On the Influence of the Algebraic Degree of F^{-1} on the Algebraic Degree of $G \circ F$

Christina Boura and Anne Canteaut

***Abstract*—We present a study on the algebraic degree of iterated permutations seen as multivariate polynomials. The main result shows that this degree depends on the algebraic degree of the inverse of the permutation which is iterated. This result is also extended to noninjective balanced vectorial functions where the relevant quantity is the minimal degree of the inverse of a permutation expanding the function. This property has consequences in symmetric cryptography since several attacks or distinguishers exploit a low algebraic degree, like higher order differential attacks, cube attacks, and cube testers, or algebraic attacks. Here, we present some applications of this improved bound to a higher degree variant of the block cipher $\mathcal{KN}$, to the block cipher Rijndael-256 and to the inner permutations of the hash functions ECHO and JH.**

***Index Terms*—Algebraic degree, block ciphers, hash functions, higher order differential attacks.**

I. Introduction

MOST of the symmetric cryptographic primitives that are used nowadays, including block ciphers and hash functions, base their designs on an inner function that is iterated a high number of times. This transformation, called the round function, is very often a permutation. The algebraic degree of this permutation, i.e., the degree of the corresponding *multivariate* polynomial, is a quantity that plays an important role on the security of the symmetric primitive. Actually, a cryptographic primitive of low algebraic degree is vulnerable to many attacks, for instance higher order differential attacks [1]–[3], algebraic attacks [4], [5], or cube attacks [6].

Here, we show that, even if the inverse of the round permutation F is never used in practice, as it is the case for Feistel ciphers or for hash functions, its degree also plays a fundamental role on the degree of the composition $G \circ F$ and in consequence on the overall degree of the primitive. Even if the degree of the round function is high, if the degree of the inverse is low, the degree of the cipher will be much lower than expected. This result helps in general the understanding of the evolution of the algebraic degree of iterated permutations. Several earlier works have established new bounds on the degree of such permutations: most notably, [7] connects the degree of $G \circ F$ with the divisibility of the Walsh spectrum of F by a high power of 2 and a recent result [8] applies to the families of functions composed of several smaller balanced functions. Here, we derive some new bounds on the degree of $G \circ F$ which involve the degree of F^{-1}. In the design of some particular ciphers, the nonlinear building blocks in the round function are not permutations. This is, for example, the case for the data encryption standard (DES) that uses a collection of eight 6 × 4 balanced functions. Obviously, the notion of inverse does not exist for such functions. However, we show that the overall degree of the cipher depends on the minimal degree of the inverse of any permutation expanding the output of the function. Thus, a result, similar to the one for permutations, can be derived.

As illustrations, we apply our results to $\mathcal{KN}'$, a variant of $\mathcal{KN}$, a cipher proposed by Knudsen and Nyberg in [9]. In this variant, the quadratic round permutation which was originally used in $\mathcal{KN}$ is replaced by a function with higher degree but derived from a permutation whose inverse has algebraic degree 2. Our new bounds are also applied to the cipher Rijndael-256 and to two hash functions accepted for Round 2 in the SHA-3 competition, ECHO and JH.

The rest of this paper is organized as follows. After some preliminaries on the algebraic degree of a vectorial function, different attack techniques that exploit a low algebraic degree are recalled in Section II. Section III presents the main result on the influence of the inverse of a permutation F on the degree of $G \circ F$ and includes some corollaries. A variant of the main result for noninjective balanced functions is presented in Section IV. Finally, an attack on a variant of the block cipher $\mathcal{KN}$ [9] is illustrated in Section V, together with applications to Rijndael-256 and to some hash functions.

II. Exploiting a Low Algebraic Degree in Cryptanalysis

The whole paper focuses on functions F from $\mathbf{F}_2^n$ into $\mathbf{F}_2^m$. The *coordinates* of such a function F are the m Boolean functions F_i, $1 \leq i \leq m$, such that $F(x) = (F_1(x), \dots, F_m(x))$ for all x. The *algebraic degree* of F is defined by the algebraic degrees of its coordinates as follows.

Definition 2.1: Let f be a function from $\mathbf{F}_2^n$ into $\mathbf{F}_2$. Then, f can be uniquely written as a multivariate polynomial in $\mathbf{F}_2[x_1, \dots, x_n]/(x_1^2 - x_1), \dots, (x_n^2 - x_n)$, named its algebraic normal form

$$f(x_1, \dots, x_n) = \sum_{u=(u_1,\dots,u_n)\in\mathbf{F}_2^n} a_u \prod_{i=1}^{n} x_i^{u_i} .$$

Manuscript received September 06, 2011; revised August 13, 2012; accepted August 14, 2012. Date of publication September 14, 2012; date of current version December 19, 2012. This work was carried out when A. Canteaut was with the Department of Mathematics, Technical University of Denmark, Lyngby. This work was supported in part by the French Agence Nationale de la Recherche through the SAPHIR2 project under Contract ANR-08-VERS-014. This paper was presented in part at Finite Fields and Applications (Fq10), Ghent, Belgium, July 2011, and in part at the third SHA-3 Candidate Conference, Washington, DC, March 2012.

C. Boura is with INRIA Paris-Rocquencourt, SECRET Project-Team, 78153 Le Chesnay Cedex, France, and also with Gemalto, 92190 Meudon, France (e-mail: christina.boura@inria.fr).

A. Canteaut is with INRIA Paris-Rocquencourt, SECRET Project-Team, 78153 Le Chesnay Cedex, France (e-mail: anne.canteaut@inria.fr).

Communicated by T. Helleseth, Associate Editor for Sequences.

Digital Object Identifier 10.1109/TIT.2012.2214203

The *(algebraic) degree* of f is then defined as

$$\deg f = \max\{wt(u) : u \in \mathbf{F}_2^n, a_u \neq 0\}$$

where wt denotes the Hamming weight of a binary vector.

For a function F from $\mathbf{F}_2^n$ into $\mathbf{F}_2^m$, $m \geq 1$, the (algebraic) degree of F is the maximal algebraic degree of its coordinates.

Moreover, the coefficients a_u in the algebraic normal form of a Boolean function f can be easily computed from $2^{wt(u)}$ pairs of inputs/outputs of f as follows [10]:

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

where $x \preceq u$ means that $x_i \leq u_i$ for all i, and the sum is computed modulo 2. It is worth noticing that, when all the 2^n values of f are known, the 2^n coefficients of the algebraic normal form can be computed all together by the Moebius transform with time complexity $\mathcal{O}(n2^n)$[11, p. 286].

From the other side, every vectorial function F from $\mathbf{F}_2^n$ into $\mathbf{F}_2^n$ can also be seen as a univariate polynomial over $\mathbf{F}_{2^n}$. This representation is possible because $\mathbf{F}_{2^n}$ can be identified with an n-dimensional vector space over $\mathbf{F}_2$. Thus, for every such F, there exists a unique univariate polynomial representation over $\mathbf{F}_{2^n}$, of degree at most $2^n - 1$

$$F(x) = \sum_{i=0}^{2^n-1} b_i x^i, \qquad b_i \in \mathbf{F}_{2^n}.$$

In this case, it can be shown that the algebraic degree of F represented in such a way is given by

$$\deg F = \max\{wt(i) : 0 \leq i < 2^n \text{ and } b_i \neq 0\}$$

where $wt(i)$ denotes the Hamming weight of the n-bit vector corresponding to the binary expansion of i (see, e.g., [12, Definition 4]).

Many statistical attacks against symmetric cryptosystems exploit the fact that a family of functions $(F_k)_{k\in\mathcal{K}}$ (resp. of permutations), whose inputs and outputs can be computed from plaintext/ciphertext pairs, is not pseudorandom. Several properties may be used to distinguish F_k from a randomly chosen function, including the fact that some given coefficients in its algebraic normal form are not distributed as it is expected for a family of randomly chosen functions. The simplest attack exploiting some property of the coefficients of the algebraic normal form is the higher order differential attack introduced by Knudsen [2]: this attack uses that, for all values of k, all coordinates of F_k have degree strictly less than n, in general, and strictly less than $n-1$, in the case of a permutation. The algebraic degree of F_k is then of primary importance since the data complexity of this cryptanalysis is proportional to $2^{\deg F_k}$[13], [14]. The higher order differential attack has been generalized to other types of symmetric primitives, especially to stream ciphers, under different names (including *cube distinguishers*) in [15]–[19]. Cube attacks [6] and algebraic attacks [4], [5] also exploit low-degree relations between some components of the cryptosystem, but they mainly aim at reducing the time complexity for recovering the secret key from a low-degree distinguisher. Finally, even if both univariate and multivariate degrees are related, all these attacks must be distinguished from the attacks exploiting a low univariate degree, like the interpolation attack and its variants [20]–[22].

In the case of iterated block ciphers, i.e., ciphers consisting of several iterations of the same round permutation parametrized by different round keys, the target function F_k usually corresponds to the encryption function where the last round is omitted. Then, the fact that F_k has a low degree can be used to recover the last-round subkey either by an exhaustive search [20], or by setting up a low-degree algebraic system in these subkey bits which can be solved with time complexity depending on the algebraic degree of the round function [3], [23]. Predicting the evolution of the degree of the cipher when the number of rounds varies is then one of the main issues in higher order differential attacks.

