THÈSE DE DOCTORAT DE
L'UNIVERSITÉ D'AIX-MARSEILLE

Discipline: **Mathématiques**

École Doctorale de Mathématiques et Informatique de Marseille

Soutenue publiquement le 11 décembre 2013 par

Stéphanie DIB

Pour obtenir le grade de
DOCTEURE DE L'UNIVERSITÉ D'AIX-MARSEILLE

Titre: **Distribution de la non-linéarité des fonctions Booléennes**

Directeur de thèse: François RODIER

RAPPORTEURS:

Daniel Augot	Directeur de recherche, INRIA/LIX, Saclay
Sihem Mesnager	Maître de conférences HDR, Université Paris VIII

COMPOSITION DU JURY

Daniel Augot	Directeur de recherche, INRIA/LIX, Saclay
Stéphane Ballet	Maître de conférences HDR, AMU/IML, Marseille
Gilles Lachaud	Directeur de recherche émérite, CNRS/IML, Marseille
Philippe Langevin	Professeur, Université de Toulon
Gregor Leander	Associate Professor, Université technique du Danemark
François Rodier	Directeur de recherche, CNRS/IML, Marseille

Contents

Chapter 1

Boolean functions and coding

1.1 Introduction

Cryptography is the practice and study of techniques that allow two people to communicate over an insecure communication channel. A third person who can see the *ciphertext*, that is the message sent through the channel instead of the original one, should not be able to determine anything about the original message, called the *plaintext*. The conversion of information from a readable plaintext to an unreadable ciphertext is called *encryption*. The encryption algorithm uses an encryption key which determines how the plaintext is to be encrypted. An authorized person, the receiver, is able to recover the message using a decryption algorithm that takes as input the ciphertext and a secret decryption key, and outputs the plaintext. In *symmetric* or *conventional cryptography*, the encryption and decryption keys are the same. Thus, people wishing to communicate must agree at first on a secret key. In *public key cryptography*, the encryption key is made available for everyone to use and encrypt messages, but the decryption key is known only by the receiver who can read the encrypted messages. A disadvantage of conventional cryptosystems is the key management. Each distinct pair of communicating persons must communicate secretly a different key, and thus the number of required keys for a group of persons to communicate pairwisely increases as about the square of the number of persons, whereas public key cryptosystems need only twice the number of communicating persons. However, the key management for conventional cryptosystems is made easier thanks to public key cryptosystems, since it allows to exchange keys without having to find a secure channel. All known public key cryptosystems need much more time to encrypt long messages than conventional cryptosystems, this is why conventional cryptography is widely used.

Error-correcting codes were invented to enable reliable data transmission over unreliable communication channels. Many communication chan-

nels are subject to channel noise, and thus errors can be introduced during the transmission from the source to a receiver. The general idea to achieve error detection and correction is to add some extra data to a message, called *redundancy*, which a receiver can use to detect errors and reconstruct the original data. A message to be sent is written as a binary string, and cut into blocks of some fixed size, say k. Each block is called *a word*, and is encoded individually into *codewords* of length n where $n \geq k$. The set of all possible codewords forms a code. With such a code, n symbols are sent when the amount of actual message is k. In this sense, the ratio $\frac{k}{n}$ measures the transmission speed, and is called the *rate* of the code. The rate must be as high as possible because messages must be sent as quickly - but also as reliably - as possible.

Boolean functions are functions from the vectorspace $\mathbb{F}_2^n$ of all binary vectors of length n to the finite field with two elements $\mathbb{F}_2$. In conventional cryptography, Boolean functions are fundamental to pseudorandom generators in stream ciphers and S-boxes in block ciphers. They also play an important role in error-correcting codes theory. Every n-variable Boolean function can be specified by a binary vector of length 2^n which gives its values at all of its 2^n arguments, and vice versa. Hence, every code with length 2^n can be defined as a subset of Boolean functions, such as Reed-Muller codes.

1.2 Boolean functions and cryptography

A *stream cipher* is a symmetric key cipher where the plaintext is added bitwise to a pseudorandom keystream of the same length, to produce the ciphertext stream. A stream cipher is a loose version of the unbreakable *one-time pad*, also known as the *Vernam cipher*. In a one-time pad cipher, the keystream must be completely random, as large as the plaintext and cannot be used more than once, which makes the system impractical and as a result, the one-time pad has not been widely used. A stream cipher uses a much smaller key, and based on this key, it generates a long pseudorandom keystream which can be added bitwise to the plaintext in a similar way to the one-time pad. But this comes with a cost, the proof of security associated to the one-time pad no longer holds, because the keystream is not truly random. To generate a pseudorandom sequence from a secret key, a stream cipher uses a *Linear-Feedback Shift Register* (*LFSR*) that outputs at every clock-cycle a bit of the keystream. A LFSR of length L consists of a shift register containing L bits $(s_{i+L-1}, s_{i+L-2}, \ldots, s_{i+1}, s_i)$ and a linear feedback function. At each clock-cycle, the right-most bit s_i is the output of the register, the others are shifted to the right, and the new left-most bit s_{i+L}

is given by the linear feedback function:

$$s_{i+L} = c_1 s_{i+L-1} + c_2 s_{i+L-2} + \cdots + c_{L-1} s_{i+1} + c_L s_i$$

where the coefficients c_i are binary. But the use of LFSRs on their own is insufficient to ensure good security, even when the feedback polynomial of the register is appropriately chosen. One technique to improve the security of LFSRs is to filter the content of a single LFSR through a Boolean function, called a *filtering function*, or to combine the outputs of n LFSRs with a n-input Boolean function, called a *combining function*. The security of both models depends on different properties of the Boolean functions involved.

A *block cipher* is a symmetric cipher that operates on a fixed length groups of bits, called *blocks*. Most block cipher algorithms are iterated block ciphers, which means that they transform a block of plaintext into a block of ciphertext of the same size, by applying repeatedely an invertible (to ensure decryption) transformation (e.g. permutations, *substitutions*...) called the *round function*, where each iteration is referred to as a *round*. The key used in a round is called a round key, and the round keys are computed from the secret key. *Substitution boxes* (*S-Boxes*) are fundamental parts of the round function and play a central role in their robustness, being the only source of nonlinearity. An S-Box substitutes a small block of input bits $(x_1, \ldots, x_n)$ with another block of output bits $(y_1, \ldots, y_m)$. The block $(y_1, \ldots, y_m)$ is the juxtaposition of the outputs of m Boolean functions whose common input is $(x_1, \ldots, x_n)$. In order to study the complexity of S-boxes, one should see it as an ensemble by studying *multi-output* Boolean functions or (n, m)*-functions*, instead of only individual Boolean functions. One can find more details on Boolean functions in cryptography in [Fon98, Car10a, Car10b].

1.3 Generalities on Boolean functions

We denote by $\mathcal{B}_n$ the set of all n-variable Boolean functions. Every Boolean function f can be identified by the vector $(f(x))_{x \in \mathbb{F}_2^n}$ that gives its values at all of its 2^n inputs, in some fixed order. The *Hamming weight* $w_H(u)$ of a vector $u \in \mathbb{F}_2^n$ is the size of its *support*, that is $supp(u) = \{1 \leq i \leq n \mid u_i \neq 0\}$. The *Hamming weight* $w_H(f)$ of a Boolean function f is the size of its support $supp(f) = \{x \in \mathbb{F}_2^n \mid f(x) \neq 0\}$. The *Hamming distance* $d_H(f, g)$ between two functions is the size of the set $\{x \in \mathbb{F}_2^n \mid f(x) \neq g(x)\}$, which equals $w_H(f + g)$.

1.3.1 Algebraic Normal Form

There exists several representations for Boolean functions, but we only mention the one we need in this document, which is the most often used representation in cryptography and coding. A Boolean function f can be represented

by a unique n-variable polynomial over $\mathbb{F}_2$ of the form

$$f(x) = \sum_{u \in \mathbb{F}_2^n} a_u \Big(\prod_{i=1}^{n} x_i^{u_i} \Big), \qquad a_u \in \mathbb{F}_2$$

where the sum is computed modulo 2. This representation, called the *Algebraic Normal Form*, belongs to $\mathbb{F}_2[x_1, \ldots, x_n]/(x_1^2 + x_1, \ldots, x_n^2 + x_n)$, since $x_i^2 = x_i$. The degree of the polynomial is called the algebraic degree of the function: $deg(f) = \max\{supp(u) \mid a_u \neq 0\,;\, u \in \mathbb{F}_2^n\}$. Boolean functions of degrees at most 1 are called *affine functions*

$$f(x) = a_0 + a_1 x_1 + \ldots + a_n x_n, \qquad a_i \in \mathbb{F}_2.$$

If $a_0 = 0$, f is said to be a linear function. Cryptographic Boolean functions must be complex and thus as far as possible from the simple functions, that is functions with degree 1 or more generally, functions with low degrees. In coding, affine functions give a definition of the Reed-Muller codes of order 1 as we shall see later.

