On the influence of the algebraic degree of F^{-1} on the algebraic degree of $G \circ F$

Christina Boura[1,2] and Anne Canteaut[1]

[1] SECRET Project-Team - INRIA Paris-Rocquencourt - B.P. 105
78153 Le Chesnay Cedex - France
[2] Gemalto - 6, rue de la Verrerie - 92447 Meudon sur Seine - France
Christina.Boura@inria.fr, Anne.Canteaut@inria.fr

Abstract. We present a study on the algebraic degree of iterated permutations seen as multivariate polynomials. Our 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 non-injective 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.

1 Introduction

Most of the symmetric cryptographic primitives that are used nowadays, including block ciphers and hash functions, base their design 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 [27,26,28] attacks, algebraic attacks [13,12] or cube attacks [16].

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 believed. 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, [11] 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 [10] 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 primitives of the round function are not permutations. This is for example the case for the DES, that uses a collection of eight 6×4 balanced functions. For such functions, the notion of inverse does not obviously exist. We show however, that the overall degree of the cipher depends on the degree of the inverse of a balanced expansion of the function and 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 [31]. 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 finalists of the SHA-3 competition, ECHO and JH.

The rest of the paper is organized as follows. After some preliminaries on the algebraic degree of a vectorial function, the technique of higher-order differential cryptanalysis is recalled in Section 2 and it is illustrated by the attack proposed by Jakobsen and Knudsen [23] against the $\mathcal{KN}$ block cipher. Section 3 presents the main result on the influence of the inverse of a permutation F to the degree of $G \circ F$ and includes some corollaries. A variant of the main result for non-injective balanced functions is presented in Section 4, while some applications are described in Section 5.

2 Exploiting a low algebraic degree in cryptanalysis

2.1 Degree of a vectorial function

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 usually defined by the algebraic degrees of its coordinates as follows.

Definition 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*:*

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

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.

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 $2^n - 1$,

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

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.

2.2 Higher-order differential cryptanalysis

Many statistical attacks against symmetric cryptosystems exploit the fact that the system involves a family of functions $(F_k)_{k\in\mathcal{K}}$ (resp. of permutations) having both following properties:

- the inputs and outputs of $(F_k)_{k\in\mathcal{K}}$ can be computed from plaintext/ciphertext pairs;

– $(F_k)_{k\in\mathcal{K}}$ is not a pseudo-random function (resp. a pseudo-random permutation).

Roughly speaking, this second property means that a randomly chosen function within this family can be distinguished with some non-negligible advantage from a randomly chosen function (resp. permutation) (see Chapter 3 in [5] for formal definitions of this notion). Several properties may be used as a distinguisher including the fact that some given coefficients in the algebraic normal forms of some Boolean functions derived from F_k are not distributed as it is expected for a family of randomly chosen functions. Indeed, the coefficients of the algebraic normal form of a Boolean function f can be easily computed from some input-output pairs of f as follows:

$$a_u = \sum_{x\in\mathbf{F}_2^n, x_i \le u_i} f(x) \bmod 2 \ .$$

In particular, this formula shows that a_u can be computed from $2^{wt(u)}$ pairs of inputs-outputs of f. 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 $\mathcal{O}(n2^n)$ (see e.g. [24, p. 286]).

The simplest attack exploiting some property of the coefficients of the algebraic normal form is the *higher-order differential attack* introduced by Knudsen [26]: this attack uses that, for all values of k, all coordinates of F_k have degree strictly less than n (resp. strictly less than $n-1$ in the case of permutations). The algebraic degree of F_k is then of primary importance since the data complexity of this cryptanalysis is proportional to $2^{\deg F_k}$. Indeed, Bhattacharyya *et al.* have recently shown that testing whether a Boolean function has degree at most d (or equivalently whether it belongs to the Reed-Muller code of order d) with constant error probability requires the knowledge of $O(2^d)$ values of the function only [8]. Moreover, this data complexity is known to be optimal [1, Corollary 7].