III. On the Degree of $G \circ F$ When F is a Permutation

A. General Problem

We now focus on the following general problem: let F be a function from $\mathbf{F}_2^n$ into $\mathbf{F}_2^n$ and G be a function from $\mathbf{F}_2^n$ into $\mathbf{F}_2^m$, for some m. Then, we aim at exhibiting some particular classes of functions F such that the trivial bound

$$\deg(G \circ F) \leq \deg(F)\deg(G)$$

can be improved.

The following two families corresponding to some common situations in cryptographic applications have been previously identified in [7] and [8].

Proposition 3.1: [7] Let F be a function from $\mathbf{F}_2^n$ into $\mathbf{F}_2^n$ and G be a function from $\mathbf{F}_2^n$ into $\mathbf{F}_2^m$. Assume that all Walsh coefficients of F, i.e., all

$$\sum_{x\in\mathbf{F}_2^n} (-1)^{b\cdot F(x)+a\cdot x},\ a, b \in \mathbf{F}_2^n$$

are divisible by 2^ℓ for some integer $\ell \geq 1$; then

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

When F is a permutation, we can deduce the following corollary which involves the degree of F^{-1}.

Corollary 3.1: Let F be a permutation of $\mathbf{F}_2^n$ and let G be a function from $\mathbf{F}_2^n$ into $\mathbf{F}_2^m$. Then, we have

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

Proof: Obviously, the sets of all Walsh coefficients of a permutation and of its inverse are the same since

$$\sum_{x\in\mathbf{F}_2^n} (-1)^{b\cdot F(x)+a\cdot x} = \sum_{x\in\mathbf{F}_2^n} (-1)^{a\cdot F^{-1}(x)+b\cdot x}\ .$$

Moreover, a lower bound of the highest power of 2 which divides all Walsh coefficients of a Boolean function can be derived from Katz theorem [24]: for any function F and any nonzero $b \in \mathbf{F}_2^n$, we have

$$\sum_{x\in\mathbf{F}_2^n} (-1)^{b\cdot F(x)+a\cdot x} \equiv \sum_{x\in\mathbf{F}_2^n} (-1)^{b\cdot F(x)} \pmod{2^{\lceil \frac{n-1}{\deg F}\rceil+1}} .$$

Since F is a permutation, any nonzero linear combination of its coordinates is balanced, which means that the right-hand side of the congruence is equal to zero. Then, by applying this result both to F and F^{-1}, we obtain that all Walsh coefficients of F are divisible by 2^ℓ with

$$\ell \geq 1 + \left\lceil \frac{n-1}{\min(\deg F, \deg F^{-1})} \right\rceil .$$

■

In particular, if F^{-1} is quadratic, Corollary 3.1 leads to

$$\deg(G \circ F) \leq \left\lfloor \frac{n-1}{2} \right\rfloor + \deg G$$

which may provide some relevant information if $\deg G \leq \lceil \frac{n-1}{2} \rceil$.

It has been recently shown in [8] that the bound given by Proposition 3.1 can be improved when F corresponds to the parallel applications of smaller balanced functions, i.e., $F = (S_1, \ldots, S_s)$. This particular situation is actually very common in cryptography for obvious implementation reasons.

B. Main Result

We now show that the upper bound given by Corollary 3.1 can be improved. This improvement relies on the following theorem which bounds the maximum degree for the product of any k coordinates of F, for all $1 \leq k \leq n$. The following notation will then be extensively used.

Definition 3.1: Let F be a function from $\mathbf{F}_2^n$ into $\mathbf{F}_2^m$. For any integer k, $1 \leq k \leq m$, $\delta_k(F)$ denotes the maximal algebraic degree of the product of any k (or fewer) coordinates of F

$$\delta_k(F) = \max_{K \subset \{1,\ldots,m\}, |K| \leq k} \deg \left(\prod_{i \in K} F_i \right) .$$

In particular, $\delta_1(F) = \deg F$.

Theorem 3.1: Let F be a permutation on $\mathbf{F}_2^n$. Then, for any integers k and ℓ, $\delta_\ell(F^{-1}) < n - k$ if and only if $\delta_k(F) < n - \ell$.

Proof: We only have to show that if $\delta_\ell(F^{-1}) < n - k$ then $\delta_k(F) < n - \ell$. Indeed, the reciprocal relation is obtained by exchanging the roles of F and F^{-1}.

Let $\pi : x \mapsto \prod_{i\in K} F_i(x)$, with $|K| \leq k$. For $L \subset \{1, \ldots, n\}$, with $|L| \leq \ell$, we denote by a_L the coefficient of the monomial $\prod_{j \notin L} x_j$ of degree $n - |L|$. We will show that $a_L = 0$:

$$\begin{aligned} a_L = &\sum_{\substack{x\in\mathbf{F}_2^n \\ x_j=0, j\in L}} \pi(x) \\ = &\#\{x \in \mathbf{F}_2^n : x_j = 0, j \in L \text{ and } F_i(x) = 1, i \in K\} \\ &\bmod 2 \\ = &\#\{y \in \mathbf{F}_2^n : y_i = 1, i \in K \text{ and } F_j^{-1}(y) = 0, j \in L\} \\ &\bmod 2 \end{aligned}$$

where the last equality comes from the fact that F is a permutation, implying that there is a one-to-one correspondence between x and $y = F(x)$. Additionally, $F_j^{-1}(y) = 0$ for all $j \in L$ if and only if $\prod_{j\in L}(1 + F_j^{-1}(y)) = 1$. Then

$$\begin{aligned} a_L = &\#\{y \in \mathbf{F}_2^n : y_i = 1, i \in K \text{ and} \\ &\prod_{j\in L}(1 + F_j^{-1}(y)) = 1\} \bmod 2 . \end{aligned} \qquad (1)$$

Now, we define the Boolean function

$$\begin{array}{rcl} H_{K,L} : \{x \in \mathbf{F}_2^n : x_i = 1, i \in K\} & \rightarrow & \mathbf{F}_2 \\ x & \mapsto & \prod_{i\in L}(1 + F_i^{-1}(x)) . \end{array}$$

We have

$$a_L = wt(H_{K,L}) \bmod 2 .$$

$H_{K,L}$ is a function of $n - k$ variables and it has degree at most $\delta_\ell(F^{-1})$. Then, as by hypothesis $\delta_\ell(F^{-1}) < n - k$, $H_{K,L}$ is of even Hamming weight and thus $a_L = 0$, which means that $\delta_k(F) < n - \ell$. ■

This theorem explains for instance the observation reported in [25] on the inverse of the quadratic permutation χ over $\mathbf{F}_2^5$ used in the hash function KECCAK [26]. Since $\delta_1(\chi) = \deg \chi = 2$, we have $\delta_2(\chi^{-1}) < 4$.

The following (less precise) result can be derived from the trivial bound on $\delta_\ell(F^{-1})$.

Corollary 3.2: Let F be a permutation of $\mathbf{F}_2^n$ and let G be a function from $\mathbf{F}_2^n$ into $\mathbf{F}_2^m$. Then, we have

$$\deg(G \circ F) < n - \left\lfloor \frac{n - 1 - \deg G}{\deg(F^{-1})} \right\rfloor .$$

Proof: Obviously, $\deg(G \circ F) \leq \delta_{\deg G}(F)$. But, the previous theorem shows that $\delta_{\deg G}(F) < n - \ell$ for some integer ℓ if and only if $\delta_\ell(F^{-1}) < n - \deg G$. However, we have from the trivial bound that $\delta_\ell(F^{-1}) \leq \ell \deg(F^{-1})$. It follows that $\delta_\ell(F^{-1}) < n - \deg G$ for any integer ℓ satisfying

$$\ell \leq \left\lfloor \frac{n - 1 - \deg G}{\deg(F^{-1})} \right\rfloor .$$

Indeed, if $n - \deg G \not\equiv 0 \pmod{\deg(F^{-1})}$, we have

$$\left\lfloor \frac{n - 1 - \deg G}{\deg(F^{-1})} \right\rfloor = \left\lfloor \frac{n - \deg G}{\deg(F^{-1})} \right\rfloor .$$

Otherwise

$$\left\lfloor \frac{n - 1 - \deg G}{\deg(F^{-1})} \right\rfloor = \frac{n - \deg G}{\deg(F^{-1})} - 1 .$$

Therefore, in all cases, we have

$$\deg(F^{-1}) \left\lfloor \frac{n - 1 - \deg G}{\deg(F^{-1})} \right\rfloor < n - \deg G$$

implying that

$$\begin{aligned} \delta_\ell(F^{-1}) &\leq \ell \deg(F^{-1}) \\ &\leq \deg(F^{-1}) \left\lfloor \frac{n - 1 - \deg G}{\deg(F^{-1})} \right\rfloor < n - \deg G . \end{aligned}$$

We then deduce that

$$\delta_{\deg G}(F) < n - \left\lfloor \frac{n-1-\deg G}{\deg(F^{-1})} \right\rfloor .$$

■

Obviously, the upper bound in the previous theorem gets better when the degree of F^{-1} decreases. Moreover, if G is balanced, this bound is relevant only if it improves the trivial bound $\deg(G \circ F) < n$. It then provides some information if $\deg G \leq n - 1 - \deg F^{-1}$, while the bound in Corollary 3.1 was relevant only for $\deg G < \left\lceil \frac{n-1}{\min(\deg F, \deg F^{-1})} \right\rceil$.