1.3.2 Walsh Hadamard Transform

Let f be a Boolean function defined on $\mathbb{F}_2^n$. The Walsh Hadamard transform of f is the discrete Fourier transform of $\chi_f = (-1)^f$ the ± 1-representation of f

$$\widehat{\chi_f}(u) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x)+u\cdot x} \qquad\qquad u \in \mathbb{F}_2^n$$

where $u \cdot x$ denotes the usual inner product $u \cdot x = \sum_{i=1}^{n} u_i x_i$. The Walsh transform satisfies Parseval's relation

$$\sum_{u \in \mathbb{F}_2^n} \big(\widehat{\chi_f}(u)\big)^2 = 2^{2n}. \tag{1.1}$$

We have

$$\widehat{\chi_f}(0) = (2^n - w_H(f))(-1)^0 + w_H(f)(-1)^1 = 2^n - 2\, w_H(f). \tag{1.2}$$

For any linear function $l_a(x) = a \cdot x = a_1 x_1 + \ldots + a_n x_n$, we have by (1.2)

$$2^n - 2\, w_H(f + l_a) = \widehat{\chi_{f+l_a}}(0) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x)+a\cdot x} = \widehat{\chi_f}(a).$$

Thus the distance between f and any linear function $l_a(x) = a \cdot x$ equals

$$d_H(f, l_a) = w_H(f + l_a) = 2^{n-1} - \frac{1}{2}\,\widehat{\chi_f}(a) \tag{1.3}$$

and as a result

$$d_H(f, 1 + l_a) = 2^{n-1} + \frac{1}{2}\,\widehat{\chi_f}(a). \tag{1.4}$$

We will use these relations to define the nonlinearity through the Walsh transform.

1.3.3 Nonlinearity of Boolean functions

Cryptographic Boolean functions must satisfy many criteria in order to resist to the known attacks, and hopefully future attacks. The main criteria are *balancedness*, having a high *algebraic degree*, a high *nonlinearity* and a high *algebraic immunity*. But it is impossible for a function to meet all optimal properties. For instance, achieving the maximum possible nonlinearity leaves out both balancedness and having maximal algebraic degree, thus trade-offs must be considered. The *nonlinearity* is the most important criterion and is related to *fast correlation attacks* on stream ciphers [CT00] and *linear attacks* on stream and block ciphers [Mat94]. We say that there is a correlation between a Boolean function f and a linear function l if $d_H(f,l) \neq 2^{n-1}$. Due to (1.1) and (1.3), any Boolean function has correlation with some linear function. But this correlation must be small. The nonlinearity $\mathcal{NL}(f)$ of f is the minimum Hamming distance between f and affine functions. According to (1.3) and (1.4), the nonlinearity of f equals

$$\mathcal{NL}(f) = 2^{n-1} - \frac{1}{2}\max_{a\in\mathbb{F}_2^n}|\widehat{\chi_f(a)}|. \tag{1.5}$$

Equation (1.1) implies that $\max_{a\in\mathbb{F}_2^n}|\widehat{\chi_f(a)}| \geq 2^{n/2}$. Thus the nonlinearity of a Boolean function satisfy

$$\mathcal{NL}(f) \leq 2^{n-1} - 2^{n/2-1}.$$

This bound, called the *covering radius bound*, is reached if and only if $|\widehat{\chi_f(a)}| = 2^{n/2}$ for every vector a by (1.1). Functions that reaches this bound are called *bent functions*. They exist only for even values of n because $2^{n-1}-2^{n/2-1}$ must be an integer. By (1.2), bent functions have weight $2^{n-1}\pm 2^{n/2-1}$, which means that they are not balanced (i.e. they do not output as many 1's as 0's), and thus are not suitable for cryptographic use, but can help building highly nonlinear balanced functions.
When n is odd, (1.3.3) cannot be tight. The maximum possible nonlinearity of n-variable Boolean functions for odd n lies between $2^{n-1} - 2^{\frac{n-1}{2}}$ and $2^{n-1} - 2^{n/2-1}$. More precisely, we know that for $n = 1, 3, 5, 7$ it equals $2^{n-1} - 2^{\frac{n-1}{2}}$ [BW72, Myk80]. For $n = 9, 11, 13$, Kavut and Yücel [KDY10] have shown that it is greater than $2^{n-1} - 2^{\frac{n-1}{2}} + 2 \times 2^{\frac{n-9}{2}}$. And for odd $n \geq 15$, Patterson and Wiedemann [PW90] have shown that it is greater than $2^{n-1} - 2^{\frac{n-1}{2}} + 20 \times 2^{\frac{n-15}{2}}$.

When n is sufficiently large, it was shown by Olejàr and Stanek [OS98], that almost all Boolean functions have nonlinearities greater than $2^{n-1} - \sqrt{2^{n-1}n}$. This means that random functions are highly nonlinear. Rodier [Rod03, Rod04, Rod06] has shown more precisely that, asymptotically, the nonlinearities of almost all Boolean functions lie in a small neighbourhood

of $2^{n-1} - \sqrt{2^{n-1} n \log 2}$. More precise estimates on the rate of convergence were given later by Litsyn and Shpunt with a different approach [LS09].

1.3.4 The r-th order nonlinearity

Using the principle of the linear attack but with non-linear approximations on block ciphers [KR96], have led to a generalization of the notion of non-linearity. The r-th order nonlinearity $\mathcal{NL}_r(f)$ of a Boolean function f is the minimum distance between f and functions of degree at most r. Cryptographic functions must have high r-th order nonlinearities. The best known asymptotic upper bound on $\mathcal{NL}_r(f)$ is [CM07]

$$\mathcal{NL}_r(f) \leq 2^{n-1} - \frac{\sqrt{15}}{2} \cdot (1+\sqrt{2})^{r-2} \cdot 2^{\frac{n}{2}} + O\left(n^{r-2}\right).$$

When n tends to infinity, Carlet [Car06] proved the good behavior of Boolean functions with respect to the r-th order nonlinearity. For every r, he showed that the density of the set of functions satisfying

$$\mathcal{NL}_r(f) > 2^{n-1} - c\sqrt{2^{n-1}\binom{n}{r}\log 2}$$

tends to 1 when n tends to infinity, if $c > 1$. This was done by counting the number of functions whose r-th order nonlinearities are bounded from above by some number. In chapter 2, we study the second order nonlinearity of Boolean functions.

1.4 Vectorial Boolean functions

Let n and m be two positive integers. Functions from the vectorspace $\mathbb{F}_2^n$ to the vectorspace $\mathbb{F}_2^m$ are called (n,m)-functions or more generally, vectorial (i.e. multi-output) Boolean functions. These functions include the (single-output) Boolean functions ($m = 1$). In cryptography, (n,m)-functions are used as S-boxes which are parts of iterative block ciphers.

A (n,m)-function f is the juxtaposition of m Boolean functions of n variables $(f_1(x), \ldots, f_m(x))$, called the *coordinate functions* of f. Since each of these Boolean functions can be represented uniquely by a n-variable polynomial in $\mathbb{F}_2[x_1, \ldots, x_n]/(x_1^2 + x_1, \ldots, x_n^2 + x_n)$, then f can be uniquely represented by a polynomial of the same form but with coefficents in $\mathbb{F}_2^m$

$$f(x) = \sum_{u \in \mathbb{F}_2^n} a_u \left(\prod_{i=1}^{n} x_i^{u_i} \right), \qquad a_u \in \mathbb{F}_2^m$$

This representation belongs to $\mathbb{F}_2^m[x_1,\ldots,x_n]/(x_1^2+x_1,\ldots,x_n^2+x_n)$ and is called the *Algebraic Normal Form* of f. The algebraic degree of f is the degree of its ANF: $deg(f) = \max\{supp(u) \mid a_u \neq (0,\ldots,0)\,;\, u \in \mathbb{F}_2^n\}$. Thus it equals the maximum algebraic degree of the coordinate functions of f. It also equals the maximum algebraic degree of the component functions of f.