In the case of iterated block ciphers, *i.e.*, ciphers consisting of several iterations of the same round permutation P parameterized by different round keys:

$$P_{k_r} \circ \ldots \circ P_{k_2} \circ P_{k_1} \ ,$$

the target function F_k whose inputs and outputs can be computed by the attacker usually corresponds to the composition of several rounds of the cipher. Typically, F_k 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 [23], 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 [34,28].

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 [33,20,22,36,3]. Cube attacks [16] and algebraic attacks [13,12] also exploit some 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 [23,2,35].

2.3 Attacking the $\mathcal{KN}$-cipher and its variant

An example of an attack exploiting the low algebraic degree of a symmetric primitive is the higher-order differential attack presented by Jakobsen and Knudsen [23] against the $\mathcal{KN}$-cipher.

This construction, proposed by Nyberg and Knudsen in [31], is a 6-round Feistel cipher over $\mathbf{F}_2^{64}$ with a 198-bit secret key. Its round permutation is defined as follows

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

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 $\mathbf{F}_2^{33}$.

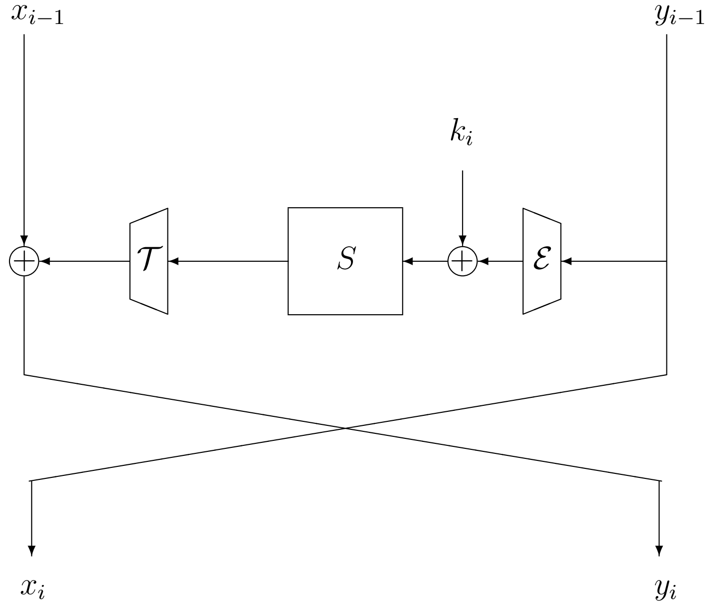

Fig. 1. Round i of the $\mathcal{KN}$ cipher

An important remark is that the decryption function is exactly the same as the encryption function, except that the round keys have to be used in reverse order. This is because the round permutation obtained with the Feistel construction is involutive. 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 [29]. More precisely, some lower bounds on the probabilities of the best differential and of the best linear approximation show that 6 rounds of this cipher are resistant to these attacks.

However, one of the main weaknesses of this cipher, identified by Jakobsen and Knudsen [23], 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 to 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 4-round and 5-round Feistel ciphers which have respective data complexity 2^{16} and 2^{32} [32]. Here, the whole encryption function can be distinguished from a random permutation with data complexity 2^{17}. Also, the last round key can be recovered by an exhaustive search: for each possible value for the last round key k_6, the attacker decrypts the last round, computes x_5 and she determines whether the function $x_0 \mapsto x_5$ has degree less than or equal to 8. This last attack recovers 33 key bits with average time complexity 2^{41} and data complexity 2^9 pairs of chosen plaintexts-ciphertexts. This attack has been improved by Shimoyama *et al.* [34] who replaced the exhaustive search for k_6 by the solving of a linear system in the bits of k_6, since the involved equations have degree $(\deg S - 1)$. The data complexity of the attack is then unchanged but the average time complexity for recovering the 33-key bits reduces to 2^{14}.

