An Upper Bound on the Number of m-Resilient Boolean Functions

Claude Carlet[1] and Aline Gouget[2]

[1] INRIA, Domaine de Voluceau, BP 105 - 78153 Le Chesnay Cedex, France; also member of GREYC-Caen and of University of Paris 8
Claude.Carlet@inria.fr
[2] GREYC, Université de Caen, 14032 Caen Cedex, France
gouget@info.unicaen.fr

Abstract. The enumeration of m-resilient Boolean functions in n variables would be a quite useful information for cryptography. But it seems to be an intractable open problem. Upper and lower bounds have appeared in the literature in the mid $80's$. Since then, improving them has been the goal of several papers. In this paper, we give a new upper bound which partially improves upon all the known bounds.

Keywords: Cryptography, Stream cipher, Boolean function, Resilient function

1 Introduction

The principle of private cryptography relies on the share-out of a private key between the sender of a message and its receiver. Symmetric cryptosystems are commonly used owing to their efficiency. Currently, there is no mathematical proof to ensure the unconditional security of the system except for the famous Vernam [13] scheme. This system produces the encoded text by adding bitwisely the plain text and the private key. Then the receiver retrieves the plain text by using the same addition of the encoded text and the private key. In practice, since the length of the private key must equal the length of the plain text, pseudo-random generators are used for stream ciphers in order to minimize the size of the private key (but the unconditional security is then no longer ensured). In order to achieve maximal security, these systems are much studied.

The basic component of a keystream generator is the Linear Feedback Shift Register (LFSR). The generic example of a keystream generator is composed of n LFSR whose outputs are combined by a Boolean function from $\mathbb{F}_2^n$ to $\mathbb{F}_2$. The security of the system relies, in a central way, on the choice of the Boolean function. Subsequently, the Boolean functions used to combine several LFSR, called combining functions, must fulfil several criteria. They must be *balanced*, *i.e.*, they must take the value 1 and the value 0 with the same probability on the set $\mathbb{F}_2^n$. They must have high algebraic degrees (see definition at section 2) so that the keystream generator resists the Berlekamp-Massey's attack [6]. The generator must also resist the Siegenthaler's correlation attack [12]. This comes down to choose a combining function which is correlation-immune of a high order m [11], *i.e.*, whose output distribution does not change when m input values (*i.e.*, m coordinates of the input vector) are fixed. If the combining function is correlation-immune of order m, the attacker has to guess the initialization of at least $m+1$ LFSR to observe a correlation between them and the output of the pseudo-random generator during a correlation attack. Combining functions must also have high non-linearities in order to prevent linear approximation. Of course these criteria are partially opponent and tradeoffs exist.

Enumerating the Boolean functions satisfying one or several of these criteria is useful for several reasons. Firstly because it indicates for which values of the parameters $(n, \ldots)$ there is a chance of finding good cryptographic functions by random search. Secondly because a large number of functions is necessary if we want to impose extra constraints on the functions or if we want to modify the cryptosystems using them by having the function as part of the secret key.
Mitchell [7] proposed a number of open problems with partial results about enumerating Boolean functions satisfying various criteria, including balancedness and correlation-immunity. The first bounds on the number of first order correlation-immune Boolean functions were lower bounds (see [7, 14, 8, 5]). In 1990, Yang and Guo published the first upper bound on such functions. Park, Lee, Sung and Kim [8] proceeded further and improved upon Yang-Guo's bound. In 1995, Schneider [10] used a new idea to improve upon previous bounds. He obtained bounds for the numbers of mth-order correlation-immune functions and of m-resilient functions. Carlet and Klapper [1] obtained a general upper bound on the number of Boolean functions whose distances to affine functions are all divisible by 2^m. They deduced an upper bound on the number of m-resilient functions and improved upon Schneider's bound for m large.

In the present paper, we obtain an upper bound on m-resilient functions ($m \geq \frac{n}{2} - 1$), and improve upon Schneider's bound for all values $m > \frac{n}{2} - 1$. We show with tables of values that our bound partially improves upon Carlet-Klapper's bound (the expressions of both bounds seem difficult to compare mathematically).

The organization of the paper is as follows. Section 2 introduces the notation and the definitions that are needed in the paper including the definition of correlation-immunity. Section 3 reviews the previous upper bounds on the numbers of first order correlation-immune functions, *i.e.*, Yang *et al*'s and Park *et al*'s bounds, and of m-resilient functions, *i.e.*, Schneider's and Carlet-Klapper's bounds. Extensions of Yang *et al.*'s and Park *et al.*'s bounds are given for the case of 1-resilient functions in this section for the first one and in appendix B for the second one. Section 4 introduces a new upper bound on the number of m-resilient functions. We give a table of values corresponding to the ratio of Schneider's bound to the new bound, and a second table corresponding to the ratio of Carlet-Klapper's bound to the new one.

2 Notation and Definitions

Let n be any positive integer. We denote by $\bigoplus$ the usual addition in $\mathbb{F}_2$ and in $\mathbb{F}_2^n$. The *Hamming weight* $w_H(u)$ of a word u in $\mathbb{F}_2^n$ is the number of its components equal to 1. We denote by $\preceq$ the partial order on the words of $\mathbb{F}_2^n$, *i.e.*, $(u_1,\ldots,u_n) \preceq (v_1,\ldots,v_n)$ if and only if $(u_i = 1) \Rightarrow (v_i = 1)$. Any Boolean function f in n variables, $f : \mathbb{F}_2^n \mapsto \mathbb{F}_2$, admits a unique Algebraic Normal Form (A.N.F.):

$$f(x_1,\ldots,x_n) = \bigoplus_{u\in\mathbb{F}_2^n} a_u \left(\prod_{i=1}^{n} x_i^{u_i}\right) = \bigoplus_{u\in\mathbb{F}_2^n} a_u\, x^u \ .$$

The function $g : u \mapsto a_u$ is called the *Möbius transform* of f. For any word u, the coefficient a_u belongs to $\mathbb{F}_2$, and can be computed thanks to the formula

$$a_u = \bigoplus_{v\in\mathbb{F}_2^n, v\preceq u} f(v) \ . \tag{1}$$

The *algebraic degree* of a Boolean function f is the degree of its algebraic normal form. The *Hamming weight* $w_H(f)$ of a Boolean function f in n variables is the size of its support, *i.e.*, the size of the set $\{x \in \mathbb{F}_2^n | f(x) = 1\}$. A Boolean function f in n variables is called *balanced* if its Hamming weight equals 2^{n-1}.