C. Some Corollaries

Some simple corollaries of Theorem 3.1 can be obtained by setting $k = 1$ in the theorem. In this case, we have $\deg(F^{-1}) < n - \ell$ if and only if $\delta_\ell(F) < n-1$. We then deduce the following result.

Corollary 3.3: Let F be a permutation of $\mathbf{F}_2^n$. Then

$$\deg(F^{-1}) = n - \min\{k \ : \ \delta_k(F) \geq n-1\} .$$

In particular, $\deg(F^{-1}) = n-1$ if and only if $\deg(F) = n-1$.

It is worth noticing that almost all permutations over $\mathbf{F}_2^n$ have maximal algebraic degree $(n-1)$, since this class includes in particular all permutations with univariate degree $(2^n - 2)$, which correspond to almost all permutations [27]–[29]. For instance, any transposition is an involution with algebraic degree $(n-1)$.

We can also deduce from Corollary 3.3 that for any integer k such that

$$k \leq \left\lceil \frac{n-1}{\deg F} \right\rceil - 1$$

we have

$$\delta_k(F) \leq k \deg F < n - 1 .$$

It follows that

$$\min\{k \ : \ \delta_k(F) \geq n-1\} \geq \left\lceil \frac{n-1}{\deg F} \right\rceil$$

implying that

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

We then recover in a different way the bound on $\deg(F^{-1})$ which can be derived from Katz theorem [24] on the divisibility of the Walsh spectrum of a permutation. Actually, all Walsh coefficients of F are divisible by $\left\lceil \frac{n-1}{\deg F} \right\rceil + 1$ and it is well known that the degree of a function whose Walsh coefficients are divisible by 2^ℓ is at most $(n+1-\ell)$ (see, e.g., [7, Proposition 3]).

Corollary 3.3 also implies the following.

Corollary 3.4: Let F be a permutation of $\mathbf{F}_2^n$. Then, the product of k coordinates of F has degree $(n-1)$ if and only if $n - \deg(F^{-1}) \leq k \leq n-1$.

In particular, $\delta_{n-1}(F) = n-1$.

Proof: Corollary 3.3 implies that the smallest k such that $\delta_k(F) \geq n-1$ is equal to $n - \deg(F^{-1})$. Moreover, it is known that $\delta_k(F) = n$ if and only if $k = n$. Finally, since $n - \deg(F^{-1}) \leq n-1$, we deduce that $\delta_{n-1}(F) = n-1$, for any permutation of $\mathbf{F}_2^n$. ■

The above results can also be used for improving the bound on $\deg(G \circ F)$ found in [8] when F is the concatenation of several smaller permutations.

Theorem 3.2: Let F be a permutation from $\mathbf{F}_2^n$ into $\mathbf{F}_2^n$ corresponding to the concatenation of s smaller permutations, $S_1, \ldots, S_s$, defined over $\mathbf{F}_2^{n_0}$. Then, for any function G from $\mathbf{F}_2^n$ into $\mathbf{F}_2^m$, we have

$$\deg(G \circ F) \leq n - \frac{n - \deg(G)}{\gamma} \tag{2}$$

where

$$\gamma = \max_{1 \leq i \leq n_0 - 1} \frac{n_0 - i}{(n_0 - \max_{1 \leq j \leq s} \delta_i(S_j))} .$$

Most notably, we have

$$\gamma \leq \max_{1 \leq j \leq s} \max \left(\frac{n_0 - 1}{n_0 - \deg(S_j)}, \ \frac{n_0}{2} - 1, \ \deg(S_j^{-1}) \right) .$$

Proof: We denote by γ_i the quantity

$$\gamma_i = \frac{n_0 - i}{n_0 - \max_{1 \leq j \leq s} \delta_i(S_j)}$$

and we will try to compute the maximal γ_i for $1 \leq i \leq n_0 - 1$, i.e., γ.

For $i = 1$

$$\gamma_1 = \max_{1 \leq j \leq s} \frac{n_0 - 1}{(n_0 - \deg(S_j))}.$$

For $2 \leq i < n_0 - \max_{1 \leq j \leq s} \deg(S_j^{-1})$, we get from Corollary 3.4 that $\max_{1 \leq j \leq s} \delta_i(S_j) \leq n_0 - 2$, and thus

$$\gamma_i = \max_{1 \leq j \leq s} \frac{n_0 - i}{(n_0 - \delta_i(S_j))} \leq \frac{n_0 - i}{2} \leq \frac{n_0 - 2}{2}.$$

Finally, for the remaining indexes, i.e., for $i \geq n_0 - \max_{1 \leq j \leq s} \deg(S_j^{-1})$, we get that

$$\gamma_i = \max_{1 \leq j \leq s} \frac{n_0 - i}{(n_0 - \delta_i(S_j))} \leq n_0 - i \leq \max_{1 \leq j \leq s} \deg(S_j^{-1}).$$

■

IV. Generalization to Balanced Functions From $\mathbf{F}_2^n$ Into $\mathbf{F}_2^m$ With $m < n$

In some symmetric primitives, the functions used to provide confusion are not permutations, but balanced functions $F : \mathbf{F}_2^n \rightarrow \mathbf{F}_2^m$, with $m < n$. An example of this design is the first encryption standard cipher, DES [30], whose round function uses the parallel application of eight different 6 × 4 Sboxes, all of them of degree 5 in six variables.

An interesting problem is to be able to predict in some manner the evolution of the algebraic degree of the cipher after few rounds of encryption. Clearly, as the Sboxes of DES are not permutations, they cannot be inverted. Nevertheless, similar results as before can be deduced.

Definition 4.1: Let $F : \mathbf{F}_2^n \to \mathbf{F}_2^m$, with $m < n$, $F = (F_1, \dots, F_m)$, be a balanced function. A permutation P of $\mathbf{F}_2^n$ is called *an expansion* of F if its first m output coordinates correspond to the coordinates of F, i.e., for all i, $1 \leq i \leq m$

$$P_i(x) = F_i(x),\ \forall x \in \mathbf{F}_2^n\ .$$

In other words, F is expanded in a permutation with n outputs in the following way: as F is balanced, each of the 2^m vectors of $\mathbf{F}_2^m$ is taken by F exactly 2^{n-m} times. We then complete all of these equal vectors by concatenating to each of them a different element of $\mathbf{F}_2^{n-m}$ in order to obtain 2^{n-m} different vectors of $\mathbf{F}_2^n$. For instance, if $(n,m) = (6,4)$, then $v = (0,1,1,0)$ is a vector in the image set of F obtained for exactly four inputs, namely a, b, c, and d in $\mathbf{F}_2^6$. Then, an expansion of F can be defined by associating with a, b, c, and d the four different vectors of $\mathbf{F}_2^6$, $(0,1,1,0,0,0)$, $(0,1,1,0,0,1)$, $(0,1,1,0,1,0)$, and $(0,1,1,0,1,1)$. These four images are obtained by concatenating $v = (0,1,1,0)$ with all elements of $\mathbf{F}_2^2$. There are $(2^{n-m}!)^{2^m}$ different expansions of a given F.

Theorem 4.1: Let F be a balanced function from $\mathbf{F}_2^n$ to $\mathbf{F}_2^m$, with $m < n$. Let k and ℓ be two integers with $1 \leq k \leq m$ and $1 \leq \ell < n$. Then, the following three properties are equivalent.

i) There exists a permutation P_F of $\mathbf{F}_2^n$ expanding F such that, in any product of ℓ coordinates of P_F^{-1}, all monomials of degree greater than or equal to $(n-k)$ have degree strictly less than $(n-m)$ in the last $n-m$ variables.

ii) For any permutation P_F of $\mathbf{F}_2^n$ expanding F, we have that, in any product of ℓ coordinates of P_F^{-1}, all monomials of degree greater than or equal to $(n-k)$ have degree strictly less than $(n-m)$ in the last $n-m$ variables.

iii) $\delta_k(F) < n - \ell$.

Proof: Let $K \subset \{1,\dots,m\}$ and $L \subset \{1,\dots,n\}$. Let π_K denote the product of the coordinates F_i for $i \in K$. Then, the coefficient $a_{K,L}$ of the monomial $\prod_{i\in\{1,\dots,n\}\setminus L} x_i$ in the algebraic normal form of F is given by

$$\begin{aligned} a_{K,L} &= \sum_{\substack{x\in\mathbf{F}_2^n \\ x_j=0, j\in L}} \pi_K(x) \\ &= \#\{x \in \mathbf{F}_2^n : x_j = 0, j \in L \text{ and} \\ &\qquad F_i(x) = 1, i \in K\} \mod 2 \\ &= \#\{x \in \mathbf{F}_2^n : x_j = 0, j \in L \text{ and} \\ &\qquad (P_F)_i(x) = 1, i \in K\} \mod 2 \end{aligned}$$

where the last equality holds for any expansion P_F of F. Then, as P_F is a permutation, setting $y = P_F(x)$ leads to

$$\begin{aligned} a_{K,L} = \#\{&y \in \mathbf{F}_2^n : y_i = 1,\ i \in K \text{ and} \\ &(P_F^{-1})_j(y) = 0,\ j \in L\} \mod 2 \end{aligned}$$

implying that $a_{K,L} = 0$ if and only if the Boolean function

$$\begin{array}{rcl} H_{K,L} : \{x \in \mathbf{F}_2^n : x_i = 1,\ i \in K\} & \to & \mathbf{F}_2 \\ x & \mapsto & \prod_{i\in L}(1 + (P_F^{-1})_i(x)) \end{array}$$

has degree strictly less than $(n-k)$.