Therefore, it is now well-known that, in an r-round block cipher, the round permutation P must be chosen such that $(\deg P)^r$ is much higher than the block size. Since a similar distinguisher can also be applied by an attacker to the decryption function, *i.e.*, to the function

$$D_k = P_{k_1}^{-1} \circ \ldots \circ P_{k_{r-1}}^{-1} \circ P_{k_r}^{-1} \ ,$$

the inverse of the round permutation must also satisfy this property, *i.e.*, $(\deg P^{-1})^r$ must be much higher than the block size. In a Feistel cipher, the condition on the degree of the round function can be refined by imposing that $(\deg S)^{r-2}$ must be much higher than half of the block size. But, in this case, the condition on the degree of the inverse of S is not necessary since S^{-1} is not involved neither in the encryption function nor in the decryption function. It may only affect the complexity of some algebraic attacks [13]. Therefore, a variant of this cipher, that we name $\mathcal{KN}'$, suggested by Nyberg and Knudsen in the same paper [31] 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 [30]. 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 $\frac{n+1}{2}$ [29]. 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 $\mathcal{KN}$:

$$\begin{aligned} \widetilde{\sigma} : \mathbf{F}_2^8 &\to \quad \mathbf{F}_2^8 \\ x &\mapsto t \circ \sigma\left(e(x)\right)) \end{aligned}$$

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 on Figure 2: it is defined by

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

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

The attack proposed by Jakobsen and Knudsen against $\mathcal{KN}$ does not apply to $\mathcal{KN}'$, since the round permutation has degree 5, and the previously used upper bound does not provide any relevant information on the degree of the left part of the output for 5 rounds or more. This example tends to show that the Sbox used in a Feistel cipher must have good cryptographic properties but, if this Sbox is a permutation, it does not seem necessary that its inverse has good cryptographic properties too. In the following, we show that, even if σ^{-1} is never involved in the $\mathcal{KN}'$ cipher, its algebraic degree affects the security of the cipher regarding higher-order differential attacks. We actually exhibit a new upper bound on the degree of the composition $G \circ F$, for any G, which involves the degree of F^{-1}.

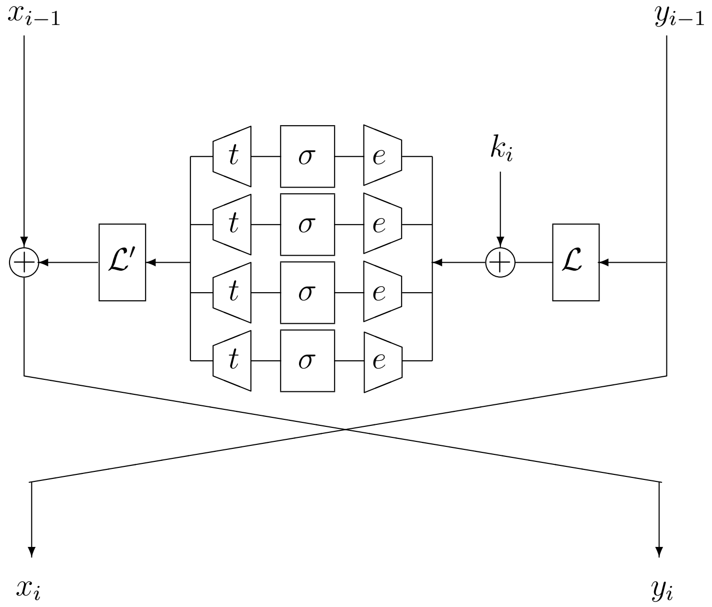

Fig. 2. Round i of the $\mathcal{KN}'$ cipher

3 On the degree of $G \circ F$ when F is a permutation

3.1 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 [11] and [10].

Proposition 1. *[11] 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 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 set 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 [25]: for any function F and any nonzero $b\in\mathbf{F}_2^n$, we have

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

Since F is a permutation, any nonzero linear combination of its coordinates is balanced. 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 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 $\deg G\leq\lceil\frac{n-1}{2}\rceil$. But, this condition on G does not hold in the problem raised by the search of a distinguisher on 5 rounds of the $\mathcal{KN}'$ cipher.

