On the nonlinearity of monotone Boolean functions

Claude Carlet

LAGA, Department of Mathematics, University of Paris 8
(and Paris 13 and CNRS), Saint–Denis cedex 02, France.
E-mail: claude.carlet@univ-paris8.fr

Abstract. We first prove the truthfulness of a conjecture on the nonlinearity of monotone Boolean functions in even dimension, proposed in the recent paper "Cryptographic properties of monotone Boolean functions", by D. Joyner, P. Stanica, D. Tang and the author, to appear in the Journal of Mathematical Cryptology. We prove then an upper bound on such nonlinearity, which is asymptotically much stronger than the conjectured upper bound and than the upper bound proved for odd dimension in this same paper. This bound shows a deep weakness of monotone Boolean functions; they are too closely approximated by affine functions for being usable as nonlinear components in cryptographic applications. We deduce a necessary criterion to be satisfied by a Boolean (resp. vectorial) function for being nonlinear.

Keywords: Boolean functions, monotone functions, Walsh–Hadamard spectrum.

1 Introduction

The present paper is a continuation (and deepening) of the main results of [9]. We shall then use the same notation as in [9]. A function from the n-dimensional vector space $\mathbb{F}_2^n$ over $\mathbb{F}_2$, and valued in $\mathbb{F}_2$ is called a Boolean function on n variables. The set of all Boolean functions on n variables is denoted by $\mathcal{B}_n$. Any element $\mathbf{x} \in \mathbb{F}_2^n$ is an n-tuple $(x_1, \ldots, x_n)$, where $x_i \in \mathbb{F}_2$ for all $i = 1, \ldots, n$. The *support* $\mathrm{supp}(\mathbf{x})$ is the set of all positions i where $x_i = 1$. The (*Hamming*) *weight* of $\mathbf{x} \in \mathbb{F}_2^n$ is the size $\sum_{i=1}^n x_i$ of its support and is denoted by $w_H(\mathbf{x})$. The Hamming weight $w_H(f)$ of a Boolean function f is the weight of its output vector, that is, the size of its support $\{\mathbf{x} \in \mathbb{F}_2^n;\ f(\mathbf{x}) = 1\}$. The additions over $\mathbb{F}_2$ and $\mathbb{F}_2^n$, are denoted by '+'. We denote the (vector) *complement* $(x_1 + 1, \ldots, x_n + 1)$ of $\mathbf{x}$ by $\overline{\mathbf{x}}$, and the (Boolean function) *complement* by $\overline{f}(\mathbf{x}) = f(\mathbf{x}) + 1$, for $f \in \mathcal{B}_n$. The cardinality of a set S is denoted by $|S|$.

For $\mathbf{x} = (x_1, \ldots, x_n), \mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{F}_2^n$, we consider the usual *inner product*:

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

For $\mathbf{u} = (u_1, \ldots, u_n), \mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{F}_2^n$, we define the partial order on $\mathbb{F}_2^n$:

$$\mathbf{u} \preceq \mathbf{v} \text{ if and only if } u_i \leq v_i, \text{ for every } i.$$

Any $f \in \mathcal{B}_n$ can be expressed in *algebraic normal form* (ANF) as

$$f(x_1, x_2, \ldots, x_n) = \sum_{\mathbf{a}=(a_1,\ldots,a_n) \in \mathbb{F}_2^n} \mu_{\mathbf{a}} \left(\prod_{i=1}^n x_i^{a_i} \right),$$

for some coefficients $\mu_{\mathbf{a}} = \mu_a(f) \in \mathbb{F}_2$. Any term $\prod_{i=1}^n x_i^{a_i}$ is called a *monomial*. The ANF of any Boolean function satisfies $\mu_{\mathbf{a}} = \sum_{\mathbf{x} \preceq \mathbf{a}} f(\mathbf{x})$ and $f(\mathbf{x}) = \sum_{\mathbf{a} \preceq \mathbf{x}} \mu_{\mathbf{a}}$.

The algebraic degree of f, $\deg(f)$, equals $\max_{\mathbf{a} \in \mathbb{F}_2^n} \{w_H(\mathbf{a}) \mid \mu_{\mathbf{a}} \neq 0\}$. Boolean functions having algebraic degree at most 1 are *affine functions*. For any two functions $f, g \in \mathcal{B}_n$, we define the (*Hamming*) *distance* $d_H(f, g) = w_H(f + g)$.

The (unnormalized) *Walsh–Hadamard transform* of $f \in \mathcal{B}_n$ at any point $\mathbf{u} \in \mathbb{F}_2^n$ is defined by

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

The multiset $[W_f(\mathbf{u}) \mid \mathbf{u} \in \mathbb{F}_2^n]$ is called the *Walsh–Hadamard spectrum* of function f. The *nonlinearity* of f is its distance from the set A_n of all n-variable affine functions:

$$nl(f) = \min_{g \in A_n} d_H(f, g) = 2^{n-1} - \frac{1}{2} \max_{\mathbf{u} \in \mathbb{F}_2^n} |W_f(\mathbf{u})|.$$

Function $f \in \mathcal{B}_n$ is called *bent* if its nonlinearity achieves the optimum $2^{n-1} - 2^{n/2-1}$.