Let us first prove that (i) implies (iii). We deduce from the previous reasoning that, if Condition (i) holds, any monomial of degree greater than or equal to $(n-k)$ in the ANF of the n-variable Boolean function

$$x \mapsto \prod_{i\in L}(1 + (P_F^{-1})_i(x))$$

is not a factor of $x_{m+1}\dots x_n$. Therefore, the restriction of such a monomial to any set $\{x \in \mathbf{F}_2^n : x_i = 1, i \in K\}$ with $K \subset \{1,\dots,m\}$ has degree strictly less than $(m-k) + (n-m) = (n-k)$. It follows that, for any choice of $K \subset \{1,\dots,m\}$, $H_{K,L}$ has degree strictly less than $(n-k)$. Then, we have: (ii) $\Rightarrow$ (i) $\Rightarrow$ (iii).

Conversely, we can prove that (iii) implies (ii). Suppose that (ii) does not hold, i.e., there exists some permutation P_F expanding F and some set $L \subset \{1,\dots,m\}$ such that the n-variable Boolean function

$$\pi'_L : x \mapsto \prod_{i\in L}(P_F^{-1})_i(x)$$

contains a monomial of the form $x_{m+1}\dots x_n \prod_{i\in I} x_i$ for some set $I \subset \{1,\dots,m\}$ of size at least $(m-k)$. We can suppose that L is the smallest such set for inclusion (otherwise, we choose the smallest $L' \subset L$ satisfying the property). Let us choose $K = \{1,\dots,m\} \setminus I$ where $x_{m+1}\dots x_n \prod_{i\in I} x_i$ is the monomial with the highest degree of this form in the ANF of π'_L. By hypothesis, the size of K is at most k, and it is greater than or equal to 1 since π'_L cannot have degree n when $|L| < n$ [8, Prop 1]. Since L is minimal for inclusion and

$$H_{K,L}(x) = \sum_{L'\subseteq L}\prod_{i\in L'}(P_F^{-1})_i(x)$$

it is clear that $H_{K,L}$ has degree $(n-k)$ if and only if the restriction of π'_L to the set $\{x \in \mathbf{F}_2^n : x_i = 1, i \in K\}$ has degree $(n-k)$. However, the algebraic normal form of π'_L contains the monomial $x_{m+1}\dots x_n \prod_{i\notin K} x_i$, implying that $H_{K,L}$ has degree at least $(n-k)$. It follows that, for these particular choices of L and K, $a_{K,L} = 1$ implying that there exists some product of k or fewer coordinates of F which has degree greater than or equal to $(n-\ell)$. Finally, it follows that all three properties are equivalent. ■

A corollary similar to Corollary 3.2 can be deduced now for the case of noninjective balanced functions.

Corollary 4.1: Let F be a balanced function from $\mathbf{F}_2^n$ into $\mathbf{F}_2^m$ and G a function from $\mathbf{F}_2^m$ into $\mathbf{F}_2^k$. For any permutation F^* expanding F, we have

$$\deg(G \circ F) < n - \left\lfloor \frac{n - 1 - \deg G}{\deg(F^{*-1})} \right\rfloor.$$

Proof: Let F^* be a permutation expanding F. We have shown in the proof of Corollary 3.2 that the trivial bound implies that $\delta_\ell(F^{*-1}) < n - \deg G$ for any

$$\ell \leq \left\lfloor \frac{n - 1 - \deg G}{\deg(F^{*-1})} \right\rfloor\ .$$

It follows that, when ℓ satisfies this condition, the product of any ℓ coordinates of F^{*-1} does not contain any monomial of degree

$(n - \deg G)$. Since Condition (i) in Theorem 4.1 is satisfied, we deduce that

$$\deg(G \circ F) \leq \delta_{\deg G}(F) < n - \left\lfloor \frac{n-1-\deg G}{\deg(F^{*-1})} \right\rfloor .$$

■

It is known that the product of k coordinates of a balanced function F with n input variables has degree n if and only if $k = n$ (see, e.g., [8, Prop 1]). Moreover, when F is a permutation, we have shown in Corollary 3.4 that the degree of F^{-1} determines whenever the product of some coordinates of F has degree $(n-1)$. Here, we provide a similar result in the case where F is a noninjective balanced function.

Corollary 4.2: Let F be a balanced function from $\mathbf{F}_2^n$ to $\mathbf{F}_2^m$, with $m < n$. Then, $\delta_m(F) \leq n-2$ if and only if, for any $y \in \mathbf{F}_2^m$, the 2^{n-m} preimages of y by F sum to zero, i.e.,

$$\sum_{x:F(x)=y} x = 0$$

where the sum corresponds to the addition in $\mathbf{F}_2^n$.

Proof: From Theorem 4.1 applied with $k = m$ and $\ell = 1$, we know that $\delta_m(F) \leq n-2$ if and only if there exists some permutation P_F expanding F such that any monomial with degree at least $(n-m)$ in the ANF of any coordinate of P_F^{-1} is not a factor of $x_{m+1}\ldots x_n$. Since a monomial of degree less than $(n-m)$ cannot be a factor of $x_{m+1}\ldots x_n$, this equivalently means that any monomial in the ANF of any coordinate of P_F^{-1} is not a factor of $x_{m+1}\ldots x_n$. Let

$$\begin{array}{rccc} f: & \mathbf{F}_2^m \times \mathbf{F}_2^{n-m} & \to & \mathbf{F}_2 \\ & (x,y) & \mapsto & [P_F^{-1}(x,y)]_i \end{array}$$

for some i. For any $(u,v) \in \mathbf{F}_2^m \times \mathbf{F}_2^{n-m}$, $a_{u,v}$ denotes the coefficient in the ANF of f of the monomial $\prod_{i,u_i \neq 0} x_i \prod_{i,v_i \neq 0} x_{m+1+i}$. Let 1_{n-m} denote the all-one vector in $\mathbf{F}_2^{n-m}$. For any $x \in \mathbf{F}_2^m$ and $y \in \mathbf{F}_2^{n-m}$, we have

$$f(x,y) = \sum_{v \preceq y} \left[\sum_{u \preceq x} a_{u,v} \right]$$

where $x \preceq y$ means that $x_i \leq y_i$ for all i. Then

$$\begin{aligned} \sum_{y \in \mathbf{F}_2^{n-m}} f(x,y) &= \sum_{y \in \mathbf{F}_2^{n-m}} \sum_{v \preceq y} \left[\sum_{u \preceq x} a_{u,v} \right] \\ &\equiv \sum_{v \in \mathbf{F}_2^{n-m}} N_v \left[\sum_{u \preceq x} a_{u,v} \right] \pmod 2 \end{aligned}$$

where

$$N_v = \#\{y \in \mathbf{F}_2^{n-m} : v \preceq y\} \bmod 2 = 2^{n-m-wt(v)} \bmod 2 .$$

Then, $N_v = 0$ except when v is the all-one vector. Therefore

$$\sum_{y \in \mathbf{F}_2^{n-m}} f(x,y) = \sum_{u \preceq x} a_{u,1_{n-m}} .$$

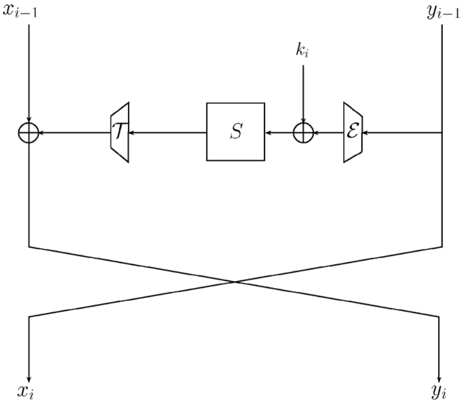

Fig. 1. Round i of the $\mathcal{KN}$-cipher.

We then deduce that all $a_{u,1_{n-m}} = 0$ for $u \in \mathbf{F}_2^m$ if and only if

$$\sum_{y \in \mathbf{F}_2^{n-m}} f(x,y) = 0$$

for all $x \in \mathbf{F}_2^m$. It is worth noticing that this property is similar to the property used in cube attacks [6, Th. 1].

Since this property holds for any coordinate f of P_F^{-1}, the required condition equivalently means that, for any $x \in \mathbf{F}_2^m$

$$\sum_{y \in \mathbf{F}_2^{n-m}} P_F^{-1}(x,y) = 0$$

where the sum is an addition in $\mathbf{F}_2^n$. By definition of P_F, all elements $P_F^{-1}(x,y)$ when $y \in \mathbf{F}_2^{n-m}$ correspond to the preimages of x under F. The condition can then be written as

$$\sum_{z:F(z)=x} z = 0 .$$

■

V. Applications to Some Symmetric Primitives

In this section, we will show how the previous results can be used in order to predict the evolution of the algebraic degrees of some chosen permutations that are the main building blocks of some well-known block ciphers and hash functions.