The nonlinearity is an important criterion for cryptographic vectorial functions for the resistance to the linear attack [Mat94] and *differential attack* [CV95]. The generalization of the nonlinearity[1] of Boolean functions to (n,m)-functions was introduced by Nyberg [Nyb93] and later studied by Chabaud and Vaudenay [CV95]. The nonlinearity of all the *component functions* of a (n,m)-function f, that is $v \cdot f$ where $v \in \mathbb{F}_2^{m*}$, must be as high as possible. The nonlinearity $\mathcal{NL}(f)$ of a (n,m)-function f is the minimum nonlinearity of all the component Boolean functions of f. Thus, we have

$$\begin{aligned}\mathcal{NL}(f) &= \min_{v\in\mathbb{F}_2^{m*}} \mathcal{NL}(v.f) \\ &= 2^{n-1} - \frac{1}{2} \max_{\substack{v\in\mathbb{F}_2^{m*}\\ u\in\mathbb{F}_2^n}} |\widehat{\chi_{v\cdot f}}(u)|.\end{aligned}$$

The covering radius bound being valid for every n-variable Boolean function, is a fortiori, valid for every (n,m)-function

$$\mathcal{NL}(f) \leq 2^{n-1} - 2^{n/2-1}. \tag{1.6}$$

(n,m)-functions which achieve the covering radius bound are called bent functions. f is bent if and only if $|\widehat{\chi_{v\cdot f}}(u)| = 2^{n/2}$ for every vector $v \in \mathbb{F}_2^{m*}$ and every $u \in \mathbb{F}_2^n$, thus if and only if all the component functions of f are bent. Nyberg proved [Nyb91] that this bound can be achieved with equality only if n is even and $m \leq n/2$. Bent functions being not balanced (i.e. they do not take every value of $\mathbb{F}_2^m$ the same number 2^{n-m} of times), are not suitable for cryptotsystems. This is why it is important to study functions with large but not optimal nonlinearities, among which some balanced functions exist. When $m \geq n-1$, we have the Sidelnikov-Chabaud-Vaudenay bound [CV95]

$$\mathcal{NL}(f) \leq 2^{n-1} - \frac{1}{2}\sqrt{3\times 2^n - 2 - 2\,\frac{(2^n-1)(2^{n-1}-1)}{(2^m-1)}},$$

which coincides with the covering radius bound for $m = n-1$ and improves it for $m \geq n$. When $m \geq n$, this bound is achieved only if $m = n$ with n odd, since the square root must be an integer. The (n,n)-functions which achieve this bound $2^{n-1} - 2^{\frac{n-1}{2}}$ are called almost bent functions. Other

[1] The notion of higher order nonlinearity is also extended to S-boxes.

bounds have been obtained in [CD07] when m is sufficiently greater than n. Finding better bounds than (4.2) in the other cases remains an open problem.

In chapter 4, we give asymptotic bounds on the nonlinearity of almost all vectorial functions.

1.5 Boolean functions and coding

The binary Reed-Muller code of order r, $0 \leq r \leq n$, of length 2^n, denoted by $\mathcal{RM}(r,n)$, is the set of the binary vectors corresponding to all n-variable Boolean functions of degree at most r. The code $\mathcal{RM}(r,n)$ has dimension $k = \sum_{i=0}^{r} \binom{n}{i}$ (since this is the number of monomials of degree at most r), and minimum distance $d_{\min} = 2^{n-r}$. Thus, $\mathcal{RM}(r,n)$ can correct $\left\lfloor \frac{1}{2}(d_{\min} - 1) \right\rfloor$ errors. Namely, for every codeword $u \in \mathcal{RM}(r,n)$, let us consider the Hamming ball $\mathcal{B}_r[u] = \{v \in \mathbb{F}_2^n \mid d_H(u,v) \leq r\}$ of radius r centered at u. When $r \leq \left\lfloor \frac{1}{2}(d_{\min} - 1) \right\rfloor$, these balls do not overlap, because the minimum Hamming distance between codewords is $d_{\min}$. Then, if a codeword u is transmitted and r errors occur, then the received vector (word) v is inside the ball $\mathcal{B}_r[u]$ and is closer to u than to any other codeword. Thus, the decoder will be able to recover the correct codeword using an efficient algorithm that outputs for every word the nearest codeword. If more than $d_{\min}/2$ errors occur, the received vector may or may not be closer to some other codeword than to the correct (transmitted) one. If it is the case, then the decoder will output the wrong codeword. In chapter 3, we give a proof on why this happens very rarely with $\mathcal{RM}(1,n)$, even when the number of errors is much larger than half the minimum distance. On the contrary, when the number of errors exceeds a certain value, the received vector will be rarely closer to the correct codeword than to any other one. This is due to the way the k-dimensional code $\mathcal{RM}(1,n)$ is spread over the n-dimensional vectorspace $\mathbb{F}_2^n$.

Chapter 2

Asymptotic Nonlinearity of Boolean Functions

2.1 Introduction

A Boolean function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ is often represented by its Algebraic Normal Form, that is the unique n-variable polynomial over $\mathbb{F}_2$ of the form

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

where $a_u \in \mathbb{F}_2$. The degree of the polynomial is called the algebraic degree of the function. We denote by $\mathcal{B}_n$ the set of Boolean functions of n variables. The r-th order nonlinearity $\mathcal{NL}_r(f)$ of a Boolean function is the minimum number of substitutions required in its truth table to change it into a function of degree at most r. For $0 \leq r \leq n$, the r-th order Reed-Muller code $\mathcal{RM}(r, n)$ of length 2^n, is the set of all n-variable Boolean functions of degree at most r, where each function is represented by a vector that gives its values in all of its 2^n arguments. The Hamming distance $d_H(f, g)$ between two functions f and g is

$$d_H(f, g) = |\{x \in \mathbb{F}_2^n \mid f(x) \neq g(x)\}| .$$

Then

$$\mathcal{NL}_r(f) = \min_{g \in \mathcal{RM}(r,n)} d_H(f, g).$$

It is examined when dealing with low-degree approximation attacks [KR96, SK98].

Note that the maximum possible r-th order nonlinearity coincides with the covering radius of $\mathcal{RM}(r, n)$. When $r = 1$, the (first order) nonlinearity of a Boolean function satisfy

$$\mathcal{NL}(f) \leq 2^{n-1} - 2^{\frac{n}{2}-1}. \tag{2.1}$$

For $r > 1$, Carlet and Mesnager [CM07] obtained an upper bound on the covering radius of $\mathcal{RM}(r, n)$, which gives

$$\mathcal{NL}_r(f) \leq 2^{n-1} - \frac{\sqrt{15}}{2} \cdot (1+\sqrt{2})^{r-2} \cdot 2^{\frac{n}{2}} + O\left(n^{r-2}\right)$$

for every n-variable Boolean function.

Our interest is the r-th order nonlinearity of random functions. When n is sufficiently large, it was shown by Carlet [Car02], Olejàr and Stanek [OS98], that almost all Boolean functions have nonlinearities greater than $2^{n-1} - c\sqrt{2^{n-1}n\log 2}$ where c is any real number greater than 1. This result was generalized later on for every r. In [Car06], Carlet proved that the density of the set of functions satisfying

$$\mathcal{NL}_r(f) > 2^{n-1} - c\sqrt{2^{n-1}\binom{n}{r}\log 2} \tag{2.2}$$

tends to 1 when n tends to infinity, if $c > 1$. This means that random functions have high r-th order nonlinearities. When $r = 1$, the gap between the lower (2.2) and upper bound (2.1) was shortened in [Rod03, Rod04, Rod06]. Using Halász method in [Hal73] concerning random trigonometric polynomials, Rodier proved that the nonlinearities of almost all (Boolean) functions lie in a small neighbourhood of $2^{n-1} - \sqrt{2^{n-1}n\log 2}$ when n tends to infinity. More precise estimates were given later by Litsyn and Shpunt with a different approach [LS09]. In [Dib10], we were enabled to extend this result to the second order nonlinearity by applying mainly fundamental combinatorics. We present this result in this chapter.