Definition 1. [11] *Let $X^{[j]} = (X_1^{[j]}, X_2^{[j]}, ..., X_n^{[j]})$ be the n-tuple of LFSR output digits at time j. The combining function f is* mth-order correlation-immune *if every m-tuple obtained by fixing m components from $X^{[j]}$ is statistically independent of the random value $Z = f(X_1, X_2, \ldots, X_n)$ associated to arbitrary outputs of LFSR.*

A characterization of mth-order correlation-immune functions was given by Guo-Zhen and Massey in [4].

Definition 2. *Let f be a Boolean function in n variables. The* Walsh Transform *of f is defined as the following real-valued function over the vector space $\mathbb{F}_2^n$,*

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

where $u \cdot x$ stands for $\sum_{i=1}^{n} u_i\, x_i$.

Theorem 1. [4] *A Boolean combining function f in n variables is mth-order correlation-immune, where $1 \leq m \leq n$, if and only if for every word u in $\mathbb{F}_2^n$ such that $1 \leq w_H(u) \leq m$, $\hat{f}(u)$ equals* 0, i.e., *$f(x) \oplus u \cdot x$ is balanced for all u such that $1 \leq w_H(u) \leq m$.*

A balanced Boolean function in n variables which is correlation-immune of order m is called m-resilient. This notion was considered for the first time by Chor *et al.* in [2].

The tradeoff between the order of correlation-immunity and the algebraic degree was given by Siegenthaler.

Theorem 2. [12] *Let f be an mth-order correlation-immune Boolean function of degree d in n variables. Then $d \leq n - m$. Furthemore, if f is balanced then $d \leq n - m - 1$ if $m < n - 1$ and $d = 1$ if $m = n - 1$.*

This result leads to the first obvious bound on the numbers of m-resilient and mth-order correlation-immune functions. The number of mth-order correlation-immune functions in n variables is upper bounded by $2^{\sum_{i=0}^{n-m} \binom{n}{i}}$, and the number of m-resilient functions in n variables is upper bounded by $2^{\sum_{i=0}^{n-m-1} \binom{n}{i}}$ if $m < n - 1$.

3 Previous Upper Bounds

The number of mth-order correlation-immune Boolean functions is still unknown (an asymptotic formula is known, due to Denisov [3]). The first upper bound on the number of correlation-immune Boolean functions, published by Yang and Guo in 1990 [14], enumerates in fact the number of Boolean functions which satisfy partially the first order correlation-immunity criterion, *i.e.*, the functions f such that for two distinct integers i_1 and i_2, $f \oplus x_{i_1}$ and $f \oplus x_{i_2}$ are balanced. This leads to:

Proposition 1. [14] *Let n be a positive integer greater than 1. The number of 1st-order correlation-immune Boolean funtions in n variables is less than:*

$$\sum_{k=0}^{2^{n-2}} \sum_{r=0}^{k} \binom{2^{n-2}}{r}^2 \binom{2^{n-2}}{k-r}^2 .$$

Yang and Guo did not study the corresponding bound for 1-resilient functions. This can be done:

Proposition 2. *Let n be a positive integer greater than 1. The number of 1-resilient Boolean functions in n variables is less than:*

$$\sum_{a=0}^{2^{n-2}} \binom{2^{n-2}}{a}^4 .$$

We give the proof of this bound in appendix A.

This work was deepened by Park, Lee, Sung and Kim for 1st-order correlation-immunity. They showed that the number of correlation-immune functions is itself upper bounded by this same number as in Proposition 2. Park *et al.* obtained this bound by numbering the Boolean functions such that for three distinct integers i_1, i_2, i_3, the functions $f \oplus x_{i_1}$, $f \oplus x_{i_2}$ and $f \oplus x_{i_3}$ are balanced. They did not study the corresponding bound for 1-resilient Boolean functions. The bound obtained is quite complicated and is given in appendix B.

The number of balanced Boolean functions such that $f(x) \oplus x_i$ is balanced for three distinct integers could be calculated by considering the solutions of a system of four equations with eight unknowns. When the number of integers increases by one, only one new equation can be obtained and the number of unknowns is doubled. Thus, enumerating the number of Boolean functions such that $f(x) \oplus x_i$ is balanced for $i \in I \subseteq \{1, \ldots, n\}$ leads to considering $|I| + 1$ equations and $2^{|I|}$ unknowns. Thus, for n greater than 3, the gap between the number of equations and the number of unknowns is too large to obtain a bound which can be computed easily.

Maitra and Sarkar [5] found a sufficient condition for a function f to be such that $f(x) \oplus x_i$ is balanced for three values of i but f is not first order correlation-immune. A lower bound on the number of such functions provides an upper bound on the number of mth-order correlation-immune functions by using the bound of Park, Lee, Sung and Kim. However, the formula given by Maitra and Sarkar cannot be computed and thus their bound cannot be compared to the other bound.

