w_2(j) + w_2(j(2^{i'}+1)))$ and $d_{alg}(H \circ G \circ F) = w_2\left(\frac{(2^{i''}+1)(2^{i'}+1)}{2^i+1} \pmod{(2^n-1)}\right)$.
For $n = 9, i = 2, i' = 3, i'' = 4$, we have $d_{alg}(H \circ G \circ F) = 6$ and the bound gives 7.
For $n = 25, i = 11, i' = 4, i'' = 12$, we have $d_{alg}(H \circ G \circ F) = 18$ and the bound gives 19.
There are also cases where the difference is large, for instance if we take $n = m = r$ and $F(x) = G(x) = H(x) = x^{2^i+1}$, with $i < n/2$, n odd and $\gcd(i,n) = 1$. The algebraic degree of F^{-1} has been given in [11], it equals $\frac{n+1}{2}$ and $d_{alg}(H \circ G \circ F) = w_2((2^i+1)^3) = w_2(2^{3i} + 2^{2i+1} + 2^{2i} + 2^{i+1} + 2^i + 1)$ equals 4, 5 or 6 (when these six powers of 2 are distinct modulo $2^n - 1$). The bound is often far from these three values. □

Remark. In the bound of Theorem 8, each function plays a distinct role: the first function F in the decomposition plays a role through the degree of its inverse, the last function H plays a role through its own degree and the function in the middle plays a role through the degree of the indicator of its graph. □

5.2.4 Comparison between Theorem 8 and the iterated known bound

Let us compare the bound of Theorem 8 with the Boura-Canteaut bound (or better, its slight improvement by Corollary 4) iterated once. For such iteration, we need that G be also bijective, with $m = n$. We take the optimum between the application of Corollary 4 to the pairs $(G \circ F, H)$ and $(F, H \circ G)$, and we

get:

$$\begin{aligned}
&d_{alg}(H\circ G\circ F)\\
&\leq n-\max\left(\left\lceil\frac{n-d_{alg}(H)}{d_{alg}(F^{-1}\circ G^{-1})}\right\rceil,\left\lceil\frac{n-d_{alg}(H\circ G)}{d_{alg}(F^{-1})}\right\rceil\right)\\
&\leq n-\max\left(\left\lceil\frac{n-d_{alg}(H)}{n-\left\lceil\frac{n-d_{alg}(F^{-1})}{d_{alg}(G^{-1})}\right\rceil}\right\rceil,\left\lceil\frac{n-d_{alg}(H)}{d_{alg}(G^{-1})\cdot d_{alg}(F^{-1})}\right\rceil\right). \quad (42)
\end{aligned}$$

If $\lambda,\mu,\nu\in[0,1]$ are such that $d_{alg}(F^{-1})=\lambda n$, $d_{alg}(G^{-1})=\mu n$ and $d_{alg}(H)=\nu n$, then (42) writes $d_{alg}(H\circ G\circ F)\leq n-\max\left(\left\lceil\frac{1-\nu}{1-\frac{\left\lceil\frac{1-\lambda}{\mu}\right\rceil}{n}}\right\rceil,\left\lceil\frac{1-\nu}{\mu\lambda n}\right\rceil\right)$ and gives then no real information since this will most often result in $d_{alg}(H\circ G\circ F)\leq n-1$. On the contrary, the bound of Theorem 8 still gives information: if additionally $\eta\in[0,1]$ is such that $d_{alg}(1_{\mathcal{G}_G})=n(1+\eta)$, this bound writes $d_{alg}(H\circ G\circ F)\leq n-\left\lceil\frac{1-\nu-\eta}{\lambda}\right\rceil$. And if this information is not significant enough, we still can try to evaluate with a finer grain the exact value given by Relation (40).

5.2.5 Case of the composition of more than three functions

Relations (19) and (20) imply in Relation (33) that if $m_0=m_1$ and G_1 is a permutation, then, denoting the coordinate functions of G_1^{-1} by $f'_1,\ldots,f'_{m_1}$, we have:

$$d_{alg}(1_{\mathcal{G}_{G_r\circ\cdots\circ G_1}})=$$

$$\max\Big\{|I|+|J|;\ I\subseteq\{1,\ldots,m_1\},J\subseteq\{1,\ldots,m_r\};d_{alg}\Big(\Big(\prod_{i\in\{1,\ldots,m_1\}\setminus I}(f'_i(y_1)+1)\Big)$$