We prove in theorem 2.6.1 that for $r \leq 2$, the density in $\mathcal{B}_n$ of the set of functions satisfying

$$\mathcal{NL}_r(f) > 2^{n-1} - c\sqrt{2^{n-1}\binom{n}{r}\log 2}$$

tends to 0 when n tends to infinity, if $0 < c < 1$. This result, together with (2.2), yield a concentration of the nonlinearities of almost all functions around $2^{n-1} - \sqrt{2^{n-1}\binom{n}{r}\log 2}$ when $r \leq 2$. To attain this goal, we consider $\mathcal{B}_n$ as a metric space with the Hamming distance $d_H(f, g)$ and weight $w_H(f) = d_H(f, 0)$, where $0 = (0, ...0)$. For a given function g and an integer ρ, we denote by $\mathcal{B}_\rho[g]$ the Hamming ball with centre g and radius ρ, defined by

$$\mathcal{B}_\rho[g] = \{f \in \mathcal{B}_n \mid d_H(f, g) \leq \rho\}.$$

Let $\delta = 2^{n-1} - c\sqrt{2^{n-1}\binom{n}{r}\log 2}$ where c is a positive real number. We want to estimate the density of the subset of functions whose r-th order

nonlinearities exceeds δ. By definition, these functions are those for which the Hamming distance to $\mathcal{RM}(r,n)$ exceeds δ. Thus

$$\mathcal{NL}_r(f) > \delta \iff \sum_{g \in \mathcal{RM}(r,n)} 1_{\mathcal{B}_\delta[g]}(f) = 0$$

where $1_{\mathcal{B}_\delta[g]}$ is the indicator function of the subset $\mathcal{B}_\delta[g]$. Let η be an integer-valued function on $\mathcal{B}_n$ defined by

$$\eta(f) = \sum_{g \in \mathcal{RM}(r,n)} 1_{\mathcal{B}_\delta[g]}(f).$$

Then

$$P\left(\mathcal{NL}_r(f) > \delta\right) = P\left(\eta = 0\right).$$

We consider all functions to be equally probable. When $\eta = 0$, η deviates from its expectation by $E(\eta)$, and by Chebyshev's inequality [AS92, chapter IV], we have

$$P(\eta = 0) \leq P\big(|\eta - E(\eta)| \geq E(\eta)\big) \leq \frac{E(\eta^2) - E^2(\eta)}{E^2(\eta)}$$

which tends to 0 if $c < 1$ as we shall prove.

The paper is organised as follows. In section 2.2, we give the main tools of our proof. In section 2.3, we give an estimation of the expected value of η. In section 2.4 and 2.5, we give an estimation of the expected value of η^2. We present our bound in section 2.6 and explain why our method fails for $r > 2$ in the final section.

2.2 Some estimations

The following lemma (see [Car06, lemma 1]) gives well-known asymptotic estimate of the sum of binomial coefficients.

Lemma 2.2.1. *Let n be a positive integer and $k \leq n$. Then*

$$\binom{2n}{k} \leq \sum_{i=0}^{k} \binom{2n}{i} < 2^{2n} \cdot \exp\left(-\frac{(n-k)^2}{n}\right). \tag{2.3}$$

When k is sufficiently close to n, the following lemma (see [GKP89, chapter IX, (9.98)], [Fel68, chapter VII]) gives a better asymptotic estimation for $\binom{2n}{k}$:

Lemma 2.2.2. *Let n be a positive integer and $|n-k| \leq n^{\frac{5}{8}}$. Then*

$$\binom{2n}{k} = \frac{2^{2n}}{\sqrt{\pi \cdot n}} \cdot \exp\left(-\frac{(n-k)^2}{n}\right) \cdot (1 + o(1)). \tag{2.4}$$

Now, we introduce a function of great importance that will occur very often in our proof (see [Fel68, Chapter IIV], [Dur91, Theorem (1.3)]):

Proposition 2.2.1. *For $y < 0$, the function defined by*

$$\mathcal{H}(y) = \int_{-\infty}^{y} e^{-u^2} du$$

satisfies

$$\left(1 - \frac{1}{2y^2}\right) \frac{\exp\left(-y^2\right)}{-2y} \leq \mathcal{H}(y) \leq \frac{\exp\left(-y^2\right)}{-2y}.$$

In particular,

$$\mathcal{H}(y) = \frac{\exp\left(-y^2\right)}{-2y} (1 + o(1)) \quad \text{as} \quad y \to -\infty. \tag{2.5}$$

Consequently, when y and z tend to $-\infty$

$$\mathcal{H}(z) = o\left(\mathcal{H}(y)\right) \qquad \text{if} \quad y^2 - z^2 \to -\infty \tag{2.6}$$

$$\mathcal{H}(z) = \mathcal{H}(y)\left(1 + o(1)\right) \qquad \text{if} \quad y^2 - z^2 \to 0. \tag{2.7}$$

We will be dealing very often with exponential sums that we evaluate through the function $\mathcal{H}(y)$. In the following lemma, we show the little manipulations we do to this purpose.

Lemma 2.2.3. *Let $f(x) = u_n\, x + v_n$ be a negative, strictly increasing real function defined on $\mathbb{R}$. Then*

$$\mathcal{H}\left(f(b_n)\right) - \mathcal{H}\left(f(a_n)\right) \leq u_n \cdot \sum_{a_n}^{b_n} \exp\left(-f^2(x)\right) \leq \mathcal{H}\left(f(b_n + 1)\right). \tag{2.8}$$

If $f^2(b_n) - f^2(a_n) \to -\infty$ and $f^2(b_n + 1) - f^2(b_n) \to 0$ when n tends to ∞, then

$$\sum_{a_n}^{b_n} \exp\left(-f^2(x)\right) = \frac{1}{u_n} \cdot \mathcal{H}\left(f(b_n)\right) \cdot (1 + o(1)). \tag{2.9}$$

Proof. f is a negative strictly increasing function, then $\exp\left(-f^2(x)\right)$ is an increasing function, and because of this monotonicity, it is clear that the

area under the graph of $\exp\left(-f^2(x)\right)$ between x_1 and x_1+1 is smaller than $\exp\left(-f^2(x_1+1)\right)$ and larger than $\exp\left(-f^2(x_1)\right)$. It follows that

$$\int_{a_n}^{b_n} \exp\left(-f^2(y)\right) dy \leq \sum_{a_n}^{b_n} \exp\left(-f^2(x)\right) \leq \int_{a_n}^{b_n+1} \exp\left(-f^2(y)\right) dy. \tag{2.10}$$

We change the variables to simplify the integrands

$$\int_{f(a_n)}^{f(b_n)} \exp\left(-z^2\right) dz \leq u_n \cdot \sum_{a_n}^{b_n} \exp\left(-f^2(x)\right) \leq \int_{f(a_n)}^{f(b_n+1)} \exp\left(-z^2\right) dz$$

and by definition of $\mathcal{H}$, we get

$$\mathcal{H}\left(f(b_n)\right) - \mathcal{H}\left(f(a_n)\right) \leq u_n \cdot \sum_{a_n}^{b_n} \exp\left(-f^2(x)\right) \leq \mathcal{H}\left(f(b_n+1)\right).$$

(2.9) is directly given by (2.6) and (2.7).

□

We will be working also with summations over integers having a fixed parity, let's say even. In this case, instead of (2.10), we have

$$\int_{a_n}^{b_n-1} \exp\left(-f^2(y)\right) dy \leq 2 \cdot \sum_{\substack{a_n \\ x \text{ even}}}^{b_n} \exp\left(-f^2(x)\right) \leq \int_{a_n}^{b_n+2} \exp\left(-f^2(y)\right) dy$$

which holds independently of the parity of a_n and b_n, and also if we are choosing odd integers instead of even. Following the steps of lemma 2.2.3 gives

$$\mathcal{H}\left(f(b_n-1)\right) - \mathcal{H}\left(f(a_n)\right) \leq 2\, u_n \sum_{\substack{a_n \\ x \text{ even}}}^{b_n} \exp\left(-f^2(x)\right) \leq \mathcal{H}\left(f(b_n+2)\right).$$

If $f^2(b_n-1) - f^2(a_n) \to -\infty$, $f^2(b_n-1) - f^2(b_n) \to 0$ and $f^2(b_n+2) - f^2(b_n) \to 0$ when n tends to ∞, then

$$\sum_{\substack{a_n \\ x \text{ even}}}^{b_n} \exp\left(-f^2(x)\right) = \frac{1}{2\, u_n} \cdot \mathcal{H}\left(f(b_n)\right) \cdot (1+o(1)) \tag{2.11}$$

by (2.6) and (2.7).

Now, we are ready to estimate the expected value of η.

2.3 Estimation of the size of a Hamming ball

Because of the linearity of E, we have

$$E(\eta) = \sum_{g \in \mathcal{RM}(r,n)} E(1_{\mathcal{B}_\delta[g]}) = \sum_{g \in \mathcal{RM}(r,n)} P(\mathcal{B}_\delta[g]).$$

Note that the size of a Hamming ball does not depend on the choice of the center for a simple translation reason. Thus,

$$E(\eta) = 2^{\sum_{i=0}^{r} \binom{n}{i}} \cdot P\left(\mathcal{B}_\delta[0]\right)$$

where $\sum_{i=0}^{r} \binom{n}{i}$ is the dimension of $\mathcal{RM}(r,n)$.