Schneider proposed a new idea in 1990 for obtaining an upper bound on the number of mth-order correlation-immune Boolean functions, and an upper bound on m-resilient Boolean functions. In [10], he presented an algorithm for producing all correlation-immune functions. This algorithm is not very efficient (the workfactor, if computed, could be comparable to the complexity of searching among all Boolean functions). But the idea of this algorithm allowed him to provide an enumeration which is quite efficient.

Theorem 3. [10] *The number of m-resilient Boolean functions in n variables is less than:*

$$\prod_{i=1}^{n-m} \binom{2^i}{2^{i-1}}^{\binom{n-i-1}{m-1}} .$$

We can compare these three bounds by giving values in the 1-resilient case. It can be observed that Schneider's bound is always better than Yang-Guo's and Park *et al.*'s bounds for $n > 4$. The case $n = 3$ can be explained: the number

Table 1. Values of previous upper bounds for first order resilient functions

n	YG (Resilient)	PLSK (Resilient)	Schneider
3	18	8	12
4	1810	648	840
5	$4.4916\ 10^{7}$	$1.1979\ 10^{7}$	$1.081\ 10^{7}$
6	$7.0667\ 10^{16}$	$1.3711\ 10^{16}$	$6.498\ 10^{15}$
7	$4.6909\ 10^{35}$	$6.5259\ 10^{34}$	$1.191\ 10^{34}$
8	$5.6935\ 10^{73}$	$5.6396\ 10^{72}$	$2.8523\ 10^{71}$

of balanced Boolean functions such that $f(x) \oplus x_i$ is balanced for three distinct values of i is then exactly the number of 1-resilient functions.

Carlet and Klapper obtained two bounds on the number of m-resilient functions, one for $2 \leq m < n/2$ and the other one for $n/2 \leq m < n$. They improved upon Schneider's bound for m large.

Theorem 4. [1] *The number of m-resilient Boolean functions in n variables, $n/2 \leq m < n$, is less than:*

$$\frac{2^{1+\sum_{i=0}^{n-m-1}\binom{n}{i}}(1+\varepsilon)}{2^{\sum_{i=0}^{n-m-1}\binom{m-1}{i}}} + 2^{\sum_{i=0}^{n-m-2}\binom{n}{i}},$$

where $\varepsilon = \frac{1}{2^{\Omega((2^n/n)^{1/2})}}$.
The number of m-resilient Boolean functions in n variables, $2 \leq m < n/2$, is less than:

$$\frac{2^{\sum_{i=0}^{n-m-1}\binom{n}{i}} - 2^{\sum_{i=0}^{n-m-2}\binom{n}{i}}}{2^{2^{2m+1}-1}} + 2^{\sum_{i=0}^{n-m-2}\binom{n}{i}} \ .$$

4 A New Bound for m-Resilient Functions

Our improvement of Schneider's bound is based on several ideas. One of them is to use more efficiently than Schneider does the bound on the degrees of m-resilient Boolean functions. Recall that, thanks to Siegenthaler's theorem, we know that, for $m < n-1$, the degree of an m-resilient function in n variables is less than or equal to $n-m-1$.

Lemma 1. *Let f be a Boolean function in n variables. If the algebraic degree of f is at most d, then f is completely determined by its values at the words $u \in \mathbb{F}_2^n$ such that $w_H(u) \leq d$.*

Proof. Consider the algebraic normal form of the function:

$$f(x) = \bigoplus_{u \in \mathbb{F}_2^n} g(u) x^u,$$

where g is the Möbius transform of f. For every word u such that $d < w_H(u) \leq n$, the coefficient $g(u)$ is equal to zero, and thus:

$$\begin{aligned} f(x) &= \textstyle\bigoplus_{u \in \mathbb{F}_2^n \mid w_H(u) \leq d} g(u)\, x^u \\ &= \textstyle\bigoplus_{u \in \mathbb{F}_2^n \mid w_H(u) \leq d, u \preceq x} g(u) \\ &= \textstyle\bigoplus_{u \in \mathbb{F}_2^n \mid w_H(u) \leq d, u \preceq x} \left(\bigoplus_{v \in \mathbb{F}_2^n \mid v \preceq u} f(v) \right) \ . \end{aligned}$$

Every v such that $v \preceq u$ where $w_H(u) \leq d$ has weight at most d. □

The number of Boolean functions of degrees less than $n-m-1$ being negligible in comparison with Schneider's bound, we shall bound the number of m-resilient functions of degree exactly $n-m-1$ and add the number of Boolean functions of degrees less than $n-m-1$. To this aim, we shall use a lemma which

was first proved in [1]. But we shall need a slightly different statement of this lemma, with extra precisions that will be useful in our context. For this reason, we give a proof of the lemma. We first introduce a notation:
Let u and v be two vectors in $\mathbb{F}_2^n$; we denote by $v \wedge u$ the vector such that, for every index i, $(v \wedge u)_i = v_i\, u_i = \min(v_i, u_i)$, *i.e.*, and by $v \vee u$ the vector such that, for every index i, $(v \vee u)_i = \max(v_i, u_i)$ (these two operations are called bitwise-AND and bitwise-OR).