A. Attacking the $\mathcal{KN}$-Cipher and Its Variant

One of the first examples in the literature of a concrete attack exploiting the low algebraic degree of a symmetric primitive is the higher order differential attack presented by Jakobsen and Knudsen [20] against the $\mathcal{KN}$-cipher. This construction, a.k.a CRADIC [31], has been proposed by Nyberg and Knudsen in [9]. It is a six-round Feistel cipher over $\mathbf{F}_2^{64}$ with a 198-bit secret key. Its round permutation is defined as follows (see Fig. 1):

$$\begin{array}{ccc} \mathbf{F}_2^{32} \times \mathbf{F}_2^{32} & \to & \mathbf{F}_2^{32} \times \mathbf{F}_2^{32} \\ (x,y) & \mapsto & (y, x + \mathcal{T} \circ S\left(\mathcal{E}(x) + k_i\right)) \end{array}$$

where k_i is the ith round subkey, $\mathcal{E}$ is a linear expansion from $\mathbf{F}_2^{32}$ into $\mathbf{F}_2^{33}$, $\mathcal{T}$ is a linear truncation from $\mathbf{F}_2^{33}$ into $\mathbf{F}_2^{32}$, and S

is the power function x^3 over $\mathbf{F}_{2^{33}}$. In this definition, the finite field $\mathbf{F}_{2^{33}}$ is identified with the vector space $\mathbf{F}_2^{33}$.

The main motivation behind this design is that the choice of S, which is the only nonlinear part in the cipher, guarantees an optimal resistance to both linear and differential attacks. Thus, x^3 over $\mathbf{F}_{2^n}$, n odd, was chosen, since it is an almost bent function [32]. More precisely, some lower bounds on the probabilities of the best differential and of the best linear approximation show that six rounds of this cipher are resistant to these attacks.

However, one of the main weaknesses of this cipher, identified by Jakobsen and Knudsen [20], is that the encryption function has a low algebraic degree. Indeed, for any r-round Feistel cipher, it can be observed that, when the right half of the input y_0 is a constant, the function which associates the left part of the output x_r with the left part of the input x_0 has degree at most $(\deg S)^{r-2}$. Therefore, since the Sbox in the $\mathcal{KN}$-cipher is quadratic, there exists a distinguisher for r rounds with data and time complexity $2^{2^{r-2}+1}$. This must be compared to the best known generic attacks against any four-round and five-round Feistel ciphers with 64-bit blocks, which have respective data complexities 2^{16} and 2^{32}[33]. Here, the whole encryption function can be distinguished from a random permutation with data complexity 2^{17}. Also, the 33-bit last round key k_6 can be recovered with average time complexity 2^{14} and data complexity 2^9 pairs of chosen plaintexts-ciphertexts [23]. Therefore, it is now well known that, in an r-round Feistel cipher, the Sbox must be chosen such that $(\deg S)^{r-2}$ is much higher than half of the block size. But, there is no condition on the degree of the inverse of S since S^{-1} is involved neither in the encryption function nor in the decryption function. The degree of S^{-1} may only affect the complexity of some algebraic attacks [4]. Therefore, a variant of this cipher, that we name $\mathcal{KN}'$, suggested by Nyberg and Knudsen in the same paper [9] does not present the same weakness. This variant is obtained by modifying S and using instead the inverse of a quadratic permutation. Actually, it is known that any permutation and its inverse present the same resistance to differential and linear cryptanalysis [34]. But, a major difference is that S and S^{-1} may have different algebraic degrees. For instance, if S is a quadratic power permutation over $\mathbf{F}_{2^n}$, n odd, i.e., $S(x) = x^{2^s+1}$ with $\gcd(s,n) = 1$, then the algebraic degree of S^{-1} is equal to $\frac{n+1}{2}$[32]. Since the implementation complexity of the inverse of x^3 over $\mathbf{F}_{2^{33}}$ is unacceptable in most applications, we consider the nonlinear function over $\mathbf{F}_2^{32}$ composed of four parallel applications of the same function $\widetilde{\sigma}$ defined over $\mathbf{F}_2^8$ like in $\mathcal{KN}$

$$\begin{array}{rccc} \widetilde{\sigma}: & \mathbf{F}_2^8 & \to & \mathbf{F}_2^8 \\ & x & \mapsto & t \circ \sigma(e(x))) \end{array}$$

where e is an affine expansion from $\mathbf{F}_2^8$ into $\mathbf{F}_2^9$ with maximal rank, t is a truncation from $\mathbf{F}_2^9$ into $\mathbf{F}_2^8$, and σ is the inverse of a quadratic power permutation $x \mapsto x^{2^s+1}$ over $\mathbf{F}_{2^9}$, e.g., $\sigma(x) = x^{171}$ which is the inverse of x^3. This function, which is the only nonlinear part of the cipher, has algebraic degree 5. It is worth noticing that it has a high univariate degree which prevents interpolation attacks. The round function of $\mathcal{KN}'$ is depicted in Fig. 2. It is defined by

$$\begin{array}{ccc} \mathbf{F}_2^{32} \times \mathbf{F}_2^{32} & \to & \mathbf{F}_2^{32} \times \mathbf{F}_2^{32} \\ (x,y) & \mapsto & (y, x + \mathcal{L}' \circ \widetilde{S}\left(\mathcal{L}(x) + k_i\right)) \end{array}$$

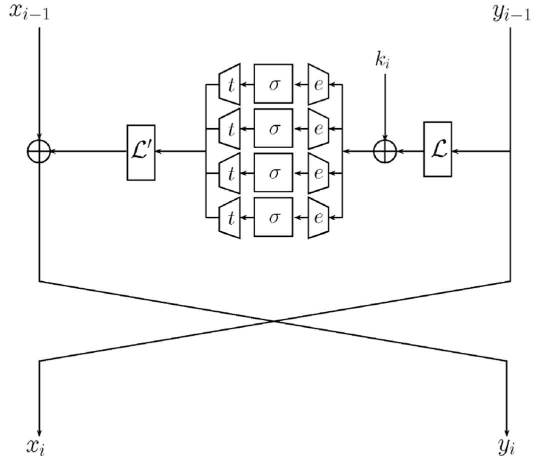

Fig. 2. Round i of the $\mathcal{KN}'$-cipher.

where $\widetilde{S}$ corresponds to four parallel applications of $\widetilde{\sigma}$, k_i is the ith 32-bit subkey, and $\mathcal{L}$ and $\mathcal{L}'$ are two linear bijections over $\mathbf{F}_2^{32}$ which aim at providing diffusion.

While the trivial bound does not provide any relevant information on the degree of the left part of the output for five rounds or more, Theorem 3.1 shows that $\mathcal{KN}'$-cipher can also be broken by the attack proposed by Jakobsen and Knudsen. At this aim, we study the algebraic degree of the function which maps x_0, the left half of the plaintext, to x_r which is the left half of the output of the cipher after r rounds. In the following, we denote by F_k the function over $\mathbf{F}_2^{32}$ defined by

$$F_k(x) = \mathcal{L}' \circ \widetilde{S}\left(\mathcal{L}(x) + k\right) \ .$$

Then, we have

$$\begin{aligned} x_2 &= x_0 + F_{k_1}(y_0) \\ x_3 &= y_0 + F_{k_2}\left(x_0 + F_{k_1}(y_0)\right) \\ x_4 &= x_0 + F_{k_1}(y_0) + F_{k_3}\left(y_0 + F_{k_2}\left(x_0 + F_{k_1}(y_0)\right)\right) \ . \end{aligned}$$

Let us now denote by x the element of $\mathbf{F}_2^{36}$ defined by

$$x = \mathcal{E}\left(\mathcal{L}(x_0 + F_{k_1}(y_0)) + k_2\right)$$

where $\mathcal{E}$ is the linear expansion from $\mathbf{F}_2^{32}$ into $\mathbf{F}_2^{36}$ composed of four applications of the smaller expansion e. Then, x_0 can be computed from x by

$$x_0 = \mathcal{L}^{-1}\left(\mathcal{E}^\star(x) + k_2\right) + F_{k_1}(y_0)$$

where $\mathcal{E}^\star$ is the function from $\mathbf{F}_2^{36}$ into $\mathbf{F}_2^{32}$ defined by $\mathcal{E}^\star\left(\mathcal{E}(x)\right) = x$ and $\mathcal{E}^\star(x) = 0$ if $x \notin Im\mathcal{E}$. Such a function exists since $\mathcal{E}$ has maximum rank. Then, x_4 can be written as a function of x

$$x_4 = \mathcal{L}^{-1}\left(\mathcal{E}^\star(x) + k_2\right) + F_{k_3}\left(y_0 + \mathcal{L}' \circ \mathcal{T} \circ S(x)\right)$$

where S is the permutation of $\mathbf{F}_2^{36}$ corresponding to four parallel applications of σ, and $\mathcal{T}$ is the function from $\mathbf{F}_2^{36}$ into $\mathbf{F}_2^{32}$ defined by four applications of the truncation t. Now, since

$$x_5 = x_3 + F_{k_4}(x_4)$$

we deduce that

$$x_5 + x_3 = F_{k_4}\left[\mathcal{L}^{-1}(\mathcal{E}^\star(x) + k_2) + F_{k_3}(y_0 + \mathcal{L}' \circ \mathcal{T} \circ S(x))\right] .$$

The degree of x_5 as a function of x_0 is at most the maximum between the degree of x_3, which is at most 5, and the degree of $x_5 + x_3$, seen as a function of x. We then focus on this last quantity. We write

$$x_5 + x_3 = G \circ S(x)$$

with

$$G(y) = F_{k_4}\left[\mathcal{L}^{-1}\big(\mathcal{E}^{\star}(S^{-1}(y)) + k_2\big) + F_{k_3}(y_0 + \mathcal{L}' \circ \mathcal{T}(y))\right].$$

Degree of G: Since F_{k_4} has degree 5, G can be decomposed as a sum of terms, each consisting of the product of i coordinates of S^{-1} multiplied by the product of at most $(5-i)$ coordinates of S. Since S^{-1} has degree 2, we get that

$$\deg G \leq \max_{0 \leq i \leq 5} \left(2i + \delta_{5-i}(S)\right).$$

From Theorem 3.1, it is known that $\delta_5(S) < 36 - \lfloor \frac{30}{2} \rfloor$, implying that $\delta_5(S) \leq 20$. Therefore, we deduce that $\deg G \leq 2 + \delta_4(S) \leq 22$.