A Boolean function f is *monotone* (increasing) if whenever $\mathbf{u} \preceq \mathbf{v}$, then $f(\mathbf{u}) \leq f(\mathbf{v})$. It is easy to see that any monomial Boolean function is monotone. Other examples are the *majority function* in odd dimension n, defined by $M(\mathbf{x}) = 1$ if $w_H(\mathbf{x}) > n/2$ and the strict (resp. large) *majority functions* in even dimension defined by $M(\mathbf{x}) = 1$ if $w_H(\mathbf{x}) > n/2$ (resp. $w_H(\mathbf{x}) \geq n/2$), and more generally the functions whose supports are the sets of vectors of Hamming weights bounded by some number from below.

Monotone Boolean functions have applications in voting theory, reliability theory, hypergraphs, learning, etc. (a non-exhaustive list of papers and monographies devoted to these connections is [2–4, 12, 15, 18]). So, it is natural to inquire about their cryptographic properties, as well.

In [9] are studied the balancedness, nonlinearity and algebraic immunity of monotone Boolean functions. It is shown that, for every even $n \geq 4$, there exists no n-variable monotone bent function and that, for n odd at least 5, every n variable monotone function has nonlinearity at most $2^{n-1} - 2^{\frac{n-1}{2}}$. It is also conjectured in [9] that, for n even sufficiently large, every n variable monotone function has nonlinearity at most $2^{n-1} - 2^{\frac{n}{2}}$, but the proof in this case is more complex and needs stronger tools as we shall see. In the present paper, we prove that the conjecture is true. Moreover, we prove another upper bound valid for every n; for n even, this bound is asymptotically much stronger than the conjectured upper bound; for n odd, it is also much larger than the upper bound $2^{n-1} - 2^{\frac{n-1}{2}}$.

2 On the nonlinearity of monotone functions

Let us first recall what is proved in [9]. Let f be any n-variable monotone Boolean function. For every $\mathbf{y} \in \mathbb{F}_2^n$ such that $f(\mathbf{y}) = 0$, we have, according to the Poisson

summation formula (which is recalled e.g. in [7]):

$$\sum_{\mathbf{u}\in\mathbb{F}_2^n;\,\mathbf{u}\preceq\overline{\mathbf{y}}} W_f(\mathbf{u}) = 2^n,$$

and this implies that $\max_{\mathbf{u}\in\mathbb{F}_2^n;\,\mathbf{u}\preceq\overline{\mathbf{y}}} |W_f(\mathbf{u})| \geq 2^{w_H(\mathbf{y})}$. Similarly, for every $\mathbf{y}$ such that $f(\mathbf{y}) = 1$, we have:

$$\sum_{\mathbf{u}\in\mathbb{F}_2^n;\,\mathbf{u}\preceq\mathbf{y}} (-1)^{(1,\ldots,1)\cdot\mathbf{u}} W_f(\mathbf{u}) = -2^n,$$

and this implies that $\max_{\mathbf{u}\in\mathbb{F}_2^n;\,\mathbf{u}\preceq\mathbf{y}} |W_f(\mathbf{u})| \geq 2^{n-w_H(\mathbf{y})}$. Since the majority function in odd dimension n has nonlinearity $2^{n-1} - \binom{n-1}{(n-1)/2}$, this allowed [9] to derive the upper bound:

Theorem 2.1 *For every odd $n \geq 5$ and every monotone n-variable function f, we have $nl(f) \leq 2^{n-1} - 2^{(n-1)/2}$.*

But no general upper bound for n even could be shown. Indeed, only the case where $f(\mathbf{x})$ differs from the majority function for at least one input $\mathbf{x}$ of Hamming weight different from $n/2$ can be easily handled thanks to the observations above. The case where $f(\mathbf{x})$ coincides with the majority function for every input $\mathbf{x}$ of Hamming weight different from $n/2$ must be handled by other means. Then [9] only conjectured the upper bound $nl(f) \leq 2^{n-1} - 2^{n/2}$ for n even large enough.

2.1 Proof of the conjecture

As explained above, we only need to handle the case, for n even, of those functions equal to the majority function at every input $\mathbf{x}$ of Hamming weight different from $n/2$. So let f be such function, that we can assume different from the strict and large majority functions, since the nonlinearity of these two functions, equal to $2^{n-1} - \binom{n-1}{n/2}$, is larger than $2^{n-1} - 2^{n/2}$ for an even number of variables large enough. We shall use the second-order Poisson summation formula[1], introduced in [5] and given as Equality (28) in [7]: given two supplementary vector subspaces E and E' in $\mathbb{F}_2^n$, we have:

$$\sum_{\mathbf{u}\in E^\perp} W_f^2(\mathbf{u}) = |E^\perp| \sum_{\mathbf{a}\in E'} \left(\sum_{\mathbf{x}\in E} (-1)^{f(\mathbf{a}+\mathbf{x})} \right)^2. \tag{2}$$

For a given $\mathbf{y}$ of Hamming weight $n/2$ and such that $f(\mathbf{y}) = 0$, let us take $E = \{\mathbf{x} \in \mathbb{F}_2^n;\, \mathbf{x} \preceq \mathbf{y}\}$. Then $E^\perp = \{\mathbf{u} \in \mathbb{F}_2^n;\, \mathbf{u} \preceq \overline{\mathbf{y}}\}$ is supplementary of E and we