Lemma 2. [1] *Let* $f(x) = \bigoplus_{u \in \mathbb{F}_2^n} a_u\, x^u$ *be an* m*-resilient Boolean function in* n *variables of degree* $n - m - 1 \geq 2$ *with* $m \geq 2$ *and let* $\bigoplus_{u \in \mathbb{F}_2^n} b_u\, x^u$ *be the ANF of the function* $f(x) \oplus x_1 \oplus \cdots \oplus x_n$ *(i.e.* $b_u = a_u$ *if* $w_H(u) > 1$ *or if* $u = 0$ *and* $b_u = a_u \oplus 1$ *if* $w_H(u) = 1$*). If* u *is a word in* $\mathbb{F}_2^n$ *of weight* $n - m - 1$ *such that* $a_u = 1$ *(i.e.* $b_u = 1$*) then for all non-zero* v *in* $\mathbb{F}_2^n$ *such that* $v \wedge u = 0$*, we have:*

$$b_v = \bigoplus_{\substack{s \vee t = u \vee v \\ s \wedge u \neq 0 \\ t \wedge u \neq 0}} b_s\, b_t \quad .$$

Proof. We know (cf. [1]) that for every word x such that $w_H(x) \geq n - m$, we have that $\displaystyle\bigoplus_{\{s,t\}\,|s \vee t = x} b_s\; b_t = 0$. We apply this to $x = v \vee u$. In the corresponding relation, the coefficient b_v appears with a non-zero coefficient only in the term $b_u\, b_v$ since if $b_{u'}\, b_v$ appears, then $u' \vee v = x$, so $u \preceq u'$. We deduce:

$$b_v = \bigoplus_{\substack{s \vee t = u \vee v \\ s,t \neq v}} b_s\, b_t \quad .$$

According to Siegenthaler's inequality, the double condition that $s \vee t = u \vee v$ and $s, t \neq v$ implies, if $b_s \neq 0$ and $b_t \neq 0$, that $s \wedge u \neq 0$ and $t \wedge u \neq 0$ since u has weight $n - m - 1$. □

Theorem 5. *Let* n *and* m *be two positive integers such that* $\frac{n}{2} - 1 \leq m < n - 2$*. The number of* m*-resilient functions of degree* $n - m - 1$ *in* n *variables is lower than:*

$$\frac{\binom{n}{n-m-1}}{2^{\binom{m+1}{n-m-1}+1}} \prod_{i=1}^{n-m} \binom{2^i}{2^{i-1}}^{\binom{n-i-1}{m-1}} \quad .$$

Thus, the number of m*-resilient functions in* n *variables is lower than:*

$$2^{\sum_{i=0}^{n-m-2} \binom{n}{i}} + \frac{\binom{n}{n-m-1}}{2^{\binom{m+1}{n-m-1}+1}} \prod_{i=1}^{n-m} \binom{2^i}{2^{i-1}}^{\binom{n-i-1}{m-1}} \quad .$$

The principle of the proof is to bound the number of different truth-tables of m-resilient functions of maximum degree ($d = n - m - 1$) by using the fact that some of their successive restrictions are balanced. The bound is then obtained by adding the number of Boolean functions of degrees at most $n - m - 2$ (which is negligible).

Proof. According to Siegenthaler's Theorem on the degrees of resilient functions and according to Lemma 1, we only need, when evaluating the number of possible truth tables of f (that is the number of choices of the values of f at words $u \in \mathbb{F}_2^n$) to consider the words u such that $0 \leq w_H(u) \leq n-m-1$. In order to bound the number of m-resilient functions of degree exactly $n-m-1$, we first bound the number of m-resilient functions whose ANF contains the monomial $x_1 \ldots x_{n-m-1}$. We proceed by induction.

- step 1: Every m-resilient Boolean function f is such that the restricted function $f(x_1, \ldots, x_{n-m}, 0, \ldots, 0)$ is balanced, *i.e.*, has weight 2^{n-m-1}. Since the monomial $x_1 \ldots x_{n-m-1}$ appears in the ANF of the function, the number of words of the support which are less than $u = 1^{n-m-1}0^{m+1}$ for the partial order is odd. Consequently, there are
$$\sum_{i \text{ odd}} \binom{2^{n-m-1}}{i} \binom{2^{n-m-1}}{2^{n-m-1}-i} = \frac{1}{2} \binom{2^{n-m}}{2^{n-m-1}}$$
different choices for the restriction of the truth-table of f at words of $\{0,1\}^{n-m} \times \{0\}^m$.
- step 2: We now consider the restrictions of f in which the $(n-m)$th variable is fixed to zero. For the values of the variables $x_{n-m+1}, \ldots, x_n$, we fix $m-1$ variables among m to zero (there are m possible different choices), and the last free one is fixed to 1 because the cases where it is fixed to 0 have already been considered at the previous step. Indeed every word v lower (for the partial order $\preceq$) than the word $u = 1^{n-m}0^m$ has been considered at the first step and, *a fortiori*, every word lower than $u' = 1^{n-m-1}0^{m+1}$ has already been considered. Thus only the words in $\{u \in \mathbb{F}_2^n | u = (u_1, \ldots, u_{n-m-1}, 0 \ldots, 0, 1, 0, \ldots, 0)\}$ will be given a value by f at this step. We do not know how many words in this set must be in the support of the considered functions since we do not know how many words in the set $\{u \in \mathbb{F}_2^n | u = (u_1, \ldots, u_{n-m-1}, 0, \ldots, 0)\}$ are already in the support. But if this latter number is i, then the former one must be $j = 2^{n-m-1} - i$. And we know that for every j we have:
$$\binom{2^{n-m-1}}{j} \leq \binom{2^{n-m-1}}{2^{n-m-2}} \ .$$
We can bound the number of choices for one such restriction by $\binom{2^{n-m-1}}{2^{n-m-2}}$, and since the number of such restrictions is m, the number of choices $\binom{2^{n-m-1}}{2^{n-m-2}}$ is raised to the mth power. At the end of this step, we have considered all the words in $\mathbb{F}_2^n$ such that $0 \leq w_H(x_{n-m}, x_{n-m+1}, \ldots, x_n) \leq 1$.
- step p: Assume we have already chosen the values on the words x such that $0 \leq w_H(x_{n-m-p+3}, \ldots, x_n) \leq p-2$.
We now consider the restrictions such that $x_{n-m-p+2} = 0$, and $m-1$ variables among $m+p-2$ are fixed to 0; the remaining free variables are fixed to 1 because the other cases have already been considered in the previous steps. Thus there are