Proposition 2.3.1. *Let r be a fixed integer, $\delta = 2^{n-1} - c\sqrt{2^{n-1}\binom{n}{r}\log 2}$ where c is a positive constant. We have*

$$P\left(\mathcal{B}_\delta[0]\right) = \frac{1}{\sqrt{\pi}} \cdot \mathcal{H}\left(-c\sqrt{\tbinom{n}{r}\log 2}\right) \cdot (1 + o(1))$$

when n tends to infinity.

Proof. The number of Boolean functions whose Hamming distance to 0 is bounded from above by some number δ equals

$$\sum_{0 \le i \le \delta} \binom{2^n}{i}.$$

We divide the sum into two sums

$$\sum_{0 \le i \le \delta} \binom{2^n}{i} = \sum_{0 \le i < 2^{n-1} - 2^{(n-1)\frac{5}{8}}} \binom{2^n}{i} + \sum_{2^{n-1} - 2^{(n-1)\frac{5}{8}} \le i \le \delta} \binom{2^n}{i}$$

that we denote by $\sum_1$ and $\sum_2$ respectively, and estimate each one separatly. Using (2.3),

$$\begin{aligned} \sum\nolimits_1 &< 2^{2^n} \exp\left(-2^{(n-1)\frac{1}{4}}\right) \\ &= O\left(2^{2^n - 2^{n/4}}\right). \end{aligned}$$

To estimate the second sum, we use (2.4)

$$\sum\nolimits_2 = \frac{2^{2^n}}{\sqrt{\pi\, 2^{n-1}}} \cdot (1 + o(1)) \cdot \sum_{2^{n-1} - 2^{(n-1)\frac{5}{8}} \le i \le \delta} \exp\left(-\frac{(2^{n-1} - i)^2}{2^{n-1}}\right).$$

According to (2.9), we have

$$\begin{aligned}\sum\nolimits_2 &= \frac{2^{2^n}}{\sqrt{\pi\, 2^{n-1}}} \cdot (1+o(1)) \cdot \sqrt{2^{n-1}} \cdot \mathcal{H}\left(-c\sqrt{\tbinom{n}{r}\log 2}\right) \cdot (1+o(1)) \\ &= \frac{2^{2^n}}{\sqrt{\pi}} \cdot \mathcal{H}\left(-c\sqrt{\tbinom{n}{r}\log 2}\right) \cdot (1+o(1)) \\ &= \frac{2^{2^n}}{\sqrt{\pi}} \cdot \frac{2^{-c^2\binom{n}{r}}}{2c\sqrt{\binom{n}{r}\log 2}} \cdot (1+o(1)),\end{aligned}$$

by (2.5). Therefore, $\sum_1$ is $o\left(\sum_2\right)$, and we obtain the required result by dividing by the total number of Boolean functions of n variables, that is 2^{2^n}. □

Estimating the expected value of η^2 is rather lengthy, this is why we divide it in two sections.

2.4 Intersection of two Hamming balls

The random variable η^2, defined on the set $\mathcal{B}_n$ of equiprobable Boolean functions, is equal to

$$\eta + \sum_{\substack{(g_1,g_2)\in\mathcal{RM}(r,n)\times\mathcal{RM}(r,n)\\ g_1\neq g_2}} 1_{\mathcal{B}_\delta[g_1]\cap\mathcal{B}_\delta[g_2]}.$$

Note that the size of intersection of any two Hamming balls depends only on the Hamming distance between their centers. Thus, we can suppose that the centers are 0 and g such that $g = g_1 + g_2$. As a result

$$\begin{aligned}E\left(\eta^2\right) &= E(\eta) + \sum_{\substack{(g_1,g_2)\in\mathcal{RM}(r,n)\times\mathcal{RM}(r,n)\\ g_1\neq g_2}} P\left(\mathcal{B}_\delta[g_1]\cap\mathcal{B}_\delta[g_2]\right) \\ &= E(\eta) + 2^{\sum_{i=0}^{r}\binom{n}{i}} \cdot \sum_{\substack{g\in\mathcal{RM}(r,n)\\ g\neq 0}} P(\mathcal{B}_\delta[0]\cap\mathcal{B}_\delta[g]).\end{aligned}$$

This means that we have to consider the weight distribution of $\mathcal{RM}(r,n)$. Let A_w be the number of functions of weight w in $\mathcal{RM}(r,n)$. The first order Reed-Muller code $\mathcal{RM}(1,n)$ has weight distribution

$$A_0 = A_{2^n} = 1, \qquad A_{2^{n-1}} = 2^{1+n} - 2. \tag{2.12}$$

The weight distribution of $\mathcal{RM}(2,n)$ is given by (see [MS77, Chapter XV]): $A_w = 0$ unless $w = 2^{n-1}$ or $w = 2^{n-1}(1 \pm 2^{-h})$ for some h, $0 \le h \le \lfloor \frac{n}{2} \rfloor$. Also, $A_0 = A_{2^n} = 1$ and

$$A_{2^{n-1}(1\pm 2^{-h})} = 2^{2h(n-h)+h} \cdot \frac{\prod_{i=0}^{2h-1}(1-2^{i-n})}{\prod_{i=1}^{h}(1-4^{-i})} \tag{2.13}$$

where $\frac{\prod_{i=0}^{2h-1}(1-2^{i-n})}{\prod_{i=1}^{h}(1-4^{-i})} < \frac{3}{2}$ (see [HKL05]), and $A_{2^{n-1}} = 2^{1+n+\binom{n}{2}} - \sum_{w \neq 2^{n-1}} A_w$. Accordingly,

$$E(\eta^2) = E(\eta) + 2^{\sum_{i=0}^{r}\binom{n}{i}} \cdot \sum_{w \neq 0} A_w \cdot P(\mathcal{B}_\delta[0] \cap \mathcal{B}_\delta[g_w]) \tag{2.14}$$

where the sum is taken over all weights in $\mathcal{RM}(r,n)$ and g_w is any codeword of weight w. When $r = 1$, we get

$$E(\eta^2) = E(\eta) + 2^{1+n} \cdot (2^{1+n} - 2) \cdot P(\mathcal{B}_\delta[0] \cap \mathcal{B}_\delta[g_{2^{n-1}}]),$$

since $\mathcal{B}_\delta[0] \cap \mathcal{B}_\delta[1] = \emptyset$, δ being smaller than $\frac{1}{2}\, d_H(0,1) = 2^{n-1}$.

Proposition 2.4.1. *Let Δ be some integer. Let g be a Boolean function in n variables and w its Hamming weight. Then the size of the subset $\mathcal{B}_\Delta[0] \cap \mathcal{B}_\Delta[g]$ is*

$$\sum_{0 \leq \delta_1, \delta_2 \leq \Delta} \binom{w}{\frac{\delta_1 - \delta_2 + w}{2}} \binom{2^n - w}{\frac{\delta_1 + \delta_2 - w}{2}}.$$

Proof. A Boolean function f belongs to the intersection of $\mathcal{B}_\Delta[0]$ and $\mathcal{B}_\Delta[g]$ if and only if

$$d_H(f,0) = \delta_1 \quad \text{and} \quad d_H(f,g) = \delta_2, \tag{2.15}$$

where δ_1, $\delta_2 \in [0, \Delta]$. In order to count the number of functions satisfying these conditions, we look at $\mathbb{F}_2^n$ as the union of the support of g (i.e. vectors for which g equals 1) of cardinality w and its complement in $\mathbb{F}_2^n$ (of cardinality $2^n - w$), and find the number of possible ways to choose the support of f of δ_1 vectors from these two sets, while respecting its distance to g, that is δ_2. Therefore, we consider

$$a_i = \left|\{x \in \mathbb{F}_2^n \,/\, f(x) = 1,\, g(x) = i\}\right|, \qquad i = 0 \,\text{or}\, 1.$$

We have by (2.15)

$$\delta_1 = a_0 + a_1 \quad \text{and} \quad \delta_2 = a_0 + (w - a_1).$$

Thus

$$a_0 = \frac{\delta_1 + \delta_2 - w}{2} \quad \text{and} \quad a_1 = \frac{\delta_1 - \delta_2 + w}{2}$$

and the number of functions in $\mathcal{B}_\Delta[0] \cap \mathcal{B}_\Delta[g]$ is

$$\sum_{0 \leq \delta_1, \delta_2 \leq \Delta} \binom{w}{a_1} \binom{2^n - w}{a_0}.$$

By definition, a_0 (resp. a_1) is a non-negative integer less than $2^n - w$ (resp. w). Hence, for values of δ_1 and δ_2 that do not satisfy these conditions, the sum is null. □

Proposition 2.4.2. *Let r be a fixed integer, $\delta = 2^{n-1} - c\sqrt{2^{n-1}\binom{n}{r}\log 2}$ where c is a positive constant. Let g be a Boolean function in n variables with Hamming weight $w = 2^{n-1}(1+t)$, where $|t| \neq 1$. Then*

$$P\left(\mathcal{B}_\delta[0] \cap \mathcal{B}_\delta[g]\right) \leq \frac{\sqrt{1+|t|}}{\pi\sqrt{1-|t|}} \cdot \mathcal{H}^2\left(-\frac{c\sqrt{\binom{n}{r}\log 2}}{\sqrt{1+|t|}}\right) \cdot (1+o(1)). \quad (2.16)$$

If $|t| = o\left(\frac{1}{\binom{n}{r}}\right)$ we have

$$P\left(\mathcal{B}_\delta[0] \cap \mathcal{B}_\delta[g]\right) \leq \frac{1}{\pi} \cdot \mathcal{H}^2\left(-c\sqrt{\binom{n}{r}\log 2}\right) \cdot (1+o(1)), \quad (2.17)$$

which is tight if $t = 0$.