[1] It is rare that this formula needs to be used rather than the simpler Poisson formula; it is interesting to find such situation (here and in the next section as well).

can then take $E' = E^{\perp}$; we obtain, f being null on E since it is monotone:

$$\sum_{\mathbf{u}\in\mathbb{F}_2^n;\,\mathbf{u}\preceq\overline{\mathbf{y}}} W_f^2(\mathbf{u}) = 2^{n/2}\sum_{\mathbf{a}\in\mathbb{F}_2^n;\mathbf{a}\preceq\overline{\mathbf{y}}}\left(\sum_{\mathbf{x}\in\mathbb{F}_2^n;\,\mathbf{x}\preceq\mathbf{y}}(-1)^{f(\mathbf{a}+\mathbf{x})}\right)^2$$
$$= 2^{3n/2} + 2^{n/2}\sum_{\mathbf{a}\in\mathbb{F}_2^n;\mathbf{a}\preceq\overline{\mathbf{y}};\mathbf{a}\neq 0}\left(\sum_{\mathbf{x}\in\mathbb{F}_2^n;\,\mathbf{x}\preceq\mathbf{y}}(-1)^{f(\mathbf{a}+\mathbf{x})}\right)^2.$$

Since the maximum of a sequence is at least equal to its mean, we deduce the inequality $\max_{\mathbf{u}\in\mathbb{F}_2^n;\,\mathbf{u}\preceq\overline{\mathbf{y}}} W_f^2(\mathbf{u}) \geq 2^n + \sum_{\mathbf{a}\preceq\overline{\mathbf{y}};\mathbf{a}\neq 0}\left(\sum_{\mathbf{x}\in\mathbb{F}_2^n;\,\mathbf{x}\preceq\mathbf{y}}(-1)^{f(\mathbf{a}+\mathbf{x})}\right)^2$.
We first show that, for every $\mathbf{a}\preceq\overline{\mathbf{y}}$ with $\mathbf{a}\neq 0$, say of Hamming weight $j>0$, the value of $\sum_{\mathbf{x}\in\mathbb{F}_2^n;\,\mathbf{x}\preceq\mathbf{y}}(-1)^{f(\mathbf{a}+\mathbf{x})}$ lies between $\sum_{i=0}^{n/2-1-j}\binom{n/2}{i} - \sum_{i=n/2+1-j}^{n/2}\binom{n/2}{i} - \binom{n/2}{n/2-j}$ and $\sum_{i=0}^{n/2-1-j}\binom{n/2}{i} - \sum_{i=n/2+1-j}^{n/2}\binom{n/2}{i} + \binom{n/2}{n/2-j}$. Indeed, if $\mathbf{x}\preceq\mathbf{y}$ has Hamming weight strictly less than $n/2-j$ then $\mathbf{a}+\mathbf{x}$ has Hamming weight strictly less than $n/2$ and $f(\mathbf{a}+\mathbf{x})$ equals 0, and if $\mathbf{x}\preceq\mathbf{y}$ has Hamming weight strictly larger than $n/2-j$ then $\mathbf{a}+\mathbf{x}$ has Hamming weight strictly larger than $n/2$ and $f(\mathbf{a}+\mathbf{x})$ equals 1. If $\mathbf{x}\preceq\mathbf{y}$ has Hamming weight $n/2-j$, then $\mathbf{a}+\mathbf{x}$ has weight $n/2$ and the value of $f(\mathbf{a}+\mathbf{x})$ is unknown.
We replace now $\binom{n/2}{i}$ by $\binom{n/2}{n/2-i}$ in the sum $\sum_{i=n/2+1-j}^{n/2}\binom{n/2}{i}$. We obtain $\sum_{i=0}^{j-1}\binom{n/2}{i}$. Then for $j<n/4$, $\left(\sum_{\mathbf{x}\in\mathbb{F}_2^n;\,\mathbf{x}\preceq\mathbf{y}}(-1)^{f(\mathbf{a}+\mathbf{x})}\right)^2 \geq \left(\sum_{i=j}^{n/2-1-j}\binom{n/2}{i} - \binom{n/2}{n/2-j}\right)^2 = \left(\sum_{i=j+1}^{n/2-1-j}\binom{n/2}{i}\right)^2$, since we have $n/2-1-j\geq j$, and if $j>n/4$, then we have $\left(\sum_{\mathbf{x}\in\mathbb{F}_2^n;\,\mathbf{x}\preceq\mathbf{y}}(-1)^{f(\mathbf{a}+\mathbf{x})}\right)^2 \geq \left(\sum_{i=n/2-j}^{j-1}\binom{n/2}{i} - \binom{n/2}{n/2-j}\right)^2 = \left(\sum_{i=n/2-j+1}^{j-1}\binom{n/2}{i}\right)^2$, since $j-1\geq n/2-j$. We then deduce that:

$$\max_{\mathbf{u}\in\mathbb{F}_2^n;\,\mathbf{u}\preceq\overline{\mathbf{y}}} W_f^2(\mathbf{u}) \geq$$

$$2^n + \sum_{1\leq j<n/4}\binom{n/2}{j}\left(\sum_{i=j+1}^{n/2-1-j}\binom{n/2}{i}\right)^2 + \sum_{n/4<j\leq n/2}\binom{n/2}{j}\left(\sum_{i=n/2-j+1}^{j-1}\binom{n/2}{i}\right)^2 =$$

$$2^n + 2\sum_{1\leq j<n/4}\binom{n/2}{j}\left(\sum_{i=j+1}^{n/2-1-j}\binom{n/2}{i}\right)^2 + \left(\sum_{i=1}^{n/2-1}\binom{n/2}{i}\right)^2 =$$

$$2^n + 2\sum_{1\leq j<n/4}\binom{n/2}{j}\left(2^{n/2} - 2\sum_{i=0}^{j}\binom{n/2}{i}\right)^2 + \left(2^{n/2}-2\right)^2.$$