$\binom{m+p-2}{m-1}$ such restrictions. For each restriction, we do not know exactly how many words should be in the support, but this number can be bounded by the maximum possible number of choices, *i.e.*, $\binom{2^{n-m-p+1}}{2^{n-m-p}}$. Since there are $\binom{m+p-2}{m-1}$ such restrictions, the number of choices $\binom{2^{n-m-p+1}}{2^{n-m-p}}$ is raised to the power $\binom{m+p-2}{m-1}$. We show now that, at the end of this step, we have considered all the words x such that $0 \le w_H(x_{n-m-p+2}, \dots, x_n) \le p-1$: if $w_H(x_{n-m-p+2}, \dots, x_n) \le p-2$ or if $w_H(x_{n-m-p+2}, \dots, x_n) = p-1$ and $x_{n-m-p+2} = 1$, then x has been considered before step p (by induction hypothesis); and if $w_H(x_{n-m-p+2}, \dots, x_n) = p-1$ and $x_{n-m-p+2} = 0$, then it has been considered at step p.

- step $n-m$: According to the property proved above, all the words such that $w_H(x_3, \dots, x_n) \le n-m-2$ have been considered at the end of step $n-m-1$. Thus, only the words of weight $n-m-1$ and such that $x_1 = x_2 = 0$ have still to be given a value by f. We first choose a value $f(x)$ for every word $x = (0, 0, x_3, \dots, x_n)$ of weight $n-m-1$ such that $x \wedge u \ne 0$, where $u = 1^{n-m-1}0^{m+1}$. The number of such choices equals $2^{\binom{n-2}{n-m-1} - \binom{m+1}{n-m-1}}$. We apply now Lemma 2 to any word v of weight $n-m-1$ and such that $v \wedge u = 0$. We deduce the value of b_v and thus of a_v. Indeed, according to relation (1), the values of all the bits b_s, b_t such that $s \vee t = u \vee v$ and $s \wedge u \ne 0$, $t \wedge u \ne 0$ can be deduced from the values of $f(x)$ already chosen since $x \preceq s$ implies that either $w_H(x) < n-m-1$ or $x \wedge u \ne 0$. The knowledge of b_v implies that of $f(v)$ because all the values $f(x)$ such that $x \prec v$ have been already chosen, and according to relation (1).

We have now proved that the number of m-resilient functions f of degree $n-m-1$ and whose ANF contains the monomial $x_1 \dots x_{n-m-1}$ is upper bounded by:

$$\frac{1}{2^{\binom{m+1}{n-m-1}+1}} \prod_{i=1}^{n-m} \binom{2^i}{2^{i-1}}^{\binom{n-i-1}{m-1}} .$$

This number does not change if we replace the monomial $x_1 \dots x_{n-m-1}$ by any other monomial μ of same degree (since the notion of resiliency is invariant under the permutation of the coordinates of x). Any m-resilient function of degree $n-m-1$ belonging to $\bigcup_\mu S_\mu$, where S_μ is the set of all m-resilient functions of degree $n-m-1$ whose ANF contains μ, we obtain a bound on the number of m-resilient functions of degree $n-m-1$ by multiplying the number above by the number of these monomials, *i.e.*, $\binom{n}{n-m-1}$. Our bound on the number of all m-resilient functions is then obtained by adding the number of Boolean functions of degrees at most $n-m-2$. □

We now give tables of values permitting to compare the bounds. We give in the first table the values of the new bound for $\lceil \frac{n}{2} \rceil \le m \le \lceil \frac{n}{2} \rceil + 5$. In the next table, we compare Schneider's bound and the new bound (which improves upon it for $m \ge \lceil \frac{n}{2} \rceil$). In the last table of values, we compare Carlet-Klapper's bound and the new one.

Table 2. New bound on the number of m-resilient functions

$n\backslash m$	$\lceil \frac{n}{2} \rceil$	$\lceil \frac{n}{2} \rceil+1$	$\lceil \frac{n}{2} \rceil+2$	$\lceil \frac{n}{2} \rceil+3$	$\lceil \frac{n}{2} \rceil+4$	$\lceil \frac{n}{2} \rceil+5$
6	$1.1\ 10^{5}$	11	–	–	–	–
7	$9.5\ 10^{5}$	12	–	–	–	–
8	$5.36\ 10^{23}$	$7.6\ 10^{6}$	14	–	–	–
9	$1.4\ 10^{31}$	$5.9\ 10^{7}$	15	–	–	–
10	$6.5\ 10^{102}$	$4.2\ 10^{39}$	$4.4\ 10^{8}$	17	–	–
11	$2.3\ 10^{145}$	$1.4\ 10^{49}$	$3.2\ 10^{9}$	18	–	–
12	$5.6\ 10^{430}$	$1.3\ 10^{199}$	$5.8\ 10^{59}$	$2.3\ 10^{10}$	20	–
13	$1.6\ 10^{638}$	$2.6\ 10^{265}$	$2.7\ 10^{71}$	$1.6\ 10^{11}$	21	–
14	$1.3\ 10^{1776}$	$1.3\ 10^{918}$	$4.8\ 10^{345}$	$1.5\ 10^{84}$	$1.2\ 10^{12}$	23
15	$3.4\ 10^{2712}$	$3.7\ 10^{1286}$	$1.9\ 10^{441}$	$9.7\ 10^{97}$	$8.0\ 10^{12}$	47
16	$3.8\ 10^{7264}$	$1.2\ 10^{4034}$	$2.0\ 10^{1761}$	$3.9\ 10^{553}$	$7.5\ 10^{112}$	$5.5\ 10^{13}$
17	$1.6\ 10^{11333}$	$2.7\ 10^{5855}$	$5.7\ 10^{2361}$	$1.0\ 10^{684}$	$6.8\ 10^{128}$	$3.7\ 10^{14}$
18	$7.6\ 10^{29577}$	$8.2\ 10^{17260}$	$1.6\ 10^{8313}$	$1.1\ 10^{3109}$	$7.8\ 10^{833}$	$7.3\ 10^{145}$
19	$1.1\ 10^{46898}$	$2.8\ 10^{25709}$	$6.5\ 10^{11567}$	$9.8\ 10^{4025}$	$4.4\ 10^{1004}$	$9.4\ 10^{163}$
20	$7.8\ 10^{120074}$	$2.4\ 10^{72742}$	$5.3\ 10^{37511}$	$2.7\ 10^{15805}$	$1.1\ 10^{5137}$	$4.3\ 10^{1197}$
21	$1.2\ 10^{192912}$	$7.7\ 10^{110527}$	$3.0\ 10^{53700}$	$1.2\ 10^{21240}$	$2.5\ 10^{6468}$	$1.8\ 10^{1414}$