Proof. Due to proposition 2.4.1, we have

$$P(\mathcal{B}_\delta[0] \cap \mathcal{B}_\delta[g]) = 2^{-2^n} \sum_{0 \leq \delta_1, \delta_2 \leq \delta} \binom{w}{\frac{\delta_1 - \delta_2 + w}{2}} \binom{2^n - w}{\frac{\delta_1 + \delta_2 - w}{2}}.$$

Whenever the binomial coefficient is not null, δ_1 and δ_2 clearly have the same parity since w is even. We divide the sum into two sums, and estimate each one separatly. Let

$$\sum\nolimits_3 = \sum_{2^{n-1} - 2^{(n-1)\frac{5}{8}} \leq \delta_1, \delta_2 \leq \delta} \binom{w}{\frac{\delta_1 - \delta_2 + w}{2}} \binom{2^n - w}{\frac{\delta_1 + \delta_2 - w}{2}}$$

and

$$\begin{aligned}\sum\nolimits_4 &= \sum_{\substack{\delta_1 < 2^{n-1} - 2^{(n-1)\frac{5}{8}};\, \delta_2 \leq \delta \\ \text{or } \delta_2 < 2^{n-1} - 2^{(n-1)\frac{5}{8}};\, \delta_1 \leq \delta}} \binom{w}{\frac{\delta_1 - \delta_2 + w}{2}} \binom{2^n - w}{\frac{\delta_1 + \delta_2 - w}{2}} \\ &\leq 2 \sum_{\delta_1 < 2^{n-1} - 2^{(n-1)\frac{5}{8}};\, \delta_2 \leq \delta} \binom{w}{\frac{\delta_1 - \delta_2 + w}{2}} \binom{2^n - w}{\frac{\delta_1 + \delta_2 - w}{2}},\end{aligned}$$

since the binomial coefficients are symmetric in δ_1 and δ_2.

δ_1 and δ_2 fulfill the conditions of (2.4) in $\sum_3$, thus

$$\begin{aligned}\sum\nolimits_3 = \frac{2^{2^n+1}}{\pi\sqrt{w(2^n-w)}} \cdot (1+o(1)) \cdot \sum_{\substack{2^{n-1} - 2^{(n-1)\frac{5}{8}} \leq \delta_1, \delta_2 \leq \delta \\ \delta_1 = \delta_2 \pmod 2}} \\ \exp\left(-\frac{\left(\frac{\delta_1 - \delta_2}{2}\right)^2}{w/2}\right) \cdot \exp\left(-\frac{\left(\frac{\delta_1 + \delta_2 - w}{2} - \frac{2^n - w}{2}\right)^2}{(2^n - w)/2}\right).\end{aligned}$$

Rearranging gives

$$\sum_3 = \frac{2^{2^n+1}}{\pi\, 2^{n-1}\sqrt{1-t^2}} \cdot (1+o(1)) \cdot \sum_{\substack{2^{n-1}-2^{(n-1)\frac{5}{8}} \le \delta_1,\delta_2 \le \delta \\ \delta_1 = \delta_2 \,(\mathrm{mod}\, 2)}} \exp\left(- \frac{\left(\delta_1 - 2^{n-1}\right)^2 + \left(\delta_2 - 2^{n-1}\right)^2 + 2\,t\left(\delta_1 - 2^{n-1}\right)\left(\delta_2 - 2^{n-1}\right)}{2^{n-1}(1-t^2)} \right).$$

We denote the latter sum by $\sum_5$ and bound it using

$$-(u^2+v^2+2uvt) \le -(u^2+v^2)(1-|t|). \tag{2.18}$$

$$\begin{aligned}
\sum_5 &\le \sum_{\substack{2^{n-1}-2^{(n-1)\frac{5}{8}} \le \delta_1,\delta_2 \le \delta \\ \delta_1 = \delta_2 \,(\mathrm{mod}\, 2)}} \exp\left(- \frac{\left(\delta_1 - 2^{n-1}\right)^2 + \left(\delta_2 - 2^{n-1}\right)^2}{2^{n-1}(1+|t|)} \right) \\
&= \left(\sum_{\substack{2^{n-1}-2^{(n-1)\frac{5}{8}} \le \delta_1 \le \delta \\ \delta_1 \text{even}}} \exp\left(-\frac{(\delta_1 - 2^{n-1})^2}{2^{n-1}(1+|t|)} \right) \right)^2 \\
&\quad + \left(\sum_{\substack{2^{n-1}-2^{(n-1)\frac{5}{8}} \le \delta_1 \le \delta \\ \delta_1 \text{odd}}} \exp\left(-\frac{(\delta_1 - 2^{n-1})^2}{2^{n-1}(1+|t|)} \right) \right)^2 .
\end{aligned}$$

According to (2.11), we have

$$\begin{aligned}
\sum_{\substack{2^{n-1}-2^{(n-1)\frac{5}{8}} \le \delta_1 \le \delta \\ \delta_1 \text{even}}} \exp\left(-\frac{(\delta_1 - 2^{n-1})^2}{2^{n-1}(1+|t|)} \right) &= \frac{\sqrt{2^{n-1}(1+|t|)}}{2} \times (1+o(1)) \\
&\quad \times \mathcal{H}\left(-\frac{c\sqrt{\binom{n}{r}\log 2}}{\sqrt{1+|t|}} \right),
\end{aligned}$$

and this estimation holds for odd values of δ_1. We have thus proved that

$$\begin{aligned}
\sum_3 &\le \frac{2^{2^n+1}}{\pi\, 2^{n-1}\sqrt{1-t^2}} \cdot 2 \cdot \frac{2^{n-1}(1+|t|)}{4} \cdot \mathcal{H}^2\left(-\frac{c\sqrt{\binom{n}{r}\log 2}}{\sqrt{1+|t|}} \right) \cdot (1+o(1)) \\
&= \frac{2^{2^n}}{\pi} \cdot \sqrt{\frac{1+|t|}{1-|t|}} \cdot \mathcal{H}^2\left(-\frac{c\sqrt{\binom{n}{r}\log 2}}{\sqrt{1+|t|}} \right) \cdot (1+o(1)) \\
&= \frac{2^{2^n}}{\pi} \cdot \sqrt{\frac{1+|t|}{1-|t|}} \cdot \frac{2^{-2c^2\binom{n}{r}/1+|t|}}{4c^2\binom{n}{r}\log 2/1+|t|} \cdot (1+o(1)),
\end{aligned} \tag{2.19}$$

by (2.5). To estimate $\sum_4$ we proceed the same way except that we use (2.3) to bound each binomial coefficient. Though (2.3) provides a much greater bound than (2.4), we end up with a much smaller bound, because this is widely compensated for by having δ_1 smaller than $2^{n-1} - 2^{(n-1)\frac{5}{8}}$.

$$\sum_4 \le 2 \cdot 2^{2^n} \cdot \sum_{\substack{\delta_1 < 2^{n-1} - 2^{(n-1)\frac{5}{8}}\,;\, \delta_2 \le \delta \\ \delta_1 = \delta_2 \,(\mathrm{mod}\, 2)}} \exp\left(- \frac{\left(\delta_1 - 2^{n-1}\right)^2 + \left(\delta_2 - 2^{n-1}\right)^2}{2^{n-1}(1+|t|)} \right).$$