It has recently be shown in [10] that the bound given by Proposition 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.

3.2 Main result

We now show that, when F is a permutation, the upper bound given by Corollary 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 2. *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 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 $a_L = 0$.

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

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,

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

Now, we define the Boolean function

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

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 [19] on the inverse of the quadratic permutation χ over $\mathbf{F}_2^5$ used in the hash function Keccak [7]. 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 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,

$$\left\lfloor \frac{n-1-\deg G}{\deg(F^{-1})} \right\rfloor = \begin{cases} \left\lfloor \frac{n-\deg G}{\deg(F^{-1})} \right\rfloor & \text{if } n - \deg G \not\equiv 0 \bmod \deg(F^{-1}) \\ \frac{n-\deg G}{\deg(F^{-1})} - 1 & \text{otherwise.} \end{cases}$$

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

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

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 of 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 obvious 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 1 was relevant only for $\deg G < \left\lceil \frac{n-1}{\min(\deg F, \deg F^{-1})} \right\rceil$.

3.3 Some corollaries

Some simple corollaries of Theorem 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 and its well-known consequence.

Corollary 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* $\deg(F) = n-1$.

Moreover, 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 [25] 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. [11, Prop. 3]).

Corollary 3 also implies the following.

Corollary 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. The previous corollary 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 [10] when F is the concatenation of several smaller permutations.

Theorem 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 $2 \leq i < n_0 - \max\limits_{1 \leq j \leq s} \deg(S_j^{-1})$, we get from Corollary 4 that $\max\limits_{1 \leq j \leq s} \delta_i(S_j) \leq n_0 - 2$, 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 $i \geq n_0 - \max\limits_{1 \leq j \leq s} \deg(S_j^{-1})$, we get that

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

□

4 Generalization to balanced functions from $\mathbf{F}_2^n$ into $\mathbf{F}_2^m$ with $m < n$

On certain occasions in some symmetric primitives, the functions used to provide confusion are not permutations, but balanced functions $F : \mathbf{F}_2^n \to \mathbf{F}_2^m$, with $m < n$. An example of this design is the first encryption standard cipher, DES [21], whose round function uses a parallel application of eight different 6 × 4 Sboxes, all of them of degree 5 in the 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 3. *Let $F : \mathbf{F}_2^n \to \mathbf{F}_2^m$, with $m < n$, $F = (F_1, \ldots, F_m)$, 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 example, if $(n, m) = (6, 4)$, $v = (0, 1, 1, 0)$ 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, one expansion of F can be defined by associating to 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 3. *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 $K \subset \{1, \ldots, m\}$ and $L \subset \{1, \ldots, n\}$. Let π_K denote the product of the coordinates F_i for $i \in K$. Then, the coefficient $a_{K,L}$ of the monomial $\prod_{i \in \{1,\ldots,n\} \setminus L} x_i$ in the algebraic normal form of F is given by

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

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

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

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 $x_{m+1} \ldots x_n$. Therefore, the restriction of such a monomial to any set $\{x \in \mathbf{F}_2^n : x_i = 1, i \in K\}$ with $K \subset \{1, \ldots, m\}$ has degree strictly less than $(m-k) + (n-m) = (n-k)$. It follows that, for any choice of $K \subset \{1, \ldots, m\}$, $H_{K,L}$ 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 $L \subset \{1,\ldots,m\}$ 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 $x_{m+1}\ldots x_n \prod_{i\in I} x_i$ for some set $I \subset \{1,\ldots,m\}$ 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 $K = \{1,\ldots,m\}\setminus I$ where $x_{m+1}\ldots x_n \prod_{i\in I} x_i$ is the monomial with 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 $|L| < n$ [10, Prop 1]. Since L is minimal for inclusion and

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

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 $x_{m+1}\ldots x_n \prod_{i\notin K} x_i$, 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 2 can be deduced now for the case of non-injective balanced functions.

Corollary 5. *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 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 3 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. [10, Prop 1]. Moreover, when F is a permutation, we have shown in Corollary 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 non-injective balanced function.