$$\Big(\prod_{t=2}^{r-1}1_{\mathcal{G}_{G_t}}(y_{t-1},y_t)\Big)\Big(\prod_{j_r\in J^c}((G_r)_{j_r}(y_{r-1})+1)\Big)=\sum_{t=1}^{r-1}m_t\Big\}, \quad (43)$$

where, for every t and every j, $(G_t)_j$ is the jth coordinate function of G_t, $(y_t)_j$ is the jth coordinate of y_t, and $J^c=\{1,\ldots,m_r\}\setminus J$. And we have:

$$d_{alg}(G_r\circ\cdots\circ G_1)=$$

$$\max_{k\in\{1,\ldots,m_r\}}\Big(\max\Big\{|I|;\ I\subseteq\{1,\ldots,m_1\};d_{alg}\Big(\Big(\prod_{i\in\{1,\ldots,m_1\}\setminus I}(f'_i(y_1)+1)\Big)$$

$$\Big(\prod_{t=2}^{r-1}1_{\mathcal{G}_{G_t}}(y_{t-1},y_t)\Big)\big((G_r)_k(y_{r-1})+1\big)=\sum_{t=1}^{r-1}m_t\Big\}\Big). \quad (44)$$

The equality $d_{alg}\Big(\big(\prod_{i\in\{1,\ldots,m_1\}\setminus I}(f_i'(y_1)+1)\big)\Big(\prod_{t=2}^{r-1}1_{\mathcal{G}_{G_t}}(y_{t-1},y_t)\Big)\big((G_r)_k(y_{r-1})+1\big)\Big)=\sum_{t=1}^{r-1}m_t$ in Relation (44) implies:

$$\sum_{t=1}^{r-1}m_t\leq(m_1-|I|)\,d_{alg}(G_1^{-1})+\sum_{t=2}^{r-1}d_{alg}(1_{\mathcal{G}_{G_t}})+d_{alg}(G_r),$$

that is, $|I|\leq m_1-\frac{\sum_{t=1}^{r-1}m_t-d_{alg}(G_r)-\sum_{t=2}^{r-1}1_{\mathcal{G}_{G_t}}}{d_{alg}(G_1^{-1})}$. We deduce:

Theorem 8 *For every functions $G_1,\ldots,G_r$ where G_t is from $\mathbb{F}_2^{m_{t-1}}$ to $\mathbb{F}_2^{m_t}$, with $m_0=m_1$, and where G_1 is a permutation, we have:*

$$d_{alg}(G_r\circ\cdots\circ G_1)\leq m_1-\left\lceil\frac{\sum_{t=1}^{r-1}m_t-d_{alg}(G_r)-\sum_{t=2}^{r-1}d_{alg}(1_{\mathcal{G}_{G_t}})}{d_{alg}(G_1^{-1})}\right\rceil.$$

6 Practical calculation of the graph indicator and of the output's algebraic degree of the r-th round in a block cipher

As we recalled in the introduction, most block ciphers (for instance in the substitution-permutation network model or in a Feistel model or in any combination of such models) have rounds which are the combinations in some order of one or several substitution boxes (that are nonlinear functions), one or several diffusion layers (that are in general linear functions) and the addition of the round key. With Relation (44), we can see that what is central in the evaluation of the algebraic degree of the output of some round, viewed as a vectorial function whose input is given by the plaintext and the round keys, is the indicator of the graph of each function involved in each round. Let us see then how the indicators of the graphs of such functions can be obtained.

6.1 Bounds involving the numerical normal form of component functions

Relation (8), applied to $l-1$ coordinate functions of F^{-1} (or less) and one coordinate function of G shows that, if the coefficient of $\prod_{i=1}^n x_i$ in the NNF of any sum (mod 2) of at most $l-1$ coordinate functions of F^{-1} and at most one coordinate function of G is divisible by 2^l, then all the products of at most $l-1$ coordinate functions of F^{-1} and one coordinate function of G have algebraic degree strictly less than n, which implies that if $d_{alg}\left((g_k\oplus 1)\prod_{i\in I^c}(f_i'\oplus 1)\right)=n$ then $|I^c|\geq l$, that is $|I|\leq n-l$. According to Relation (37), we deduce:

Proposition 4 *For every (n,n)-permutation F and any (n,m)-function G, if for some $l \leq n$, the coefficient of $\prod_{i=1}^{n} x_i$ in the numerical normal form of any sum (mod 2) of at most $l-1$ coordinate functions of F^{-1} and at most one coordinate function of G is divisible by 2^l, then we have:*

$$d_{alg}(G \circ F) \leq n - l.$$

We know, according to Relation (13), that the coefficient $\lambda_{\{1,\ldots,n\}}$ of $\prod_{i=1}^{n} x_i$ in the numerical normal form of a function h equals $\frac{(-1)^{n-1}W_h(1_n)}{2}$, where 1_n is the all-1 vector. Then, the hypothesis of Proposition 4 is equivalent to assuming that $W_{\bigoplus_{i\in J} f'_i}(1_n)$ and $W_{g_k \oplus \bigoplus_{i\in J} f'_i}(1_n)$ are divisible by 2^{l+1} for every $k \in \{1,\ldots,m\}$ and every $J \subseteq \{1,\ldots,n\}$ such that $|J| \leq l-1$. Hence, changing l into $l-1$ so as to be able to compare with the Canteaut-Videau bound, we have:

Corollary 7 *For every (n,n)-permutation F and any (n,m)-function G, if for some $l \leq n$ and for every $v \in \mathbb{F}_2^n$ such that $w_H(v) \leq l-2$ and every $k \in \{1,\ldots,m\}$, both numbers $W_{g_k \oplus v\cdot F^{-1}}(1_n)$ and $W_{v\cdot F^{-1}}(1_n) = W_{F^{-1}}(1_n, v) = W_F(v, 1_n)$ are divisible by 2^l, then we have:*

$$d_{alg}(G \circ F) \leq n - l + 1.$$

This bound can be compared to the Canteaut-Videau bound. Each bound has advantages and disadvantages:

- Corollary 7 assumes F bijective, contrary to the Canteaut-Videau bound, which has then a weaker hypothesis from this viewpoint.
- In Corollary 7, the condition concerns $W_{g_k \oplus v\cdot F^{-1}}(1_n)$, while the Canteaut-Videau does not, and has then a weaker hypothesis from this viewpoint as well.
- In Corollary 7, the condition also concerns $W_{v\cdot F^{-1}}(u) = W_F(v,u)$, but only for $u = 1_n$ and $w_H(v) \leq l-2$, while the condition for the Canteaut-Videau bound deals with all the Walsh coefficients of F; the Canteaut-Videau bound has then a more demanding hypothesis from this viewpoint.
- Corollary 7 limits the algebraic degree to $n-l+1$, while the Canteaut-Videau limits it to $n-l+d_{alg}(G)$, and is then weaker from this viewpoint.

We see that the two bounds are complementary, since they are neither comparable from the viewpoint of their hypotheses, nor comparable from the viewpoint of their values.

We can also apply Relation (8) by keeping the factor $(g_k \oplus 1)$ apart. If, for some values $l, d \leq n$, all the coefficients of the terms of degrees strictly larger than d in any sum (mod 2) of at most l coordinate functions of F^{-1} are divisible by 2^l, then all the products of at most l coordinate functions of F^{-1} have algebraic degree at most d, which implies that if $d_{alg}\left((g_k \oplus 1)\prod_{i\in I^c}(f'_i \oplus 1)\right) =$

n then we have $n \leq d_{alg}(G) + d$, and therefore, taking $d = n - d_{alg}(G) - 1$ then $d_{alg}\left((g_k \oplus 1)\prod_{i\in I^c}(f'_i \oplus 1)\right) = n$ implies $|I^c| > l$, that is $|I| \leq n - l - 1$. According to Relation (37), we deduce then:

Proposition 5 *Let F be an (n,n)-permutation and G an (n,m)-function, and let $l \leq n$. If all the coefficients of the terms of degrees at least $n - d_{alg}(G)$ in the numerical normal forms of the sums (mod 2) of at most l coordinate functions of F^{-1} are divisible by 2^l, then we have:*

$$d_{alg}(G \circ F) \leq n - l - 1.$$

Relations (12) and (13) allow to translate by means of $W_{F^{-1}}$, and therefore of W_F, the fact that the coefficients of the terms x^I of degree $|I| \geq n - d_{alg}(G)$ in the numerical normal form of a linear combination $u \cdot F^{-1}$ are all divisible by 2^l. Relations (12) and (13), with v in the place of u and applied to $f = u \cdot F^{-1}$, give:

Corollary 8 *Let F be an (n,n)-permutation and G an (n,m)-function, and let $l \leq n$. If for every $u, v \in \mathbb{F}_2^n$ such that $w_H(u) \leq l$ and $w_H(v) \geq n - d_{alg}(G)$, the value $W_{F^{-1}}(v,u) = W_F(u,v)$ is divisible by $2^{n-w_H(v)+l+1}$, then we have:*

$$d_{alg}(G \circ F) \leq n - l - 1.$$

Note that, if a number is divisible by $2^{l+d_{alg}(G)+1}$, then it is divisible by $2^{n-w_H(v)+l+1}$, for every v such that $w_H(v) \geq n - d_{alg}(G)$. Then, replacing l by $l - d_{alg}(G) - 1$ in Corollary 8, so as to be able to compare with the Canteaut-Videau bound, and weakening the corollary (by strengthening its hypothesis), we deduce:

Corollary 9 *Let F be an (n,n)-permutation and G an (n,m)-function, and let $l \leq n - d_{alg}(G) - 1$. If for every $u \in \mathbb{F}_2^n$ such that $w_H(u) \leq l - d_{alg}(G) - 1$ and every $v \in \mathbb{F}_2^n$ such that $w_H(v) \geq n - d_{alg}(G)$, the value of $W_{F^{-1}}(v,u) = W_F(u,v)$ is divisible by 2^l, then we have:*

$$d_{alg}(G \circ F) \leq n - l + d_{alg}(G).$$

Let us now compare Corollary 9 with the Canteaut-Videau bound. The differences are:

- Corollary 9 assumes F bijective; this is a restriction with respect to the Canteaut-Videau bound.
- In Corollary 9, the condition concerns $W_F(v,u)$ only for $w_H(v) \leq l - d_{alg}(G) - 1$ and $w_H(u) \geq n - d_{alg}(G)$; this is an extension with respect to the Canteaut-Videau bound.
- Both bounds limit the algebraic degree to $n - l + d_{alg}(G)$, so there is no difference from this viewpoint.

6.1.1 Case of the composition of more than two functions

The Canteaut-Videau bound and the bounds of Proposition 4 and 5 or of Corollaries 7 and 9 address only the compositions of two functions. Let us see how we can extend them to bounding the algebraic degree of the composition of (say) three functions F, G, H.
A first possibility is to replace G by $H \circ G$ in the Canteaut-Videau bound, for instance, or in the bound of Corollary 9 (note that using Corollary 7 would be complex). Then, we need to bound $d_{alg}(H \circ G)$. A possibility is to assume that G satisfies the hypothesis on F in either the Canteaut-Videau bound, or the bound of Corollary 7, or that of Corollary 9, but in all cases, this means that the conditions needed are very strong for both F and G. We can also use the Boura-Canteaut bound or, better, the bound of Corollary 4, assuming that G is a permutation; then if (say) F and G are (n,n)-permutations and H is an (n,r)-function and if W_F has all its values divisible by 2^l, we have:

$$d_{alg}(H \circ G \circ F) \leq 2n - l - \left\lceil \frac{n - d_{alg}(H)}{d_{alg}(G^{-1})} \right\rceil . \tag{45}$$

Another solution, which may be more efficient, at least in some cases, is to use Theorem 6, that gives an exact value and does not need G bijective. So let us take for G any (n,m)-function. Relation (40): $d_{alg}(H \circ G \circ F) =$

$$\max_{k \in \{1,\ldots,r\}} \left(\max \left\{ |I|; d_{alg}\big(\big(\prod_{i \in I^c} (f_i'(y) \oplus 1) \big) \big(1_{\mathcal{G}_G}(y,z) \big) \big(h_k(z) \oplus 1 \big) \big) = n + m \right\} \right),$$

makes that all the methods we developed above for two functions can be used for three. For instance, let us assume again that, for some values $l, d \leq n$, all the coefficients of the terms of degrees strictly larger than d in any sum (mod 2) of at most l coordinate functions of F^{-1} are divisible by 2^l, then again all the products of at most l coordinate functions of F^{-1} have algebraic degree at most d, which implies that if $d_{alg}\big(\big(\prod_{i \in I^c} (f_i'(y) \oplus 1) \big) \big(1_{\mathcal{G}_G}(y,z) \big) \big(h_k(z) \oplus 1 \big) \big) = n + m$ then we have $n+m \leq d_{alg}(1_{\mathcal{G}_G}) + d_{alg}(H) + d$, and therefore that if $d = n+m-d_{alg}(1_{\mathcal{G}_G}) - d_{alg}(H) - 1$ then $d_{alg}\big(\big(\prod_{i \in I^c} (f_i'(y) \oplus 1) \big) \big(1_{\mathcal{G}_G}(y,z) \big) \big(h_k(z) \oplus 1 \big) \big) = n+m$ implies $|I^c| > l$, that is $|I| \leq n - l - 1$. According to Relation (37), we deduce then:

Proposition 6 *For every (n,n)-permutation F, any (n,m)-function G and any (m,r)-function H, if for some $l \leq n$, all the coefficients of the terms of degrees at least $n + m - d_{alg}(1_{\mathcal{G}_G}) - d_{alg}(H)$ in the numerical normal forms of the sums (mod 2) of at most l coordinate functions of F^{-1} are divisible by 2^l, then we have:*

$$d_{alg}(H \circ G \circ F) \leq n - l - 1.$$

We can, here also, deduce a bound involving the Walsh transform of F: according to Relations (12) and (13) applied with v in the place of u to the function $f = u \cdot F^{-1}$, saying that the coefficient of any term x^I such that $|I| \geq n + m - d_{alg}(1_{\mathcal{G}_G}) - d_{alg}(H)$ in the numerical normal form of $u \cdot F^{-1}$ (with

$w_H(u) \leq l$) is divisible by 2^l is equivalent to saying that, for every $u \in \mathbb{F}_2^n$ such that $w_H(u) \leq l$ and every $v \in \mathbb{F}_2^n$ such that $w_H(v) \geq n+m-d_{alg}(1_{\mathcal{G}_G})-d_{alg}(H)$, the value $W_{F^{-1}}(v,u) = W_F(u,v)$ is divisible by $2^{n-w_H(v)+l+1}$. Note that the divisibility by $2^{d_{alg}(1_{\mathcal{G}_G})-m+d_{alg}(H))+l+1}$ implies the divisibility by $2^{n-w_H(v)+l+1}$ when $w_H(v) \geq n+m-d_{alg}(1_{\mathcal{G}_G})-d_{alg}(H)$. We have then, replacing l by $l-d_{alg}(1_{\mathcal{G}_G})+m-d_{alg}(H))-1$:

Corollary 10 *For every (n,n)-permutation F, any (n,m)-function G and any (m,r)-function H, if for some $l \leq n-d_{alg}(1_{\mathcal{G}_G})+m-d_{alg}(H))-1$ and for every $u \in \mathbb{F}_2^n$ such that $w_H(u) \leq l-d_{alg}(1_{\mathcal{G}_G})+m-d_{alg}(H)-1$ and every $v \in \mathbb{F}_2^n$ such that $w_H(v) \geq n+m-d_{alg}(1_{\mathcal{G}_G})-d_{alg}(H)$, the value of $W_{F^{-1}}(v,u) = W_F(u,v)$ is divisible by $2^{l-d_{alg}(1_{\mathcal{G}_G})+m-d_{alg}(H)-1}$, then we have:*

$$d_{alg}(H \circ G \circ F) \leq n-m-l+d_{alg}(1_{\mathcal{G}_G})+d_{alg}(H).$$

Let us compare the bound of Corollary 10 with (45), for $m=n$ and G bijective (since this is needed by (45)). The bound of Corollary 10 writes $d_{alg}(H \circ G \circ F) \leq d_{alg}(1_{\mathcal{G}_G})+d_{alg}(H)-l$, to be compared with $d_{alg}(H \circ G \circ F) \leq 2n-l-\left\lceil \frac{n-d_{alg}(H)}{d_{alg}(G^{-1})} \right\rceil$. We can see that (45) is more efficient when the algebraic degrees of $1_{\mathcal{G}_G}$ and H are large and Corollary 10 is better when they are smaller. But note that the hypothesis of Corollary 10 does not assume the divisibility by 2^l for all Walsh transform values of F and is then weaker. Moreover, the advantage of the approach by Corollary 10 is that, if this bound is inefficient, we can try to use to the exact value given by Theorem 6 and bound it in a more precise way.