Table 3. (Schneider's bound/new bound) for m-resilient functions

$n\backslash m$	$\lceil \frac{n}{2} \rceil$	$\lceil \frac{n}{2} \rceil+1$	$\lceil \frac{n}{2} \rceil+2$	$\lceil \frac{n}{2} \rceil+3$	$\lceil \frac{n}{2} \rceil+4$	$\lceil \frac{n}{2} \rceil+5$
6	8.5	8.7	–	–	–	–
7	$9.8\ 10^{1}$	$1.5\ 10^{1}$	–	–	–	–
8	$3.7\ 10^{1}$	$2.3\ 10^{3}$	$2.7\ 10^{1}$	–	–	–
9	$2.5\ 10^{4}$	$1.2\ 10^{5}$	$5.0\ 10^{1}$		–	–
10	$3.1\ 10^{2}$	$5.7\ 10^{8}$	$1.2\ 10^{7}$	$9.0\ 10^{1}$	–	–
11	$2.1\ 10^{8}$	$8.7\ 10^{14}$	$2.5\ 10^{9}$	$1.7\ 10^{2}$	–	–
12	$5.3\ 10^{3}$	$4.8\ 10^{18}$	$1.8\ 10^{23}$	$1.1\ 10^{12}$	$3.1\ 10^{2}$	–
13	$1.1\ 10^{14}$	$2.4\ 10^{35}$	$9.3\ 10^{33}$	$9.2\ 10^{14}$	$5.7\ 10^{2}$	–
14	$1.8\ 10^{5}$	$8.5\ 10^{34}$	$3.3\ 10^{60}$	$2.6\ 10^{47}$	$1.6\ 10^{18}$	$1.1\ 10^{3}$
15	$7.7\ 10^{21}$	$4.8\ 10^{72}$	$3.2\ 10^{96}$	$7.4\ 10^{63}$	$5.7\ 10^{21}$	$2.0\ 10^{3}$
16	$1.2\ 10^{7}$	$4.1\ 10^{59}$	$5.4\ 10^{135}$	$1.1\ 10^{146}$	$4.4\ 10^{83}$	$4.1\ 10^{25}$
17	$1.3\ 10^{32}$	$1.9\ 10^{135}$	$8.4\ 10^{234}$	$1.4\ 10^{212}$	$1.1\ 10^{107}$	$6.0\ 10^{29}$
18	$1.6\ 10^{9}$	$1.4\ 10^{95}$	$1.5\ 10^{274}$	$6.2\ 10^{383}$	$1.4\ 10^{298}$	$2.3\ 10^{134}$
19	$1.2\ 10^{45}$	$1.0\ 10^{234}$	$2.7\ 10^{512}$	$7.9\ 10^{598}$	$4.2\ 10^{407}$	$7.8\ 10^{165}$
20	$4.3\ 10^{11}$	$1.6\ 10^{144}$	$9.5\ 10^{511}$	$5.1\ 10^{899}$	$1.3\ 10^{900}$	$3.1\ 10^{544}$
21	$1.1\ 10^{61}$	$2.6\ 10^{382}$	$2.3\ 10^{1028}$	$1.7\ 10^{1503}$	$7.8\ 10^{1310}$	$9.4\ 10^{712}$