Corollary 6. *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 3 applied to $k = m$ 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 $x_{m+1} \dots x_n$. Since a monomial of degree less than $(n-m)$ cannot be a factor of $x_{m+1} \dots x_n$, this equivalently means that any monomial in the ANF of any coordinate of P_F^{-1} is not a factor of $x_{m+1} \dots x_n$. Let

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

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

$$f(x,y) = \sum_{v \preceq y} \left[\sum_{u \preceq x} a_{u,v} \right] \bmod 2 \ ,$$

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

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

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,

$$\sum_{y \in \mathbf{F}_2^{n-m}} f(x,y) \equiv \sum_{u \preceq x} a_{u,1_{n-m}} \bmod 2 \ .$$

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

$$\sum_{y \in \mathbf{F}_2^{n-m}} f(x,y) \bmod 2 = 0$$

for all $x \in \mathbf{F}_2^m$. It is worth noticing that this property is similar to the property used in cube attacks (see [16, Theorem 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

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

□

5 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 degree of some chosen permutations that are the main building blocks of some well-known block ciphers and hash functions. We will start with the case of the block cipher $\mathcal{KN}'$ described in Section 2.3.

5.1 Attacking the $\mathcal{KN}'$-cipher

We will show now how Theorem 1 can be used to attack the $\mathcal{KN}'$-cipher. 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. Therefore, we need to express x_r as a function of x_0. 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 4 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

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

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

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

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

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

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

Degree of $G \circ S$. We now apply Corollary 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 5 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 2 which additionally exploits the fact that S corresponds to the concatenation of 4 permutations σ defined over $\mathbf{F}_2^9$.

Variant with non-bijective Sboxes. The nonlinear function in $\mathcal{KN}'$ can also be seen as the concatenation of 4 balanced Sboxes σ' from $\mathbf{F}_2^9$ into $\mathbf{F}_2^8$. Instead of applying Corollary 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 6 applies and also shows that x_5 is a function of degree at most 29 of x_0.

5.2 On the algebraic degree of Rijndael-256

Rijndael-128 [14] is the algorithm selected by the NIST in 2000 as the winner of the 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 a N_b-bit state, that is represented as a $4 \times t$-byte matrix $A = (a_{i,j})$, with $t = N_b/32$. The states for Rijndael-128 and Rijndael-256 are for example depicted on Figure 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. The states of Rijndael-256 and Rijndael-128

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

- `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.
- `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.
- `MixColumns`: Linear transformation that applies in parallel to every column of the state.
- `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 of the key size. These values can be found in Table 1.

Table 1. Number of rounds for the Rijndael block cipher.

		$\boldsymbol{N_b}$				
		128	**160**	**192**	**224**	**256**
$\boldsymbol{N_k}$	**128**	10	11	12	13	14
	192	12	12	12	13	14
	256	14	14	14	14	14

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 3 rounds of encryption are enough for achieving the maximal degree.

We will show using the results of Section 3, that the above estimations are way too pessimistic. We will see in particular that for Rijndael-256, at least 7 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 [15], 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 a so-called SDS transformation: it consists of two layers of four 8×8 balanced Sboxes of degree 7, separated by a linear layer. Therefore, we can use Theorem 2 of [10] 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. We can apply thus 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 [10] gives now

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

Let $F = R^3$. F is a permutation of degree at most 113 and clearly $F^2 = R^6$. By bounding thus the degree of F^2 we get a bound for the degree of Rijndael-256 after six rounds. From

Theorem 2, we get that the constant γ associated to this permutation is at most 127 and we deduce finally that

$$\deg F^2 = \deg R^6 \leq 256 - \frac{256 - 113}{127} < 255.$$

Therefore, at least 7 rounds are needed to achieve the maximal degree 255.