Conclusion

We have shown how the indicators of the graphs of vectorial functions can be used for studying the algebraic degree of vectorial functions. This approach has led to an exact expression of the algebraic degree of composite functions and to an efficient upper bound on it, that is valid without any condition on the functions. We have seen how this allows to prove more simply the known bounds (that all assume conditions), and to clarify why they work. It has in particular completely clarified why, when F is a permutation, the algebraic degree of $G \circ F$ depends on the algebraic degree of G and of the algebraic degree of F^{-1} (rather than that of F), and why the divisibility of the Walsh transform values of F plays a role. The approach by graph indicators has also led to new bounds that involve the numerical normal form of component functions. We have now three types of bounds: a first bound, that is completely general and however efficient, but may be delicate to be precisely evaluated; a second bound, that is simpler to evaluate, but is often weaker and assumes that F is bijective; and a third series of bounds, assuming strong conditions on the divisibility by powers of 2 of some coefficients in the numerical normal form of some functions or (equivalently) of the values taken by the Walsh transform. Moreover, we have derived a general upper bound on the algebraic degree of the composition of three functions, and

for $H \circ G \circ F$, where F is bijective, we have shown that the algebraic degree of the composite function essentially depends on the algebraic degrees of H and F^{-1}, and of the algebraic degree of the graph indicator of G. These bounds generalize to more than three functions. We have also generalized to three functions the bounds assuming divisibility by powers of 2. Our results give more insight for the designer of a block cipher on how optimizing the choice of S-boxes from the viewpoint of the algebraic degree of the round functions. In a future work, we shall study in detail some main block ciphers, starting in particular with the algebraic degree of the so-called SDS function, that intervenes in the two first rounds of the AES.

References

[1] C. Boura and A. Canteaut. On the Influence of the Algebraic Degree of F^{-1} on the Algebraic Degree of $G \circ F$. *IEEE Transactions on Information Theory* 59 (1), pp. 691-702, 2013. 3, 4, 9

[2] L. Budaghyan, C. Carlet, A. Pott. New Classes of Almost Bent and Almost Perfect Nonlinear Functions. *IEEE Transactions on Information Theory* 52 (3), pp. 1141-1152, March 2006. 2, 23

[3] A. Canteaut and M. Videau. Degree of composition of highly nonlinear functions and applications to higher order differential cryptanalysis. *Proceedings of EUROCRYPT 2002, Lecture Notes in Computer Science* 2332, pp. 518-533, 2002. See also https://hal.inria.fr/inria-00072221/document 3, 4, 7, 9

[4] C. Carlet. Boolean Functions for Cryptography and Error Correcting Codes. Chapter of the monography *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*, Y. Crama and P. Hammer eds, Cambridge University Press, pp. 257-397, 2010. 5, 8

[5] C. Carlet. Vectorial Boolean Functions for Cryptography. Chapter of the monography *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*, Y. Crama and P. Hammer eds, Cambridge University Press, pp. 398-469, 2010. 2, 5, 7, 9

[6] C. Carlet. Handling vectorial functions by means of their graph indicators. To appear in *IEEE Transactions on Information Theory*, 2020. 2, 3, 4, 10, 11, 12, 14, 20, 21

[7] C. Carlet. Boolean Functions for Cryptography and Coding Theory. Monograph to appear in Cambridge University Press, 2020. 2, 5, 7, 8, 9

[8] C. Carlet, P. Charpin, and V. Zinoviev. Codes, bent functions and permutations suitable for DES-like cryptosystems. *Designs, Codes and Cryptography*, 15(2), pp. 125-156, 1998. 2

[9] C. Carlet and P. Guillot. A new representation of Boolean functions. *Proceedings of AAECC'13, Lecture Notes in Computer Science* 1719, pp. 94-103, 1999. 3

[10] J. Daemen and V. Rijmen. AES proposal: Rijndael, 1999. See http://www.quadibloc.com/crypto/co040401.htm 14

[11] G. Kyureghyan and V. Suder. On inversion in $\mathbb{Z}_{2^n-1}$. *Finite Fields and Their Applications* 25, pp. 234-254, 2014. 15, 21, 24

[12] F. J. MacWilliams and N. J. Sloane. *The theory of error-correcting codes*, Amsterdam, North Holland. 1977.

[13] S. Mesnager. Linear codes from functions. *A Concise Encyclopedia of Coding Theory*. To appear. 2

[14] C.E. Shannon. Communication theory of secrecy systems. *Bell system technical journal*, 28, pp. 656-715, 1949. 2, 5