Table 4. (Carlet-Klapper's bound/new bound) for m-resilient functions

$n\backslash m$	$\lceil\frac{n}{2}\rceil$	$\lceil\frac{n}{2}\rceil+1$	$\lceil\frac{n}{2}\rceil+2$	$\lceil\frac{n}{2}\rceil+3$	$\lceil\frac{n}{2}\rceil+4$	$\lceil\frac{n}{2}\rceil+5$
6	$5.5\ 10^{-2}$	1.27	–	–	–	–
7	$2.6\ 10^{-2}$	1.12	–	–	–	–
8	$5.5\ 10^{-4}$	$1.2\ 10^{-2}$	1	–	–	–
9	$4.4\ 10^{-5}$	$6.7\ 10^{-3}$	$9.0\ 10^{-1}$		–	–
10	$3.3\ 10^{-4}$	$2.4\ 10^{-6}$	$3.6\ 10^{-3}$	$8.2\ 10^{-1}$	–	–
11	$2.0\ 10^{-6}$	$9.7\ 10^{-8}$	$1.9\ 10^{-3}$	$7.6\ 10^{-1}$	–	–
12	$7.3\ 10^{10}$	$1.4\ 10^{-9}$	$2.6\ 10^{-9}$	$1.1\ 10^{-3}$	$7.0\ 10^{-1}$	–
13	$4.1\ 10^{12}$	$7.0\ 10^{-14}$	$4.7\ 10^{-11}$	$6.1\ 10^{-4}$	$6.5\ 10^{-1}$	–
14	$2.0\ 10^{99}$	$4.5\ 10^{12}$	$1.7\ 10^{-19}$	$5.9\ 10^{-13}$	$3.5\ 10^{-4}$	$6.1\ 10^{-1}$
15	$2.7\ 10^{142}$	$9.1\ 10^{9}$	$1.3\ 10^{-26}$	$5.0\ 10^{-15}$	$2.0\ 10^{-4}$	$5.7\ 10^{-1}$
16	$4.2\ 10^{511}$	$1.6\ 10^{194}$	$2.2\ 10^{3}$	$2.4\ 10^{-35}$	$2.9\ 10^{-17}$	$1.2\ 10^{-4}$
17	$9.3\ 10^{785}$	$2.3\ 10^{253}$	$2.5\ 10^{-9}$	$6.1\ 10^{-46}$	$1.1\ 10^{-19}$	$6.9\ 10^{-5}$
18	$1.1\ 10^{2256}$	$6.4\ 10^{1158}$	$4.1\ 10^{317}$	$3.9\ 10^{-28}$	$1.4\ 10^{-58}$	$2.9\ 10^{-22}$
19	$3.7\ 10^{3610}$	$1.2\ 10^{1649}$	$2.4\ 10^{383}$	$1.6\ 10^{-55}$	$2.2\ 10^{-73}$	$5.1\ 10^{-25}$
20	$2.3\ 10^{9330}$	$5.9\ 10^{5571}$	$1.3\ 10^{2275}$	$4.7\ 10^{445}$	$2.0\ 10^{-93}$	$1.5\ 10^{-90}$
21	$5.4\ 10^{15353}$	$2.5\ 10^{8328}$	$4.1\ 10^{3053}$	$8.1\ 10^{497}$	$7.1\ 10^{-144}$	$2.6\ 10^{-110}$
22	$3.7\ 10^{37456}$	$1.5\ 10^{24442}$	$5.2\ 10^{12102}$	$2.0\ 10^{3998}$	$5.2\ 10^{532}$	$3.6\ 10^{-209}$

Remark 1. A slight improvement of our bound is possible: let k be a positive integer; the number of Boolean functions of degree at most $n-m-1$ and whose ANF contains at most $k-1$ monomials of degree $n-m-1$ equals $2^{\sum_{i=0}^{n-m-2}\binom{n}{i}}\left(\sum_{j=0}^{k-1}\binom{\binom{n}{n-m-1}}{j}\right)$. We deduce that the number of m-resilient functions in n variables is lower than:

$$2^{\sum_{i=0}^{n-m-2}\binom{n}{i}}\left(\sum_{j=0}^{k-1}\binom{\binom{n}{n-m-1}}{j}\right)+\frac{\binom{n}{n-m-1}}{k\,2^{\binom{m+1}{n-m-1}+1}}\prod_{i=1}^{n-m}\binom{2^i}{2^{i-1}}^{\binom{n-i-1}{m-1}}.$$

We have checked that for almost every n, some values of $k\leq\binom{n}{n-m-1}$ permit to improve upon our bound.

5 Conclusion

We have obtained for $m\geq\frac{n}{2}$ an improvement of Schneider's bound on the number of m-resilient functions in n variables. The tables computed show that our bound also partially improves upon Carlet-Klapper's bound. Notice that the values of m for which this happens in the tables are those among which the best satisfactory tradeoffs between resiliency order, nonlinearity (limited by Sarkar-Maitra's bound [9]) and degree (limited by Siegenthaler's bound) can be obtained (since none of these parameters must be small). Moreover, we can conjecture that, asymptotically, the new bound improves upon Carlet-Klapper's bound when $m-n/2$ is fixed and n tends to infinity (recall that Carlet-Klapper's bound improves upon Schneider's one when $n-m$ is fixed and n tends to infinity).

References

1. Carlet, C., Klapper, A.: Upper bounds on the number of resilient functions and of bent functions. Proceedings of 23rd Symposium on Information Theory in the Benelux (2002)
2. Chor, B., Goldreich, O., Hastad, J., Rudich, S., Smolensky, R.: The bit extraction problem or t-resilient functions. In 26th IEEE Symposium on Foundations of Computer Science (1985) 396–407
3. Denisov, O.: An asymptotic formula for the number of correlation-immune of order k Boolean functions. Discrete Mathematics and Applications **2 (4)** (1992) 407–426
4. Guo-Zhen, X., Massey, J. L.: A spectral characterization of correlation-immune combining functions. IEEE Transaction on Information Theory **34 (3)** (1988) 569–571
5. Maitra, S., Sarkar, P.: Enumeration of correlation-immune Boolean functions. ACISP (1999) 12–25
6. Massey, J. L.: Shift-register synthesis and BCH decoding. IEEE Transactions on Information Theory **15** (1969) 122–127
7. Mitchell, C. J.: Enumerating Boolean functions of cryptographic signifiance. Journal of Cryptology **2 (3)** (1990) 155–170
8. Park, S. M., Lee, S., Sung, S. H., Kim, K.: Improving bounds for the number of correlation-immune Boolean functions. Information Processing Letters **61** (1997) 209–212
9. Sarkar, P., Maitra, S.: Nonlinearity bounds and constructions of resilient Boolean functions. Crypto 2000 (2000) 515–532
10. Schneider, M.: A note on the construction and upper bounds of correlation-immune functions. 6th IMA Conference (1997) 295–306
11. Siegenthaler, T.: Correlation-immunity of nonlinear combining functions for cryptographic applications. IEEE Transactions on Information Theory **30 (5)** (1984) 776–780
12. Siegenthaler, T.: Decrypting a class of stream cipher using ciphertext only. IEEE Transactions on Computers **34 (1)** (1985) 81–85
13. Vernam, G.: Cipher printing telegraph systems for secret wire and radio telegraphic communication. Journal of the American Institute of Electrical Engineers **45** (1926) 109–115
14. Yang, Y. X., Guo, B.: Further enumerating Boolean functions of cryptographic signifiance. Journal of Cryptology **8 (3)** (1995) 115–122