It is easily seen that for every $n \geq 10$, we have $2\sum_{1\leq j<n/4}\binom{n/2}{j}\left(2^{n/2}-2\sum_{i=0}^{j}\binom{n/2}{i}\right)^2+\left(2^{n/2}-2\right)^2 \geq 3\cdot 2^n$. Indeed, the expression of n equal to

$$2^{-n}\left[2\sum_{1\leq j<n/4}\binom{n/2}{j}\left(2^{n/2}-2\sum_{i=0}^{j}\binom{n/2}{i}\right)^2+\left(2^{n/2}-2\right)^2\right]$$

is clearly non-decreasing and is larger than 3 for $n=10$. We deduce then:

Theorem 2.2 *For every even $n \geq 10$ and every monotone n-variable function f, we have $nl(f) \leq 2^{n-1}-2^{n/2}$.*

Observe that Theorem 2.1 provides an alternative proof, for $n \geq 10$, of the inexistence of monotone bent functions (originally proved in [9]).

2.2 A stronger bound, valid for every n

The inequalities $\max_{\mathbf{u}\in\mathbb{F}_2^n}|W_f(\mathbf{u})| \geq 2^{w_H(\mathbf{y})}$ for $\mathbf{y}\in\mathbb{F}_2^n$ such that $f(\mathbf{y})=0$ and $\max_{\mathbf{u}\in\mathbb{F}_2^n}|W_f(\mathbf{u})| \geq 2^{n-w_H(\mathbf{y})}$ for $\mathbf{y}\in\mathbb{F}_2^n$ such that $f(\mathbf{y})=1$ can be refined by using Equality (2) again. For every $\mathbf{y}\in\mathbb{F}_2^n$ such that $f(\mathbf{y})=0$, we have, since $f(\mathbf{x})$ is null on $E=\{\mathbf{x}\in\mathbb{F}_2^n;\ \mathbf{x}\preceq\mathbf{y}\}$:

$$\begin{aligned}\sum_{\mathbf{u}\in\mathbb{F}_2^n;\,\mathbf{u}\preceq\overline{\mathbf{y}}} W_f^2(\mathbf{u}) &= 2^{n-w_H(\mathbf{y})}\sum_{\mathbf{a}\in\mathbb{F}_2^n;\,\mathbf{a}\preceq\overline{\mathbf{y}}}\left(\sum_{\mathbf{x}\in\mathbb{F}_2^n;\,\mathbf{x}\preceq\mathbf{y}}(-1)^{f(\mathbf{a}+\mathbf{x})}\right)^2\\ &= 2^{n+w_H(\mathbf{y})}+2^{n-w_H(\mathbf{y})}\sum_{\mathbf{a}\in\mathbb{F}_2^n;\,\mathbf{a}\preceq\overline{\mathbf{y}};\mathbf{a}\neq 0}\left(\sum_{\mathbf{x}\in\mathbb{F}_2^n;\,\mathbf{x}\preceq\mathbf{y}}(-1)^{f(\mathbf{a}+\mathbf{x})}\right)^2,\end{aligned}$$

which implies

$$\max_{\mathbf{u}\in\mathbb{F}_2^n;\,\mathbf{u}\preceq\overline{\mathbf{y}}} W_f^2(\mathbf{u}) \geq 2^{2w_H(\mathbf{y})}+\sum_{\mathbf{a}\in\mathbb{F}_2^n;\,\mathbf{a}\preceq\overline{\mathbf{y}};\mathbf{a}\neq 0}\left(\sum_{\mathbf{x}\in\mathbb{F}_2^n;\,\mathbf{x}\preceq\mathbf{y}}(-1)^{f(\mathbf{a}+\mathbf{x})}\right)^2. \qquad (3)$$

1. If there exist vectors of Hamming weight strictly larger than $n/2$ whose image by f is null, let then $\mathbf{y}$ have maximal Hamming weight among all vectors satisfying $f(\mathbf{y})=0$ and denote this Hamming weight by w. For every $\mathbf{a}\preceq\overline{\mathbf{y}}$ (of Hamming weight $j\leq n-w$), we have $f(\mathbf{a}+\mathbf{x})=1$ for every $\mathbf{x}\preceq\mathbf{y}$ such that $\mathbf{a}+\mathbf{x}$ has Hamming weight at least $w+1$ (that is, for every $\mathbf{x}\preceq\mathbf{y}$ of Hamming weight at least $w-j+1$), and we deduce $\sum_{\mathbf{x}\in\mathbb{F}_2^n;\,\mathbf{x}\preceq\mathbf{y}}(-1)^{f(\mathbf{a}+\mathbf{x})} \leq 2^w-2\sum_{i=w-j+1}^{w}\binom{w}{i}$. Note that we have $2^w-2\sum_{i=w-j+1}^{w}\binom{w}{i}\leq 0$ if and only if $w-j+1\leq\frac{w}{2}$, that is, $j\geq\frac{w}{2}+1$. We have then $\sum_{\mathbf{a}\in\mathbb{F}_2^n;\,\mathbf{a}\preceq\overline{\mathbf{y}};\mathbf{a}\neq 0}\left(\sum_{\mathbf{x}\in\mathbb{F}_2^n;\,\mathbf{x}\preceq\mathbf{y}}(-1)^{f(\mathbf{a}+\mathbf{x})}\right)^2 \geq$