5.3 Application to the ECHO hash function

The ECHO [6] 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,

- `BIG.SubWords` is a nonlinear transformation applied independently to every 128-bit cell. It consists of two AES rounds.
- 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 Theorem 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 2, the constant γ associated to this permutation is at most 466, as the degrees of $\mathtt{R}^2$ and of its inverse are both upper-bounded by 466, therefore

$$\deg F^2 = \deg \mathtt{R}^4 \leq 2048 - \frac{2048 - 466}{466} < 2046.$$

The same bounds hold for the inverse round transformation. Due to this observation, we are able to distinguish the inner permutation in ECHO from a random one. This can be done for instance by constructing zero-sum structures [9,4]. By choosing the intermediate states after 4 rounds of the permutation in the cosets of any subspace V with dimension 2^{2046}, we get zero-sum partitions for the entire P permutation.

5.4 Application to the JH hash function

JH [37] 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 3 different layers: an Sbox layer, a linear layer and a permutation layer P_d.

- 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 round constant bits, which are not xored to the state as done in other constructions.
- The **linear layer** mixes the 2^d words two by two.
- The **permutation** P_d permutes the words of the state.

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

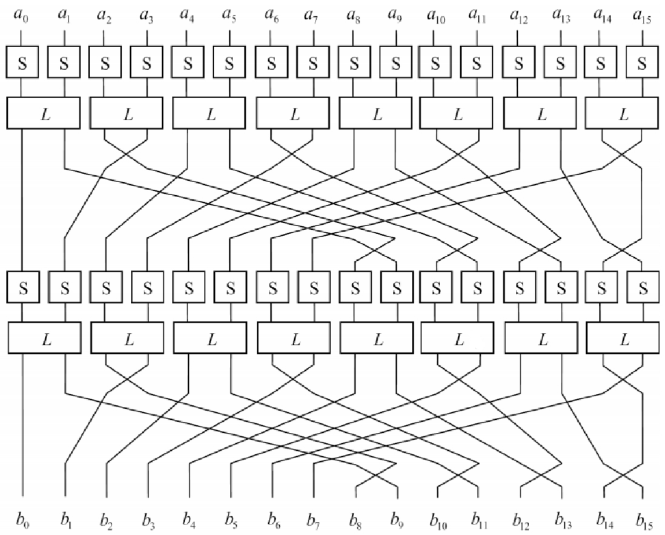
a_0 a_1 a_2 a_3 a_4 a_5 a_6 a_7 a_8 a_9 a_{10} a_{11} a_{12} a_{13} a_{14} a_{15}
S
L
b_0 b_1 b_2 b_3 b_4 b_5 b_6 b_7 b_8 b_9 b_{10} b_{11} b_{12} b_{13} b_{14} b_{15}

Fig. 4. Two rounds of R_4

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 6 rounds is at most $\deg(R_8^6) \leq 3^6 = 729$ and consequently the maximal degree seems to be reached just after 7 rounds of encryption. We will show again by applying the results of Section 3 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 $\mathbf{F}_2^{2^{r+1}}$. Thus, for $2 \leq r \leq 8$ a bound on the degree of R_8^r can be obtained with the help of Theorem 2 in [10]:

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

The bounds on the degree up to 8 rounds of the permutation, given by the above formula, can be seen in Table 2. The same bounds hold for the inverse permutation.

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

Table 2. Upper bounds on the degree of up to 8 rounds of the JH permutation.

Using now Theorem 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

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

6 Conclusions

Our work 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 non-injective function, the minimal degree of the inverse of a permutation expanding the function, has some influence on the degree of the iterated function. This observation can be used for exhibiting non-ideal 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 [28,18]. 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 [16,17].

References

1. N. Alon, T. Kaufman, M. Krivelevich, S. Litsyn, and D. Ron. Testing Reed-Muller codes. *IEEE Transactions on Information Theory*, 51(11):4032–4039, 2005.
2. K. Aoki. Efficient evaluation of security against generalized interpolation attack. In *Selected Areas in Cryptography - SAC'99*, volume 1758 of *Lecture Notes in Computer Science*, pages 135–146. Springer, 2000.