And according to (2.11)

$$\begin{aligned}
&\sum_{\substack{\delta_1 < 2^{n-1} - 2^{(n-1)\frac{5}{8}}\,;\, \delta_2 \le \delta \\ \delta_1,\, \delta_2 \text{ even}}} \exp\left(- \frac{\left(\delta_1 - 2^{n-1}\right)^2 + \left(\delta_2 - 2^{n-1}\right)^2}{2^{n-1}(1+|t|)} \right) \\
&\quad = \sum_{\substack{\delta_1 < 2^{n-1} - 2^{(n-1)\frac{5}{8}} \\ \delta_1 \text{even}}} \exp\left(-\frac{\left(\delta_1 - 2^{n-1}\right)^2}{2^{n-1}(1+|t|)} \right) \sum_{\substack{\delta_2 \le \delta \\ \delta_2 \text{even}}} \exp\left(-\frac{\left(\delta_2 - 2^{n-1}\right)^2}{2^{n-1}(1+|t|)} \right) \\
&\quad = \frac{2^{n-1}(1+|t|)}{4} \cdot \mathcal{H}\left(-\frac{c\sqrt{\binom{n}{r}\log 2}}{\sqrt{1+|t|}} \right) \cdot \mathcal{H}\left(-\frac{2^{(n-1)\frac{1}{8}}}{\sqrt{1+|t|}} \right) \cdot (1+o(1)),
\end{aligned}$$

and the same estimation holds for odd values of δ_1 and δ_2. Thus

$$\begin{aligned}
\sum_4 &\le 2^{2^n+n-1} \cdot (1+|t|) \cdot \mathcal{H}\left(-\frac{c\sqrt{\binom{n}{r}\log 2}}{\sqrt{1+|t|}} \right) \cdot \mathcal{H}\left(-\frac{2^{(n-1)\frac{1}{8}}}{\sqrt{1+|t|}} \right) \cdot (1+o(1)) \\
&= O\left(2^{2^n+n-\left(c^2\binom{n}{r}+2^{n/4}\right)/(1+|t|)} \right),
\end{aligned}$$

by (2.5), and is much less than (2.19). Thus, we obtain (2.16). Now, when $|t| = o\left(\frac{1}{\binom{n}{r}}\right)$, we get (2.17) straightforwardly from (2.7). Finally, we note that when $t = 0$, (2.18) is tight, and then also (2.19) and (2.17).

□

In our estimation of (2.14), we shall use (2.17) when $t = 0$, or $t = \pm 2^{-h}$ so that $2^{-h} \le n^{-3}$, hence for $w = 2^{n-1}$ or $w = 2^{n-1}(1 \pm 2^{-h})$ where $h \ge 3\log_2 n$.

2.5 A bound on the total size of intersections

When $r = 1$, we have by (2.17)

$$E(\eta^2) = E(\eta) + \left(2^{1+n}\right) \cdot \left(2^{1+n} - 2\right) \cdot \frac{1}{\pi} \cdot \mathcal{H}^2\left(-c\sqrt{n\log 2}\right) \cdot (1+o(1)).$$

But for $\mathcal{RM}(2,n)$, it will take few more steps. Let g_h^+ (resp. g_h^-) be any codeword of weight $2^{n-1}(1+2^{-h})$ (resp. $2^{n-1}(1-2^{-h})$). We recall that $A_{2^{n-1}(1+2^{-h})} = A_{2^{n-1}(1-2^{-h})}$. Then

$$\sum_{w\neq 0} A_w \cdot P(\mathcal{B}_\delta[0] \cap \mathcal{B}_\delta[g_w]) = \mathcal{S}_1(h) + \mathcal{S}_2(h),$$

where

$$\begin{aligned}\mathcal{S}_1(h) = &\sum_{3\log_2 n \le h \le \lfloor \frac{n}{2} \rfloor} A_{2^{n-1}(1+2^{-h})} \cdot \left(P(\mathcal{B}_\delta[0] \cap \mathcal{B}_\delta[g_h^+]) + P(\mathcal{B}_\delta[0] \cap \mathcal{B}_\delta[g_h^-])\right) \\ &+ A_{2^{n-1}} \cdot P(\mathcal{B}_\delta[0] \cap \mathcal{B}_\delta[g_{2^{n-1}}])\end{aligned}$$

and

$$\mathcal{S}_2(h) = \sum_{0<h<3\log_2 n} A_{2^{n-1}(1+2^{-h})} \cdot \left(P(\mathcal{B}_\delta[0] \cap \mathcal{B}_\delta[g_h^+]) + P(\mathcal{B}_\delta[0] \cap \mathcal{B}_\delta[g_h^-])\right).$$

Lemma 2.5.1. *Let $\delta = 2^{n-1} - c\sqrt{2^{n-1}\binom{n}{2}\log 2}$ where c is a positive constant. We have*

$$\mathcal{S}_1(h) \le 2^{1+n+\binom{n}{2}} \cdot \frac{1}{\pi} \cdot \mathcal{H}^2\left(-c\sqrt{\tbinom{n}{2}\log 2}\right) \cdot (1+o(1)) \tag{2.20}$$

Proof. According to (2.17)

$$\begin{aligned}\mathcal{S}_1(h) \le &\left(2^{1+n+\binom{n}{2}} - 2\sum_{0\le h<3\log_2 n} A_{2^{n-1}(1+2^{-h})}\right) \times \frac{1}{\pi} \\ &\times \mathcal{H}^2\left(-c\sqrt{\tbinom{n}{2}\log 2}\right) \times (1+o(1)).\end{aligned}$$

What we are about to find out is that almost all codewords of $\mathcal{RM}(2,n)$ have weights setting close to 2^{n-1}. Knowing the weight distribution of the code, this property can be easily found. By (2.13), we have

$$\begin{aligned}&\sum_{0\le h<3\log_2 n} A_{2^{n-1}(1+2^{-h})} \\ &< \frac{3}{2} \cdot \sum_{0\le h<3\log_2 n} 2^{2h(n-h)+h} \\ &= \frac{3}{2} \cdot \sum_{0\le h<3\log_2 n} 2^{-2\left(\left(h-\frac{2n+1}{4}\right)^2 - \frac{(2n+1)^2}{16}\right)} \\ &= \frac{3}{2} \cdot 2^{\frac{(2n+1)^2}{8}} \cdot \sum_{0\le h<3\log_2 n} \exp\left(-2\log 2\left(h-\frac{2n+1}{4}\right)^2\right).\end{aligned}$$

Using the right-hand inequality of (2.8), we get

$$\begin{aligned}\sum_{0\le h<3\log_2 n} \exp\left(-2\log 2\left(h-\frac{2n+1}{4}\right)^2\right) &\le \mathcal{H}\left(\sqrt{2\log 2}\left(3\log_2 n-\frac{2n+1}{4}+1\right)\right)\\ &= O\left(2^{-\frac{(2n+1)^2}{8}+3(2n+1)\log_2 n}\right),\end{aligned}$$

by (2.5). Hence,

$$\begin{aligned}\sum_{0\le h<3\log_2 n} A_{2^{n-1}(1+2^{-h})} &= O\left(2^{3(2n+1)\log_2 n}\right) \qquad (2.21)\\ &= o\left(2^{1+n+\binom{n}{2}}\right).\end{aligned}$$

□

Lemma 2.5.2. *Let* $\delta = 2^{n-1} - c\sqrt{2^{n-1}\binom{n}{2}\log 2}$ *where* $0 < c < 1$. *We have*

$$\sum_{w\neq 0} A_w \cdot P(\mathcal{B}_\delta[0]\cap\mathcal{B}_\delta[g_w]) \le 2^{1+n+\binom{n}{2}}\cdot\frac{1}{\pi}\cdot\mathcal{H}^2\left(-c\sqrt{\binom{n}{2}\log 2}\right)\cdot(1+o(1))$$

Proof. One can see that the intersection between two Hamming balls is larger when the distance between their centers is smaller. As a result

$$\mathcal{S}_2(h) \le 2\cdot P(\mathcal{B}_\delta[0]\cap\mathcal{B}_\delta[g_1^-])\cdot\sum_{0<h<3\log_2 n} A_{2^{n-1}(1+2^{-h})}$$

where g_1^- is a codeword having the smallest possible non-zero weight, that is $2^{n-1}(1-2^{-1})$. According to proposition 2.4.2,

$$\begin{aligned}P(\mathcal{B}_\delta[0]\cap\mathcal{B}_\delta[g_1^-]) &\le \frac{\sqrt{3}}{\pi}\cdot\mathcal{H}^2\left(-\frac{c\sqrt{\binom{n}{2}\log 2}}{\sqrt{3/2}}\right)\cdot(1+o(1))\\ &= O\left(2^{-\frac{4c^2}{3}\binom{n}{2}}\right).\end{aligned}$$