A Proof of Proposition 2

We prove that the number of 1-resilient functions in n variables is less than $\sum_{k=0}^{2^{n-2}} \binom{2^{n-2}}{a}^4$.
Every Boolean function f in n variables can be considered as the concatenation of four Boolean functions in $n-2$ variables, $f = f_1 f_2 f_3 f_4$. The ANF of the function is

$$f = (1-x_n)(1-x_{n-1})f_1 \oplus (1-x_n)x_{n-1}f_2 \oplus x_n(1-x_{n-1})f_3 \oplus x_n x_{n-1} f_4 \ .$$

We have:

$$w_H(f) = 2^{n-1} \Leftrightarrow w_H(f_1) + w_H(f_2) + w_H(f_3) + w_H(f_4) = 2^{n-1} \tag{2}$$

$$w_H(f_{|x_n=0}) = 2^{n-2} \Leftrightarrow w_H(f_1) + w_H(f_2) = 2^{n-2} \tag{3}$$

$$w_H(f_{|x_{n-1}=0}) = 2^{n-2} \Leftrightarrow w_H(f_1) + w_H(f_3) = 2^{n-2} \tag{4}$$

Thus,

$$(3), (4) \ \Rightarrow w_H(f_2) = w_H(f_3) \tag{5}$$

$$(2), (3), (5) \ \Rightarrow w_H(f_1) = w_H(f_4) \tag{6}$$

The bound of Proposition 2 is then a direct consequence of equations (3), (5) and (6). Indeed, we can deduce:

$$\sum_{w_H(f_1)=0}^{2^{n-2}} \binom{2^{n-2}}{w_H(f_1)}^2 \binom{2^{n-2}}{2^{n-2} - w_H(f_1)}^2 .$$

□

B Park, Lee Sung and Kim's bound in the case of first order resilient Boolean function

Proposition 3. *Let n be a positive integer greater than 1. The number of 1-resilient Boolean functions in n variables is less than:*

$$\sum_{a,b,c,d=0}^{2^{n-3}} \binom{2^{n-3}}{a}\binom{2^{n-3}}{b}\binom{2^{n-3}}{c}\binom{2^{n-3}}{d}\binom{2^{n-3}}{2^{n-2}-a-b-c}$$
$$\times \binom{2^{n-3}}{2^{n-2}-a-c-d}\binom{2^{n-3}}{c+d-b}\binom{2^{n-3}}{a+b-d} .$$

Proof. Every Boolean function in n variables f can be considered as the concatenation of eight functions in $n-3$ variables, *i.e.*, $f = f_1 f_2 f_3 f_4 f_5 f_6 f_7 f_8$. The corresponding ANF of the function is

$$\begin{aligned} f = {} & (1-x_n)(1-x_{n-1})(1-x_{n-2})f_1 \oplus (1-x_n)(1-x_{n-1})x_{n-2}f_2 \\ & \oplus (1-x_n)x_{n-1}(1-x_{n-2})f_3 \oplus (1-x_n)x_{n-1}x_{n-2}f_4 \\ & \oplus x_n(1-x_{n-1})(1-x_{n-2})f_5 \oplus x_n(1-x_{n-1})x_{n-2}f_6 \\ & \oplus x_n x_{n-1}(1-x_{n-2})f_7 \oplus x_n x_{n-1} x_{n-2} f_8 \ . \end{aligned}$$

We have the following equations:

$$w_H(f) = 2^{n-1} \Leftrightarrow \sum_{i=1}^{8} w_H(f_i) = 2^{n-1} \tag{7}$$

$$w_H(f_{|x_n=0}) = 2^{n-2} \Leftrightarrow w_H(f_1) + w_H(f_2) + w_H(f_3) + w_H(f_4) = 2^{n-2} \quad (8)$$

$$w_H(f_{|x_{n-1}=0}) = 2^{n-2} \Leftrightarrow w_H(f_1) + w_H(f_2) + w_H(f_5) + w_H(f_6) = 2^{n-2} \quad (9)$$

$$w_H(f_{|x_{n-2}=0}) = 2^{n-2} \Leftrightarrow w_H(f_1) + w_H(f_3) + w_H(f_5) + w_H(f_7) = 2^{n-2} \quad (10)$$

We obtain:

$$(8),(9) \Rightarrow w_H(f_3) + w_H(f_4) = w_H(f_5) + w_H(f_6) \quad (11)$$

$$(7),(11) \Rightarrow w_H(f_1) + w_H(f_2) = w_H(f_7) + w_H(f_8) \quad (12)$$

Assume that the values of $w_H(f_1)$, $w_H(f_2)$, $w_H(f_3)$ and $w_H(f_7)$ are fixed, then

$$(8) \Rightarrow w_H(f_4) = 2^{n-2} - w_H(f_1) - w_H(f_2) - w_H(f_3) \quad (13)$$

$$(10) \Rightarrow w_H(f_5) = 2^{n-2} - w_H(f_1) - w_H(f_3) - w_H(f_7) \quad (14)$$

$$(9),(14) \Rightarrow w_H(f_6) = w_H(f_3) + w_H(f_7) - w_H(f_2) \quad (15)$$

$$(12) \Rightarrow w_H(f_8) = w_H(f_1) + w_H(f_2) - w_H(f_7) \quad (16)$$

Since we know that the values of $w_H(f_1)$, $w_H(f_2)$, $w_H(f_3)$ and $w_H(f_7)$ vary between 0 and 2^{n-3}, we can deduce the formula with the equations (13), (14), (15) and (16). $\square$