3. J.-P. Aumasson, E. Käsper, L.R. Knudsen, K. Matusiewicz, R. Ødegård, T. Peyrin, and M. Schläffer. Distinguishers for the compression function and output transformation of Hamsi-256. In *Information Security and Privacy - ACISP 2010*, volume 6168 of *Lecture Notes in Computer Science*, pages 87–103. Springer, 2010.
4. 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 of Cryptographic Hardware and Embedded Systems - CHES 2009, 2009.
5. M. Bellare and P. Rogaway. *Introduction to Modern Cryptography*. 2005. Available at `http://cseweb.ucsd.edu/~mihir/cse207`.
6. R. Benadjila, O. Billet, H. Gilbert, G. Macario-Rat, T. Peyrin, M. Robshaw, and Y. Seurin. SHA-3 Proposal: ECHO. Submission to NIST (Round 2), available at `http://crypto.rd.francetelecom.com/echo`, 2009.
7. G. Bertoni, J. Daemen, M. Peeters, and G. Van Assche. The KECCAK reference. Submission to NIST (Round 3), available at `http://keccak.noekeon.org/Keccak-reference-3.0.pdf`, 2011.
8. A. Bhattacharyya, S. Kopparty, G. Schoenebeck, M. Sudan, and D. Zuckerman. Optimal testing of Reed-Muller codes. In *IEEE Symposium on Foundations of Computer Science - FOCS 2010*, pages 488–497. IEEE Computer Society, 2010.
9. C. Boura and A. Canteaut. Zero-sum distinguishers for iterated permutations and application to Keccak-f and Hamsi-256. In *Selected Areas in Cryptography - SAC 2010*, volume 6544 of *Lecture Notes in Computer Science*, pages 1–17. Springer, 2010.
10. C. Boura, A. Canteaut, and C. De Cannière. Higher-order differential properties of Keccak and Luffa. In *Fast Software Encryption - FSE 2011*, volume 6733 of *Lecture Notes in Computer Science*, pages 252–269. Springer, 2011.
11. A. Canteaut and M. Videau. Degree of composition of highly nonlinear functions and applications to higher order differential cryptanalysis. In *Advances in Cryptology - EUROCRYPT 2002*, volume 2332 of *Lecture Notes in Computer Science*, pages 518–533. Springer-Verlag, 2002.
12. N. Courtois and W. Meier. Algebraic attacks on stream ciphers with linear feedback. In *Advances in Cryptology - EUROCRYPT 2003*, volume 2656 of *Lecture Notes in Computer Science*, pages 345–359. Springer-Verlag, 2003.
13. N. Courtois and J. Pieprzyk. Cryptanalysis of block ciphers with overdefined systems of equations. In *Advances in Cryptology - ASIACRYPT'02*, volume 2501 of *Lecture Notes in Computer Science*, pages 267–287. Springer-Verlag, 2002.
14. J. Daemen and V. Rijmen. *The Design of Rijndael: AES - The Advanced Encryption Standard*. Springer, 2002.
15. J. Daemen and V. Rijmen. Understanding Two-Round Differentials in AES. In *Security and Cryptography for Networks - SCN 2006*, volume 4116 of *Lecture Notes in Computer Science*. Springer, 2006. pp. 78-94.
16. I. Dinur and A. Shamir. Cube attacks on tweakable black box polynomials. In *Advances in Cryptology - EUROCRYPT 2009*, volume 5479 of *Lecture Notes in Computer Science*, pages 278–299. Springer, 2009.
17. I. Dinur and A. Shamir. Breaking Grain-128 with dynamic cube attacks. In *Fast Software Encryption - FSE 2011*, volume 6733 of *Lecture Notes in Computer Science*, pages 167–187. Springer, 2011.
18. I. Dinur and A. Shamir. An improved algebraic attack on Hamsi-256. In *Fast Software Encryption - FSE 2011*, volume 6733 of *Lecture Notes in Computer Science*, pages 88–106. Springer, 2011.
19. M. Duan and X. Lai. Improved zero-sum distinguisher for full round Keccak-f permutation. IACR ePrint Report 2011/023, January 2011. `http://eprint.iacr.org/2011/023`.
20. H. Englund, T. Johansson, and M. S. Turan. A framework for chosen IV statistical analysis of stream ciphers. In *Progress in Cryptology - INDOCRYPT 2007*, volume 4859 of *Lecture Notes in Computer Science*, pages 268–281. Springer, 2007.
21. FIPS PUB 46-3. Data Encryption Standard (DES). Federal Information Processing Standards Publication 46-3, 1999. U.S. Department of Commerce/National Bureau of Standards.
22. S. Fischer, S. Khazaei, and W. Meier. Chosen IV statistical analysis for key recovery attacks on stream ciphers. In *AFRICACRYPT 2008*, volume 5023 of *Lecture Notes in Computer Science*, pages 236–245. Springer, 2008.
23. T. Jakobsen and L.R. Knudsen. The interpolation attack on block ciphers. In *Fast Software Encryption - FSE'97*, volume 1267 of *Lecture Notes in Computer Science*. Springer-Verlag, 1997.
24. A. Joux. *Algorithmic cryptanalysis*. Chapman & Hall/CRC Press, 2009.
25. N. Katz. On a theorem of Ax. *American Journal of Mathematics*, 93:485–499, 1971.
26. L. R. Knudsen. Truncated and higher order differentials. In *Fast Software Encryption - FSE'94*, volume 1008 of *Lecture Notes in Computer Science*, pages 196–211. Springer-Verlag, 1995.
27. X. Lai. Higher order derivatives and differential cryptanalysis. In *Proc. "Symposium on Communication, Coding and Cryptography", in honor of J. L. Massey on the occasion of his 60'th birthday*. Kluwer Academic Publishers, 1994.
28. S. Moriai, T. Shimoyama, and T. Kaneko. Higher order differential attak of CAST cipher. In *Fast Software Encryption - FSE'98*, volume 1372 of *Lecture Notes in Computer Science*, pages 17–31. Springer, 1998.