Degree of $G \circ S$: We now apply Corollary 3.2 for upper bounding the degree of $G \circ S$, exploiting the fact that S^{-1} has degree 2. Then, we get

$$\deg(G \circ S) < 36 - \left\lfloor \frac{35 - 22}{2} \right\rfloor$$

or equivalently

$$\deg(G \circ S) \leq 29$$

and we finally find that x_5 is a function of degree at most 29 of x_0. This leads to a distinguisher on five rounds of $\mathcal{KN}'$ with data complexity 2^{30} that improves the generic distinguisher. It is worth noticing that the same upper bound can be derived from Theorem 3.2 which additionally exploits the fact that S corresponds to the concatenation of four permutations σ defined over $\mathbf{F}_2^9$.

Variant with nonbijective Sboxes: The nonlinear function in $\mathcal{KN}'$ can also be seen as the concatenation of four balanced Sboxes σ' from $\mathbf{F}_2^9$ into $\mathbf{F}_2^8$. Instead of applying Corollary 3.2 based on the degree of the inverse of the nonlinear function S, we can then rely on the existence of a permutation S^* expanding the 36 × 32 Sbox, with $\deg((S^*)^{-1}) = 2$. Then, Corollary 4.2 applies and also shows that x_5 is a function of degree at most 29 of x_0.

It is worth noticing that, if we consider another variant of the $\mathcal{KN}$-cipher using the inverse of x^3 over $\mathbf{F}_{2^{33}}$ as an Sbox, Corollary 3.2 leads to a distinguisher on four rounds exploiting that x_4 has degree at most 25. But, finding a relevant bound on the degree of x_5 remains an open problem.

B. On the Algebraic Degree of Rijndael-256

Rijndael-128 [35] is the algorithm selected by the National Institute of Standards and Technology (NIST) in 2000 as the winner of the advanced encryption standard (AES) competition in order to replace the DES. Rijndael-N_b, with $N_b \in \{128, 160, 192, 224, 256\}$ has the form of a substitution-permutation-network. The key size N_k varies between 128, 192, and 256 bits. Its round transformation applies to an N_b-bit state, that is represented by a $4 \times t$-byte matrix $A = (a_{i,j})$, with $t = N_b/32$. For instance, the states for Rijndael-128 and Rijndael-256 are depicted in Fig. 3.

$a_{0,0}$	$a_{0,1}$	$a_{0,2}$	$a_{0,3}$	$a_{0,4}$	$a_{0,5}$	$a_{0,6}$	$a_{0,7}$
$a_{1,0}$	$a_{1,1}$	$a_{1,2}$	$a_{1,3}$	$a_{1,4}$	$a_{1,5}$	$a_{1,6}$	$a_{1,7}$
$a_{2,0}$	$a_{2,1}$	$a_{2,2}$	$a_{2,3}$	$a_{2,4}$	$a_{2,5}$	$a_{2,6}$	$a_{2,7}$
$a_{3,0}$	$a_{3,1}$	$a_{3,2}$	$a_{3,3}$	$a_{3,4}$	$a_{3,5}$	$a_{3,6}$	$a_{3,7}$

$a_{0,0}$	$a_{0,1}$	$a_{0,2}$	$a_{0,3}$
$a_{1,0}$	$a_{1,1}$	$a_{1,2}$	$a_{1,3}$
$a_{2,0}$	$a_{2,1}$	$a_{2,2}$	$a_{2,3}$
$a_{3,0}$	$a_{3,1}$	$a_{3,2}$	$a_{3,3}$

Fig. 3. States of Rijndael-256 and Rijndael-128.

TABLE I
NUMBER OF ROUNDS FOR THE RIJNDAEL BLOCK CIPHER

		N_b				
		128	160	192	224	256
N_k	128	10	11	12	13	14
	192	12	12	12	13	14
	256	14	14	14	14	14

Four basic layers are composing a round of the Rijndael-N_b transformation.

1) `SubBytes`: The only nonlinear transformation of the cipher. Every byte is updated by an 8 × 8 Sbox of degree 7. The inverse transformation has the same degree.
2) `ShiftRows`: Linear transformation that rotates to the left the bytes in each row by a certain offset. This offset depends on the block size N_b. The offset is for example $\{0,1,2,3\}$ for Rijndael-128 and $\{0,1,3,4\}$ for Rijndael-256.
3) `MixColumns`: Linear transformation that applies in parallel to every column of the state.
4) `AddRoundKey`: The combination of the state with the round subkey using bitwise XOR.

A round R of the transformation applied to a state $\mathcal{S}$ corresponds thus to

$$\texttt{AddRoundKey} \circ \texttt{MixColumns} \circ \texttt{ShiftRows} \circ \texttt{SubBytes}(\mathcal{S}).$$

The number of rounds depends on the block size and on the key size. These values can be found in Table I.

As seen from the description, the only source of nonlinearity for Rijndael-N_b is the `SubBytes` transformation. This transformation has algebraic degree 7. By using the trivial bound as an estimation for the degree, we can see that the degree after two rounds is at most $7^2 = 49$ and after three rounds it is bounded by $\max(N_b - 1, 7^3)$. Thus, it may be believed that only three rounds of encryption are enough for achieving the maximal degree.

We will show, using the results of Section III, that the above estimates of the required number of rounds are way too small. We will see in particular that for Rijndael-256, at least seven rounds are needed to achieve the maximal degree.

We start by giving a bound for the degree of two rounds of Rijndael-256. By using the superSbox view [36], we can see these two rounds as the parallel application of eight copies of a function S_{32} operating on 32-bit words, followed by a linear transformation. S_{32} corresponds to the so-called SDS transformation: it consists of two layers of four 8 × 8 balanced Sboxes

TABLE II
UPPER BOUNDS FOR r ITERATIONS OF THE RIJNDAEL-256 ROUND PERMUTATION OBTAINED BY USING THE TRIVIAL BOUND, THE TRIVIAL BOUND TOGETHER WITH THE SUPERSBOX VIEW (OR THE BOUND FROM [7]) AND THE NEW RESULTS, RESPECTIVELY

# rounds	trivial bound	superSbox and [7]	this paper
1	7	7	7
2	49	31	28
3	–	127	113
4	–	–	235
5	–	–	250
6	–	–	254

of degree 7, separated by a linear layer. Therefore, we can use [8, Th. 2] and get that

$$\deg R^2 = \deg S_{32} \leq 32 - \frac{32-7}{7} < 29 \ .$$

As the state of Rijndael-256 is wide, after two rounds of the permutation, not all the parts of the state have been mixed together. Thus, we can apply a similar approach as before and see three rounds of the permutation as the parallel application of two copies of a function S_{128}, operating now on 128-bit words, followed again by a linear layer. Theorem 2 of [8] gives now

$$\deg R^3 = \deg S_{128} \leq 128 - \frac{128-28}{7} < 114.$$

Let $F = R^2$. F is a permutation of degree at most 28 and its inverse has degree at most 28 too. Using that $(R^3) \circ F = R^5$, we get a bound for the degree of Rijndael-256 after five rounds. From Theorem 3.2, we get that the constant γ associated with $F = R^2$ is at most 28 and we deduce finally that

$$\deg R^5 \leq 256 - \frac{256-113}{28} < 251.$$

We get a similar result for six rounds, by considering $F = R^3$, which has degree at most 113. Since $\deg F^{-1} \leq 113$, the corresponding constant γ is at most 113, leading to

$$\deg R^6 \leq 256 - \frac{256-113}{113} < 255.$$

Therefore, at least seven rounds are needed to achieve the maximal degree 255.

In order to make a comparison with previously known bounds and to see at what extension they are improved, we present in Table II upper bounds for Rijndael-256 coming from four different sources. The first column presents the results obtained by using the trivial bound, and the second column combines it with the superSbox view. Since the Walsh spectrum of the AES Sbox is divisible by 4 only, the bound from [7] provides the same results as the superSbox view. The last column illustrates the new results. A $-$ in the table means that the obtained bound corresponds to the maximal degree of a permutation, and thus provides no information.

C. *Application to the ECHO Hash Function*

The ECHO [37] hash function has been designed by Benadjila *et al.* for the NIST SHA-3 competition. It uses the HAIFA mode of operation. Its compression function has a 2048-bit input (corresponding to the chaining value and a message block whose respective lengths depend on the size of the message digest), and it outputs a 512-bit or a 1024-bit value. It relies on a 2048-bit AES-based permutation P.

The permutation P updates a 2048-bit state, which can be seen as a 4×4 AES state, composed of 128-bit words. In every round $\mathtt{R}$, three operations modify the state. These are the `BIG.SubWords`, `BIG.ShiftRows`, and `BIG.MixColumns` transformations. These transformations can be seen as generalizations of the three classical AES transformations. In particular,

1) `BIG.SubWords` is a nonlinear transformation applied independently to every 128-bit cell. It consists of two AES rounds.
2) The `BIG.ShiftRows` and `BIG.MixColumns` transformations are exact analogues of the AES `ShiftRows` and `MixColumns` transformations, respectively, with the only difference that they do not operate on bytes but on 128-bit words.