$\sum_{j=\lceil \frac{w}{2} \rceil+1}^{n-w} \binom{n-w}{j} \left(2\sum_{i=w-j+1}^{w} \binom{w}{i} - 2^w\right)^2 = \sum_{j=\lceil \frac{w}{2} \rceil+1}^{n-w} \binom{n-w}{j} \left(2^w - 2\sum_{i=0}^{w-j} \binom{w}{i}\right)^2$.
We deduce then from (3) that:

$$\max_{\mathbf{u}\in\mathbb{F}_2^n;\, \mathbf{u}\preceq\overline{\mathbf{y}}} W_f^2(\mathbf{u}) \geq 2^{2w} + \sum_{j=\lceil \frac{w}{2} \rceil+1}^{n-w} \binom{n-w}{j} \left(2^w - 2\sum_{i=0}^{w-j} \binom{w}{i}\right)^2.$$

Denoting $2w = n + k$ (where $k > 0$ has the same partity as n), we have then:

$$\max_{\mathbf{u}\in\mathbb{F}_2^n;\, \mathbf{u}\preceq\overline{\mathbf{y}}} W_f^2(\mathbf{u}) \geq 2^{n+k} + \sum_{j=\lceil \frac{n+k}{4} \rceil+1}^{\frac{n-k}{2}} \binom{\frac{n-k}{2}}{j} \left(2^{\frac{n+k}{2}} - 2\sum_{i=0}^{\frac{n+k}{2}-j} \binom{\frac{n+k}{2}}{i}\right)^2.$$

Hence, we have:

$$nl(f) \leq 2^{n-1} - \frac{1}{2}\sqrt{2^{n+k} + \sum_{j=\lceil \frac{n+k}{4} \rceil+1}^{\frac{n-k}{2}} \binom{\frac{n-k}{2}}{j} \left(2^{\frac{n+k}{2}} - 2\sum_{i=0}^{\frac{n+k}{2}-j} \binom{\frac{n+k}{2}}{i}\right)^2}. \quad (4)$$

2. Assume now that there exist vectors of Hamming weight smaller than $n/2$ and whose image by f equals 1. Let $\mathbf{y}$ have minimal Hamming weight w such that $f(\mathbf{y}) = 1$ ($w < n/2$). Applying the upper bound (4) to the monotone function $f(\overline{\mathbf{x}}) + 1$, whose nonlinearity equals that of f, and denoting $w' = n - w = \frac{n+k'}{2}$, where $k' > 0$ has the same partity as n, we have:

$$nl(f) \leq 2^{n-1} - \frac{1}{2}\sqrt{2^{n+k'} + \sum_{j=\lceil \frac{n+k'}{4} \rceil+1}^{\frac{n-k'}{2}} \binom{\frac{n-k'}{2}}{j} \left(2^{\frac{n+k'}{2}} - 2\sum_{i=0}^{\frac{n+k'}{2}-j} \binom{\frac{n+k'}{2}}{i}\right)^2}.$$

3. If none of the two cases above happens, this means that f coincides with the majority function at every input $\mathbf{x}$ of Hamming weight different from $n/2$ and we have seen above in Subsection 2.1 that either (i) f is a majority function and $nl(f)$ equals then $2^{n-1} - \binom{n-1}{n/2}$ if n is even and $2^{n-1} - \binom{n-1}{(n-1)/2}$ if n is odd, or (ii) n is even and $nl(f) \leq 2^{n-1} - \frac{1}{2}\sqrt{A}$ where A equals:

$$2^n + 2\sum_{1\leq j<n/4} \binom{n/2}{j} \left(2^{n/2} - 2\sum_{i=0}^{j} \binom{n/2}{i}\right)^2 + \left(2^{n/2} - 2\right)^2.$$

We deduce:

Theorem 2.3 *For every n and every monotone n-variable function f, we have $nl(f) \leq 2^{n-1} - \frac{1}{2}\sqrt{M}$, where $M = \min(A, B, C)$ if n is even and $M = \min(B, C)$ if n is odd, with*

$$A = 2^n + 2 \sum_{1 \leq j < n/4} \binom{n/2}{j} \left(2^{n/2} - 2 \sum_{i=0}^{j} \binom{n/2}{i} \right)^2 + \left(2^{n/2} - 2 \right)^2,$$

$$B = \min_{\substack{1 \leq k \leq n/2 \\ n+k \, even}} \left(2^{n+k} + \sum_{j=\lceil \frac{n+k}{4} \rceil + 1}^{\frac{n-k}{2}} \binom{\frac{n-k}{2}}{j} \left(2^{\frac{n+k}{2}} - 2 \sum_{i=0}^{\frac{n+k}{2}-j} \binom{\frac{n+k}{2}}{i} \right)^2 \right),$$

and $C = \begin{cases} \left[2\binom{n-1}{n/2} \right]^2 & \text{if } n \text{ is even} \\ \left[2\binom{n-1}{(n-1)/2} \right]^2 & \text{if } n \text{ is odd.} \end{cases}$