29. K. Nyberg. Differentially uniform mappings for cryptography. In *Advances in Cryptology - EUROCRYPT'93*, volume 765 of *Lecture Notes in Computer Science*, pages 55–64. Springer-Verlag, 1993.
30. K. Nyberg. S-boxes and round functions with controllable linearity and differential uniformity. In *Fast Software Encryption - FSE'94*, volume 1008 of *Lecture Notes in Computer Science*, pages 111–130. Springer-Verlag, 1995.
31. K. Nyberg and L.R. Knudsen. Provable security against a differential attack. *Journal of Cryptology*, 8(1):27–37, 1995.
32. J. Patarin. Security of random Feistel schemes with 5 or more rounds. In *Advances in Cryptology - CRYPTO 2004*, volume 3152 of *Lecture Notes in Computer Science*, pages 106–122. Springer, 2004.
33. M.-J. O. Saarinen. Chosen-IV statistical attacks on eStream ciphers. In *SECRYPT 2006 - International Conference on Security and Cryptography*, pages 260–266. INSTICC Press, 2006.
34. T. Shimoyama, S. Moriai, and T. Kaneko. Improving the higher order differential attack and cryptanalysis of the *KN* cipher. In *Information Security - ISW'97*, volume 1396 of *Lecture Notes in Computer Science*, pages 32–42. Springer, 1998.
35. B. Sun, L. Qu, and C. Li. New cryptanalysis of block ciphers with low algebraic degree. In *Fast Software Encryption - FSE 2009*, volume 5665 of *Lecture Notes in Computer Science*, pages 180–192. Springer, 2009.
36. M. Vielhaber. Breaking ONE.FIVIUM by AIDA an algebraic IV differential attack. Cryptology ePrint Archive, Report 2007/413, 2007. `http://eprint.iacr.org/2007/413`.
37. H. Wu. The hash function JH. Submission to NIST (Round 3) available at `http://www3.ntu.edu.sg/home/wuhj/research/jh/`, 2011.