The number of rounds r is specified to be 8 for the 256-bit candidate. Finally, each bit in the output of the compression function is defined as a linear combination of some output bits of P and some input bits.

We will see how the algebraic degree of the permutation P varies with the number of rounds. We will show that the degree does not increase as predicted and reaches its maximum value much later than expected. The algebraic degree of the permutation P was believed to be high, as in every round $\mathtt{R}$ the input has to pass twice through the Sbox layer, of degree 7. As $7^4 = 2401$, two rounds seemed to be enough to achieve the highest possible degree.

`BIG.SubWords` is the only source of nonlinearity in the round permutation. It is a 128-bit transformation corresponding to two rounds of AES. Its degree thus matches the degree of the S_{32} transformation of Rijndael-256 and is hence at most 28. The two-round permutation $\mathtt{R}^2$ is a permutation of the set of 2048-bit states, but it can be decomposed as four parallel applications of a permutation S_{512} operating on 512-bit words, followed by a linear layer. We will determine the degree of any of these four applications. After the first round of the permutation P, every bit of the state consists of polynomials of degree at most 28. By applying to this state, the first layer of Sboxes in every `BIG.SubWords`, the degree gets at most $7 \cdot 28 = 196$. We can apply now the bound of [8, Th. 2] to get the following bound on the degree of $\mathtt{R}^2$:

$$\deg \mathtt{R}^2 = \deg S_{512} \leq 512 - \frac{512-196}{7} < 467 \ .$$

Let $F = \mathtt{R}^2$. F is then a permutation of degree at most 466. From Theorem 3.2, the constant γ associated with this permutation is at most 466, as the degrees of $\mathtt{R}^2$ and of its inverse are both upper bounded by 466, therefore

$$\deg F^2 = \deg \mathtt{R}^4 \leq 2048 - \frac{2048-466}{466} < 2045.$$

As for Rijndael, these results compared with the previously known bounds are summarized in Table III.

The same bounds hold for the inverse round transformation. Due to this observation, we are able to distinguish the inner per-

TABLE III
UPPER BOUNDS FOR r ITERATIONS OF THE ROUND PERMUTATION OF ECHO OBTAINED BY USING THE TRIVIAL BOUND, THE TRIVIAL BOUND TOGETHER WITH THE SUPERSBOX VIEW (OR THE BOUND FROM [7]) AND THE NEW RESULTS, RESPECTIVELY

# rounds	trivial bound	superSbox and [7]	this paper
1	49	31	28
2	–	511	466
3	–	–	1991
4	–	–	2044

mutation in ECHO from a random one. This can be done for instance by constructing many zero-sum partitions of size 2^{2045}, i.e., partitions of the input set $\mathbf{F}_2^{2048}$ into eight sets $X_1, \dots, X_8$ of size 2^{2045} such that all elements in each X_i sum to zero and the corresponding images $P(x), x \in X_i$ sum to zero too [38], [39]. Such a partition can be constructed by the method introduced in [39] and detailed in [38, Proposition 2]. Let V be any subspace of $\mathbf{F}_2^{2048}$ with codimension 3 and W be its complement. Then, the eight sets

$$X_i = \{(\mathtt{R}^4)^{-1}(a_i + v),\ v \in V\},\ a_i \in W$$

form a zero-sum partition of $\mathbf{F}_2^{2048}$ for P of size 2^{2045}.

D. Application to the JH Hash Function

JH [40] is a hash function family, having some members submitted to the NIST hash function competition. It has been chosen in late 2010 to be one of the five finalists of the contest.

The compression function in JH is constructed from a block cipher with constant key. This compression function is based on an inner permutation, named E_d and is composed of 42 steps of a round function R_d, where $d = 8$ for the SHA-3 candidate. R_d applies to a state of 2^{d+2} bits, divided into 4-bit words. It consists of three different layers: an Sbox layer, a linear layer, and a permutation layer P_d.

1) The *Sbox layer* corresponds to the parallel application of 2^d Sboxes to the state. Two different Sboxes, S_0 and S_1, are used in JH. Both of them, as also their inverses, are of degree 3. The selection of the Sbox to use is made by the bits of the round constant, which are not xored to the state as done in other constructions.
2) The *linear layer* mixes the 2^d words two by two.
3) The *permutation P_d* permutes the words of the state.

Two rounds of R_d, for $d = 4$, can be seen in Fig. 4.

A round of the permutation is of algebraic degree 3, as the only source of nonlinearity of the cipher comes from the 4-bit Sboxes. Thus, if we try to estimate the evolution of the degree by using the trivial bound, we can see that the degree of the permutation after six rounds is at most $\deg(R_8^6) \leq 3^6 = 729$ and consequently the maximal degree seems to be reached just seven rounds of encryption only. We will show by applying the results of Section III that the algebraic degree of JH does not increase as expected.

An important observation on the structure of the R_8 permutation is that for $r \leq 8$, r rounds of R^8, denoted by R_8^r, can be seen as the concatenation of 2^{9-r} permutations S_r over $\mathbf{F}_2^{2^{r+1}}$.

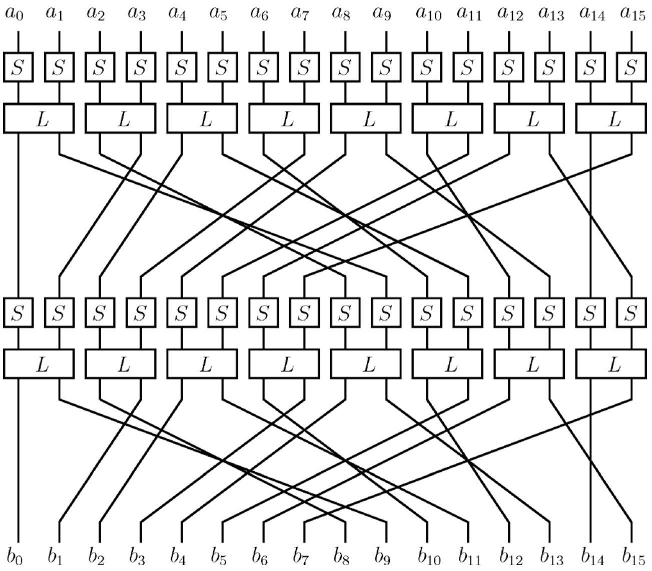
a_0 a_1 a_2 a_3 a_4 a_5 a_6 a_7 a_8 a_9 a_{10} a_{11} a_{12} a_{13} a_{14} a_{15}
S
L
S
L
b_0 b_1 b_2 b_3 b_4 b_5 b_6 b_7 b_8 b_9 b_{10} b_{11} b_{12} b_{13} b_{14} b_{15}

Fig. 4. Two rounds of R_4.

TABLE IV
UPPER BOUNDS ON THE DEGREE OF UP TO EIGHT ROUNDS OF THE JH PERMUTATION

# Rounds	Bound on $\deg(R_8^r)$
1	3
2	6
3	12
4	25
5	51
6	102
7	204
8	409

Thus, for $2 \leq r \leq 8$, a bound on the degree of R_8^r can be obtained with [8, Th. 2]

$$\deg(R_8^r) \leq 2^{r+1} - \frac{2^{r+1} - \deg(R_8^{r-1})}{3}.$$

The bounds on the degree up to eight rounds of the permutation, given by the above formula can be seen in Table IV. The same bounds hold for the inverse permutation.

Using now Theorem 3.2, we get that the constant $\gamma(S_8)$ of the permutation S_8 over $\mathbf{F}_2^{512}$ is at most 409. Thus, we have that

$$\deg R_8^{16} \leq 1024 - \frac{1024 - \deg(R_8^8)}{\gamma(S_8)} < 1023.$$

The same technique applied to 9 to 16 rounds leads to the results presented in Table V, with a comparison with the previous best results.

VI. CONCLUSION

Our study points out that, in many situations, the algebraic degree of an iterated function does not grow as fast as expected with the number of rounds. In particular, the degree of the inverse of the iterated permutation or, in the case of a noninjective function, the minimal degree of the inverse of a permutation expanding the function, has some influence on the degree

TABLE V
UPPER BOUNDS FOR r ITERATIONS OF THE ROUND PERMUTATION OF JH, OBTAINED BY USING THE TRIVIAL BOUND, THE SUPERSBOX VIEW (OR THE BOUND FROM [7]) AND THE NEW RESULTS, RESPECTIVELY

# rounds	trivial bound	superSbox and [7]	this paper
1	3	3	3
2	9	7	6
3	27	15	12
4	81	31	25
5	243	63	51
6	729	127	102
7	–	255	204
8	–	511	409
9	–	–	819
10	–	–	921
11	–	–	972
12	–	–	999
13	–	–	1011
14	–	–	1017
15	–	–	1020
16	–	–	1022

of the iterated function. This observation can be used for exhibiting nonideal behaviors in some cryptographic primitives, like block ciphers or hash functions. However, turning such distinguishers into real attacks, like a key-recovery attack on a cipher or a (second)-preimage attack on a hash function, is a difficult problem. The most promising approach consists in combining some properties of the algebraic normal form of an inner function (e.g., its low degree) and the solving of some algebraic system, as proposed in [3] and [41]. Another open problem is to determine the impact of our result on some stream ciphers which appear to be vulnerable to several attacks exploiting the existence of some function with a low degree [6], [42].