Let us study the asymptotical behavior of A, B and C when n tends to infinity:

- for $j \leq \frac{n}{4}\lambda$, where λ is any number strictly smaller than 1, we know, see e.g. [1], that $\sum_{i=0}^{j} \binom{n/2}{i}$ is negligeable with respect to $2^{n/2}$. Indeed, we have $\frac{2^{\frac{n}{2}H_2(\frac{\lambda}{2})}}{\sqrt{n\lambda(2-\lambda)}} \leq \sum_{0 \leq i \leq \frac{n}{4}\lambda} \binom{\frac{n}{2}}{i} \leq 2^{\frac{n}{2}H_2(\frac{\lambda}{2})} < 2^{\frac{n}{2}} e^{-\frac{n}{4}(1-\lambda)^2}$, where $H_2(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ is the binary entropy function. Asymptotically and for every $\lambda < 1$, A is then larger than $2^n \left(2 + 2\sum_{1 \leq j \leq \frac{n}{4}\lambda} \binom{n/2}{j} \right)$, which is larger than $\frac{2^{n+\frac{n}{2}H_2(\frac{\lambda}{2})}}{\sqrt{n\lambda(2-\lambda)}}$;
- for $\frac{n+k}{2} - j \leq \frac{n+k}{4}\lambda$, where λ is any number strictly smaller than 1, that is, for $j \geq \frac{n+k}{4}\mu$, where μ is any number strictly larger than 1, $\sum_{i=0}^{\frac{n+k}{2}-j} \binom{\frac{n+k}{2}}{i}$ is negligeable with respect to $2^{\frac{n+k}{2}}$ and, for every $\mu > 1$, B is then asymptotically larger than:

$$\min_{\substack{1 \leq k \leq n/2 \\ n+k \, even}} \left(2^{n+k} \left(1 + \sum_{j=\frac{n+k}{4}\mu}^{\frac{n-k}{2}} \binom{\frac{n-k}{2}}{j} \right) \right) =$$

$$\min_{\substack{1 \leq k \leq n/2 \\ n+k \, even}} \left(2^{n+k} \left(1 + \sum_{j=0}^{\frac{n-k}{2} - \frac{n+k}{4}\mu} \binom{\frac{n-k}{2}}{j} \right) \right),$$

  which is between $\frac{2^{n+k+(\frac{n-k}{2})H_2\left(\frac{1-\frac{n+k}{2(n-k)}\mu}{2}\right)}}{\sqrt{(n-k)(1-\frac{n+k}{2(n-k)}\mu)(1+\frac{n+k}{2(n-k)}\mu)}}$ and $2^{n+k+(\frac{n-k}{2})H_2\left(\frac{1-\frac{n+k}{2(n-k)}\mu}{2}\right)}$;
- $\binom{n-1}{n/2}$ and $\binom{n-1}{(n-1)/2}$ are both equivalent to $\frac{1}{\sqrt{2\pi n}}2^n$, according to the Stirling formula; hence, C is equivalent with $\frac{2^{2n+1}}{\pi n}$.

Hence, $\min(A, B, C)$ is asymptotically equivalent to an expression of n at least equal to $2^{\frac{3n\lambda_n}{2}}$ for some λ_n tending to 1.

In the tables below, we indicate for each value of n between 4 and 31 the value given by the upper bound of Theorem 2.3 and of λ_n such that this upper bound equals $2^{n-1} - 2^{\frac{3n\lambda_n}{4}}$, with the indication whether it is A, B or C which is minimal, and in the case it is B, the value of k for which the minimum is achieved.

n	4	5	6	7	8	9	10	11	12
Upper bound:	5.8	12	27	55.5	114.4	237.1	478.5	977.6	1975.1
λ_n:	0.39	0.53	0.52	0.59	0.63	0.63	0.68	0.67	0.69
$\min(A, B, C)$:	20	64	100	292	740	1424	4496	8596	21284
Minimum:	A	$B, k = 1$	A	$B, k = 1$	A	$B, k = 1$	$B, k = 2$	$B, k = 1$	$B, k = 2$

n	13	14	15	16	17	18	19	20
Upper bound:	3975.2	8013.1	16046.0	32298.1	64575.1	129723.8	259354.3	520377.3
λ_n:	0.71	0.71	0.75	0.74	0.78	0.77	0.80	0.80
$\min(A, B, C)$:	58328	128060	456948	883072	3693520	7271104	31129636	61175140
Minimum:	$B, k = 1$	$B, k = 2$	$B, k = 1$	$B, k = 2$	$B, k = 1$	$B, k = 2$	$B, k = 1$	$B, k = 2$

n	21	22	23	24	25	26
Upper bound:	1040509.3	2085531.4	4170870.5	8354450.6	16709542.2	33453505.5
λ_n:	0.82	0.82	0.84	0.84	0.86	0.85
$\min(A, B, C)$:	260285480	540151100	2196515260	4666906924	18318973664	40744645256
Minimum:	$B, k = 1$	$B, k = 2$	$B, k = 1$	$B, k = 2$	$B, k = 1$	$B, k = 2$

n	27	28	29	30	31
Upper bound:	66913246.3	133922618.2	267872592.2	536007586.2	1072122368.7
λ_n:	0.87	0.87	0.88	0.88	0.89
$\min(A,B,C)$:	153065065624	348359171016	1267262490224	2981325910789	10490542374645
Minimum:	$B, k=1$	$B, k=2$	$B, k=1$	$B, k=2$	$B, k=1$