ACKNOWLEDGMENT

The authors would like to greatly thank the anonymous reviewers and Gilles Van Assche for their helpful comments.

REFERENCES

[1] X. Lai, “Higher order derivatives and differential cryptanalysis,” in *Proc. Symp. Commun., Coding Cryptography*, 1994, pp. 227–233, in honor of J. L. Massey on the occasion of his 60th birthday, Kluwer Academic Publishers.

[2] L. R. Knudsen, “Truncated and higher order differentials,” in *Proc. Fast Software Encryption - FSE’94*, pp. 196–211.

[3] S. Moriai, T. Shimoyama, and T. Kaneko, “Higher order differential attack of CAST cipher,” in *Proc. Fast Software Encryption - FSE’98*, pp. 17–31.

[4] N. Courtois and J. Pieprzyk, “Cryptanalysis of block ciphers with overdefined systems of equations,” in *Proc. Advances in Cryptology - ASIACRYPT 2002*, pp. 267–287.

[5] N. Courtois and W. Meier, “Algebraic attacks on stream ciphers with linear feedback,” in *Proc. Advances in Cryptology - EUROCRYPT 2003*, pp. 345–359.

[6] I. Dinur and A. Shamir, “Cube attacks on tweakable black box polynomials,” in *Proc. Advances in Cryptology - EUROCRYPT 2009*, pp. 278–299.

[7] A. Canteaut and M. Videau, “Degree of composition of highly nonlinear functions and applications to higher order differential cryptanalysis,” in *Proc. Advances in Cryptology - EUROCRYPT 2002*, pp. 518–533.

[8] C. Boura, A. Canteaut, and C. De Cannière, “Higher-order differential properties of Keccak and Luffa,” in *Proc. Fast Software Encryption - FSE 2011*, pp. 252–269.

[9] K. Nyberg and L. Knudsen, “Provable security against a differential attack,” *J. Cryptol.*, vol. 8, no. 1, pp. 27–37, 1995.

[10] C. Carlet, “Boolean functions,” in *Encyclopedia of Cryptography and Security*, H. C. A. van Tilborg and S. Jajodia, Eds., 2nd ed. New York: Springer-Verlag, 2011, pp. 162–165.

[11] A. Joux*, Algorithmic Cryptanalysis*. London, U.K.: Chapman & Hall/CRC Press, 2009.

[12] C. Carlet, P. Charpin, and V. Zinoviev, “Codes, bent functions and permutations suitable for DES-like cryptosystems,” *Des. Codes Cryptography*, vol. 15, no. 2, pp. 125–156, 1998.

[13] A. Bhattacharyya, S. Kopparty, G. Schoenebeck, M. Sudan, and D. Zuckerman, “Optimal testing of Reed-Muller codes,” in *Proc. IEEE Symp. Found. Comput. Sci.*, Oct. 2010, pp. 488–497.

[14] N. Alon, T. Kaufman, M. Krivelevich, S. Litsyn, and D. Ron, “Testing Reed-Muller codes,” *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 4032–4039, Nov. 2005.

[15] M.-J. O. Saarinen, “Chosen-IV statistical attacks on eStream ciphers,” in *Proc. International Conference on Security and Cryptography - SECRYPT 2006*, pp. 260–266.

[16] H. Englund, T. Johansson, and M. S. Turan, “A framework for chosen IV statistical analysis of stream ciphers,” in *Proc. Progress in Cryptology - INDOCRYPT 2007*, pp. 268–281.

[17] S. Fischer, S. Khazaei, and W. Meier, “Chosen IV statistical analysis for key recovery attacks on stream ciphers,” in *Proc. Progress in Cryptology - AFRICACRYPT 2008*, pp. 236–245.

[18] M. Vielhaber, Breaking ONE.FIVIUM by AIDA an algebraic IV differential attack Cryptology ePrint Archive, 2007, Rep. 2007/413.

[19] J.-P. Aumasson, E. Käsper, L. Knudsen, K. Matusiewicz, R. Ødegård, T. Peyrin, and M. Schläffer, “Distinguishers for the compression function and output transformation of Hamsi-256,” in *Proc. Information Security and Privacy - ACISP 2010*, pp. 87–103.

[20] T. Jakobsen and L. Knudsen, “The interpolation attack on block ciphers,” in *Proc. Fast Software Encryption - FSE’97*, pp. 28–40.

[21] K. Aoki, “Efficient evaluation of security against generalized interpolation attack,” in *Proc. Selected Areas in Cryptography - SAC’99*, pp. 135–146.

[22] B. Sun, L. Qu, and C. Li, “New cryptanalysis of block ciphers with low algebraic degree,” in *Proc. Fast Software Encryption - FSE 2009*, pp. 180–192.

[23] T. Shimoyama, S. Moriai, and T. Kaneko, “Improving the Higher Order Differential Attack and Cryptanalysis of the KN Cipher,” in *Proc. Information Security - ISW’97*, pp. 32–42.

[24] N. Katz, “On a theorem of Ax,” *Am. J. Math.*, vol. 93, pp. 485–499, 1971.

[25] M. Duan and X. Lai, Improved zero-sum distinguisher for full round Keccak-f permutation IACR ePrint, 2011, Rep. 2011/023.

[26] G. Bertoni, J. Daemen, M. Peeters, and G. V. Assche, The KECCAK reference, NIST, 2011 [Online]. Available: http://keccak.noekeon.org/Keccak-reference-3.0.pdf

[27] S. Konyagin and F. Pappalardi, “Enumerating permutation polynomials over finite fields by degree,” *Finite Fields Their Appl.*, vol. 8, no. 4, pp. 548–553, 2002.

[28] P. Das, “The number of permutation polynomials of a given degree over a finite field,” *Finite Fields Their Appl.*, vol. 8, no. 4, pp. 478–490, 2002.

[29] C. Wells, “The degrees of permutation polynomials over finite fields,” *J. Comb. Theory*, vol. 7, no. 1, pp. 49–55, 1969.

[30] *Data Encryption Standard (DES)*, U.S. Department of Commerce/National Bureau of Standards, 1999, Federal Information Processing Standards Publication 46-3, FIPS PUB 46-3.

[31] K. Nyberg, ““Provable” security against differential and linear cryptanalysis,” in *Proc. Fast Software Encryption - FSE 2012*, pp. 1–8.

[32] K. Nyberg, “Differentially uniform mappings for cryptography,” in *Proc. Advances in Cryptology - EUROCRYPT’93*, pp. 55–64.

[33] J. Patarin, “Security of random Feistel schemes with 5 or more rounds,” in *Proc. Advances in Cryptology - CRYPTO 2004*, pp. 106–122.

[34] K. Nyberg, “S-boxes and round functions with controllable linearity and differential uniformity,” in *Proc. Fast Software Encryption - FSE’94*, pp. 111–130.

[35] J. Daemen and V. Rijmen*, The Design of Rijndael: AES—The Advanced Encryption Standard*. New York: Springer-Verlag, 2002.

[36] J. Daemen and V. Rijmen, “Understanding two-round differentials in AES,” in *Proc. Security and Cryptography for Networks - SCN 2006*, pp. 78–94.

[37] R. Benadjila, O. Billet, H. Gilbert, G. Macario-Rat, T. Peyrin, M. Robshaw, and Y. Seurin, SHA-3 proposal: ECHO, NIST, 2009 [Online]. Available: http://crypto.rd.francetelecom.com/echo

[38] C. Boura and A. Canteaut, "Zero-sum distinguishers for iterated permutations and application to Keccak-f and Hamsi-256," in *Proc. Selected Areas in Cryptography - SAC 2010*, pp. 1–17.

[39] J.-P. Aumasson and W. Meier, "Zero-sum distinguishers for reduced Keccak-f and for the core functions of Luffa and Hamsi," presented at the Rump Session Cryptographic Hardware Embedded Syst., 2009.

[40] H. Wu, The hash function JH, NIST, 2011 [Online]. Available: http://www3.ntu.edu.sg/home/wuhj/research/jh/

[41] I. Dinur and A. Shamir, "An improved algebraic attack on Hamsi-256," in *Proc. Fast Software Encryption - FSE 2011*, pp. 88–106.

[42] I. Dinur and A. Shamir, "Breaking Grain-128 with dynamic cube attacks," in *Proc. Fast Software Encryption - FSE 2011*, pp. 167–187.

Christina Boura received a diploma in Mathematics from the Athens National University. Since 2010, she is a PhD student in the SECRET project-team at INRIA, the French National Research Institute in Computer Science. She is working on symmetric cryptography.

Anne Canteaut received the French engineer's degree from the École Nationale Supérieure de Techniques Avancées in 1993 and the Ph.D. degree in computer science from the University of Paris VI, France, in 1996. Since 1997, she has been a researcher with the French National Research Institute in Computer Science (INRIA), Paris-Rocquencourt. She is currently Director of Research and the scientific head of the SECRET research team at INRIA. Her research interests include cryptography and coding theory.

Dr. Canteaut has served on program committees for several international conferences such as Eurocrypt, Crypto and FSE. She served on the Editorial Board of the IEEE TRANSACTIONS ON INFORMATION THEORY (2005 to 2008).