Table 1. VALUES OF THE UPPER BOUND OF THEOREM 2.3 AND OF λ_n

These tables seem to confirm that the nonlinearity of any monotone Boolean function in n variables is bounded above by $2^{n-1} - 2^{\frac{3n\lambda_n}{4}}$ for some λ_n tending to 1. The nonlinearity of monotone functions is then much worse than what was suggested by the upper bounds obtained (resp. conjectured) in [9].

A part of the interest of Theorem 2.3 is to give knowledge on the nonlinearity of those functions differing with the majority function at vectors of weight $n/2$ only. Some observations were made in [14] about these functions.

3 Characterization of a class of weak Boolean functions and vectorial functions

Theorem 2.3 shows that for having a reasonably large nonlinearity, a Boolean function f should not be linearly equivalent to a monotone function (that is, there should not exist a monotone function g and a linear automorphism L such that $f = g \circ L$). And a vectorial function $F : \mathbb{F}_2^n \mapsto \mathbb{F}_2^m$ should not have a component function $\ell \circ F$ linearly equivalent to a monotone function (where ℓ is a nonzero linear form over $\mathbb{F}_2^m$). Let us try to characterize such weak functions.

3.1 Characterization of Boolean functions linearly equivalent to monotone functions

A Boolean function g is monotone if, for every $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n$ such that $supp(\mathbf{x}) \subset supp(\mathbf{y})$, we have $g(\mathbf{x}) \leq g(\mathbf{y})$. Denoting by B_0 the natural basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ of $\mathbb{F}_2^n$ (that is, the family of all vectors of Hamming weight 1) and by $\mathbf{x}^\perp$ the orthogonal space of $\{0, \mathbf{x}\}$, we have $B_0 \cap \mathbf{x}^\perp = \{\mathbf{e}_i,\ i \notin supp(\mathbf{x})\}$. Then g is monotone if and only if, for every $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n$, we have $(B_0 \cap \mathbf{y}^\perp \subset B_0 \cap \mathbf{x}^\perp) \Longrightarrow (g(\mathbf{x}) \leq g(\mathbf{y}))$. If $g' = g \circ L$ where L is a linear automorphism, the condition on g is equivalent to: $(B_0 \cap [L(\mathbf{y})]^\perp \subset B_0 \cap [L(\mathbf{x})]^\perp) \Longrightarrow (g'(\mathbf{x}) \leq g'(\mathbf{y}))$. Denoting by L^* the adjoint operator of L, we have $[L(\mathbf{x})]^\perp = (L^*)^{-1}(\mathbf{x}^\perp)$ and the condition on g is equivalent to: $(L^*(B_0) \cap \mathbf{y}^\perp \subset L^*(B_0) \cap \mathbf{x}^\perp) \Longrightarrow (g'(\mathbf{x}) \leq g'(\mathbf{y}))$. We deduce:

Proposition 3.1 *Let f be any n-variable Boolean function. Then, f is linearly equivalent to a monotone function if and only if there exists a basis B of the $\mathbb{F}_2$-vector space $\mathbb{F}_2^n$ such that*

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n,\ (B \cap \mathbf{y}^\perp \subset B \cap \mathbf{x}^\perp) \Longrightarrow (f(\mathbf{x}) \leq f(\mathbf{y})).$$

Equivalently:

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n, \, (f(\mathbf{x}) = 1 \text{ and } f(\mathbf{y}) = 0) \Longrightarrow (B \cap (\mathbf{y}^\perp \setminus \mathbf{x}^\perp) \neq \emptyset).$$

Corollary 3.2 *A Boolean function f is linearly inequivalent to monotone Boolean functions if and only if, for every basis B of the $\mathbb{F}_2$-vector space $\mathbb{F}_2^n$, there exists $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n$ such that $f(\mathbf{x}) = 1$ and $f(\mathbf{y}) = 0$ and $B \cap (\mathbf{y}^\perp \setminus \mathbf{x}^\perp) = \emptyset$.*

3.2 Characterization of vectorial functions with a least one component function linearly equivalent to monotone functions

Given a vectorial (n, m)-function $F : \mathbb{F}_2^n \mapsto \mathbb{F}_2^m$, according to Proposition 3.1, the component function $c \cdot F$, $c \neq 0$, is linearly equivalent to a monotone function if and only if:

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n, \, (B \cap \mathbf{y}^\perp \subset B \cap \mathbf{x}^\perp) \Longrightarrow (c \cdot F(\mathbf{x}) \leq c \cdot F(\mathbf{y})),$$

that is

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n, \, (c \cdot F(\mathbf{x}) = 1 \text{ and } c \cdot F(\mathbf{y}) = 0) \Longrightarrow (B \cap (\mathbf{y}^\perp \setminus \mathbf{x}^\perp) \neq \emptyset).$$

Then:

Proposition 3.3 *Lert F be any a vectorial (n, m)-function. Then there exists a component function of F which is linearly equivalent to a monotone function if and only if:*

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n, \exists c \in \mathbb{F}_2^n, \, (c \in (F(\mathbf{y}))^\perp \setminus (F(\mathbf{x}))^\perp) \Longrightarrow (B \cap (\mathbf{y}^\perp \setminus \mathbf{x}^\perp) \neq \emptyset).$$

Corollary 3.4 *A vectorial (n, m)-function F has all its component functions linearly inequivalent to monotone Boolean functions if and only if, for every basis B of $\mathbb{F}_2^n$ and every nonzero $c \in \mathbb{F}_2^m$, there exists $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n$ such that $(F(\mathbf{y}))^\perp \setminus (F(\mathbf{x}))^\perp$ contains c and $B \cap (\mathbf{y}^\perp \setminus \mathbf{x}^\perp) = \emptyset$.*

Acknowledgement. The author deeply thanks Stjepan Picek for his kind help for generating Table 1. We also thank Pante Stănică for having suggested, when we worked on [9], the problem of finding an upper bound on the nonlinearity of monotone Boolean functions.

References

1. N. Alon and J.H. Spencer. *The probabilistic method*. Wiley-VCH, 2000 (second edition).
2. A. Blum, C. Burch and J. Langford. On learning monotone Boolean functions. In Proc. 39th FOCS, pp. 408415. IEEE Comp. Soc. Press, 1998.
3. A. Blum, M. Furst, M. Kearns and R. Lipton. Cryptographic primitives based on hard learning problems. In Proc. Advances in Cryptology CRYPTO93, number 773 in LNCS, pp. 278291. Springer, 1993

4. N. Bshouty and C. Tamon. On the Fourier spectrum of monotone functions. J. ACM, 43(4):747770, 1996
5. A. Canteaut, C. Carlet, P. Charpin and C. Fontaine. On cryptographic properties of the cosets of $R(1,m)$. *IEEE Transactions on Information Theory* vol. 47, no 4, pp. 1494-1513, 2001.
6. C. Carlet, *Two new classes of bent functions*, Adv. in Crypt. – Eurocrypt 1993, LNCS 765 (1994), Springer–Verlag, pp. 77–101.
7. C. Carlet, *Boolean functions for cryptography and error correcting codes*, In: Y. Crama, P. Hammer (eds.), Boolean Methods and Models, Cambridge Univ. Press, Cambridge, pp. 257–397, 2010. Available at: `www.math.univ-paris13.fr/~carlet/pubs.html`
8. C. Carlet, *Vectorial Boolean functions for cryptography*, In: Y. Crama, P. Hammer (eds.), Boolean Methods and Models, Cambridge Univ. Press, Cambridge, pp. 398–469, 2010. Available at: `www.math.univ-paris13.fr/~carlet/pubs.html`
9. C. Carlet, D. Joyner, P. Stănică and D. Tang. Cryptographic properties of monotone Boolean functions. To appear in the *Journal of Mathematical Cryptography*, 2015.
10. C. Celerier, D. Joyner, C. Melles, D. Phillips, *On the Walsh–Hadamard transform of monotone Boolean functions*, Tbilisi Math. J., Special Issue on Sage and Research, vol. 5 (2012), 19–35.
11. N. Courtois, W. Meier, *Algebraic attacks on stream ciphers with linear feedback*, Adv. in Cryp. – Eurocrypt 2003, LNCS 2656, Springer–Verlag, 2003, pp. 345–359.
12. Y. Crama, P. L. Hammer, Boolean functions. Theory, algorithms, and applications, Cambridge University Press, Cambridge, 2011.
13. T. W. Cusick, P. Stănică, Cryptographic Boolean functions and applications, Elsevier–Academic Press, 2009.
14. D. K. Dalai, S. Maitra, S. Sarkar, *Basic theory in construction of Boolean functions with maximum possible annihilator immunity*, Des. Codes Cryptogr. 40:1, pp. 41–58, 2006.
15. D. Dachman-Soled, H. K. Lee, T. Malkin, R. A. Servedio, A. Wan and H. Wee. Optimal Cryptographic Hardness of Learning Monotone Functions. THEORY OF COMPUTING, Volume 5, pp. 257282, 2009.
16. J. F. Dillon, *Elementary Hadamard difference sets*, Proc. Sixth S. E. Conf. Combinatorics, Graph Theory and Computing, Utility Mathematics, Winnipeg, 1975, pp. 237–249.
17. F. J. MacWilliams and N. J. A. Sloane, The theory of error correcting codes, North-Holland, Amsterdam, 1977.
18. E. Mossel and R. O'Donnell. On the noise sensitivity of monotone functions. Random Struct. Algorithms, 23(3):333350, 2003.
19. W. Meier, E. Pasalic, C. Carlet, *Algebraic Attacks and Decomposition of Boolean Functions*, Advances in Cryptology – Eurocrypt 2004, LNCS 3027, Springer–Verlag, 2004, pp. 474–491.
20. R. O'Donnell and R. Servidio. Learning monotone decision trees in polynomial time. SIAM J. Comput., 37(3):827844, 2007
21. O. S. Rothaus, *On bent functions*, J. Combin. Theory Ser. A 20 (1976), 300–305.