Using (2.21), we obtain

$$\begin{aligned}\mathcal{S}_2(h) &= O\left(2^{-\frac{4c^2}{3}\binom{n}{2}+3(2n+1)\log_2 n}\right)\\ &= o\left(2^{(1-2c^2)\binom{n}{2}}\right),\end{aligned}$$

provided that $c < 1$. Taking into account (2.20), we obtain the required result. □

2.6 An asymptotic upper bound

Putting all together, we proved that for $r \leq 2$ and $c > 0$

$$E(\eta) = 2^{\sum_{i=0}^{r} \binom{n}{i}} \cdot \frac{1}{\sqrt{\pi}} \cdot \mathcal{H}\left(-c\sqrt{\tbinom{n}{r}\log 2}\right) \cdot (1 + o(1))$$

$$E(\eta^2) \leq E(\eta) + 2^{2\cdot\sum_{i=0}^{r} \binom{n}{i}} \cdot \frac{1}{\pi} \cdot \mathcal{H}^2\left(-c\sqrt{\tbinom{n}{r}\log 2}\right) \cdot (1 + o(1)) \quad \text{if} \quad c < 1.$$

Thus

$$E(\eta^2) \leq E(\eta) + E^2(\eta)(1 + o(1)) \quad \text{if} \quad c < 1. \tag{2.22}$$

Finally, we can state our result.

Theorem 2.6.1. *Let* $\delta = 2^{n-1} - c\sqrt{2^{n-1}\binom{n}{r}\log 2}$ *where* $0 < c < 1$ *and* $r \leq 2$. *Then*

$$P\left(\mathcal{NL}_r(f) > \delta\right) \underset{n\to+\infty}{\longrightarrow} 0.$$

Proof. We have

$$P\left(\mathcal{NL}_r(f) > \delta\right) = P(\eta = 0) \leq \frac{E(\eta^2) - E^2(\eta)}{E^2(\eta)}.$$

And by (2.22)

$$\frac{E(\eta^2) - E^2(\eta)}{E^2(\eta)} \leq \frac{1}{E(\eta)} + o(1).$$

Or

$$E(\eta) = 2^{\sum_{i=0}^{r} \binom{n}{i} - c^2\binom{n}{r}} \times \frac{(1 + o(1))}{2c\sqrt{\binom{n}{r}\pi\log 2}}$$

tends to infinity with n, if $c < 1$. □

2.7 Asymptotic r-th order nonlinearity

We proved that $2^{n-1} - \sqrt{2^{n-1}\binom{n}{2}\log 2}$ represents a concentration point of the second-order nonlinearity of Boolean functions, when n tends to infinity. The situation gets critical with our method when dealing with Reed-Muller codes of higher orders because of the little knowledge we have about their weight distribution. One can prove that almost all codewords have weights setting around 2^{n-1}, but the problem arises regarding the sum taken over the weights far from 2^{n-1}, that is $\mathcal{S}_2(h)$, that we were unable to enhance in order to give the needed precisions. Recently, Schmidt [Sch13] extended this result for every $r \geq 1$, proving that the r-th order nonlinearities of almost all Boolean functions lie close to $2^{n-1} - \sqrt{2^{n-1}\binom{n}{r}\log 2}$.

Chapter 3

Decoding of first order Reed-Muller codes

3.1 Introduction

Let $\mathbb{F}_2^{2^n}$ be the set of all binary vectors of length 2^n. For any vector $f \in \mathbb{F}_2^{2^n}$, the set

$$f + \mathcal{RM}(1,n) = \{f + g \mid g \in \mathcal{RM}(1,n)\}$$

is called a coset of $\mathcal{RM}(1,n)$ and contains 2^{n+1} vectors. One can easily check that two cosets are either disjoint or coincide. This means

$$\text{if } f \in f' + \mathcal{RM}(1,n) \Longrightarrow f + \mathcal{RM}(1,n) = f' + \mathcal{RM}(1,n).$$

Therefore, the set $\mathbb{F}_2^{2^n}$ can be partitioned into $2^{2^n-(n+1)}$ cosets of $\mathcal{RM}(1,n)$:

$$\mathbb{F}_2^{2^n} = \bigcup_{i=0}^{2^{2^n-(n+1)}-1} (f_i + \mathcal{RM}(1,n)), \qquad f_i \in \mathbb{F}_2^{2^n}$$

where $(f_i + \mathcal{RM}(1,n)) \cap (f_j + \mathcal{RM}(1,n)) = \emptyset$ for $i \neq j$. Suppose the decoder receives the vector h, h belongs to the coset $h + \mathcal{RM}(1,n)$. If the codeword g was sent, then the error vector that have occured is $(h - g) \in h + \mathcal{RM}(1,n)$. Thus, the possible error vectors are the vectors in the coset containing h. In maximum-likelihood decoding, the decoder's strategy is, given h, to choose a minimum weight vector f in $h + \mathcal{RM}(1,n)$, and to decode h as $h - f$. The minimum weight vector in a coset is called the coset leader, and when there is more than one vector of minimum weight in a coset, any one of them can be selected as the coset leader. Only coset leaders are correctable errors, which means that $2^{2^n-(n+1)}$ errors can be corrected with this decoding. Since the minimum distance of $\mathcal{RM}(1,n)$ is 2^{n-1}, a given error f of weight $\leq (2^{n-2}-1)$ is a coset leader (i.e. correctable error) because $\forall g \in \mathcal{RM}(1,n) \setminus \{0\}$:

$$w_H(f+g) \geq w_H(g) - w_H(f) \geq 2^{n-1} - (2^{n-2}-1) = 2^{n-2} + 1 > w_H(f)$$

and thus f is the only minimum weight vector in $f+\mathcal{RM}(1,n)$, which makes it a coset leader. On the other hand, every error of weight larger than the covering radius is uncorrectable since the covering radius is the largest weight of a correctable error.

Whenever we refer to results from the previous chapter, we use the correspondance between the binary vectors of length 2^n and the Boolean functions of n variables, and hope there will be no ambiguity. Let $f \in \mathbb{F}_2^{2^n}$, f is associated to a n-variable Boolean function. We have

$$\begin{aligned}\mathcal{NL}(f) &= \min_{g\in\mathcal{RM}(1,n)} d_H(f,g) \\ &= \min_{g\in\mathcal{RM}(1,n)} w_H(f+g),\end{aligned}$$

which is the minimum weight of the words of $f + \mathcal{RM}(1,n)$, and thus the weight of the coset leader. That being said, the associated result to that of the first chapter from a coding view, is that almost all coset leaders of $\mathcal{RM}(1,n)$ have weights lying close to $2^{n-1} - \sqrt{2^{n-1}n\log 2}$, when n tends to infinity. In other words, the distribution of the nonlinearity of Boolean functions coincides with the weight distribution of the coset leaders of $\mathcal{RM}(1,n)$. In this chapter, we study more closely the weight distribution of coset leaders. Our aim is to find the largest weight t such that the fraction of all errors of weight t that are correctable (i.e. coset leaders) tends to 1 when n tends to infinity. Note that t is somewhere between half the minimum distance of the code and its covering radius

$$2^{n-2} \leq t \leq 2^{n-1} - 2^{n/2-1}.$$

Let $\delta = 2^{n-1} - c\sqrt{2^{n-1}n\log 2}$ where c is a positive real number. We prove that the code can correct almost all errors of weight smaller than δ if $c > 1$ (theorem 3.2.1), and fails to correct almost all errors of weight larger than δ if $c < 1$ (theorem 3.3.1). This means that the limit (upper bound) established for the asymptotic (first order) nonlinearity of Boolean functions is the same for the number of almost always corrected errors in maximum-likelihood decoding of $\mathcal{RM}(1,n)$. Proving the lower bound for the number of almost always corrected errors is an easy task considering the estimations already given in the previous chapter, but proving that this bound is optimal turns out to be more difficult. This result is not new, however, our method underlines the link between the nonlinearity of Boolean functions and decoding of $\mathcal{RM}(1,n)$. When decoding the $[m = 2^n, k = n+1, d_{\min} = 2^{n-1}]$ code $\mathcal{RM}(1,n)$, Helleseth, Klove and Levenshtein [HKL05] proved that, for $0 < \epsilon < 1$, the largest t such that the fraction of all errors of weight t that are correctable is $\geq \epsilon$ equals

$$\frac{m}{2} - \frac{1}{2}\sqrt{\frac{m-d_{\min}}{d_{\min}}\, m\, k \log 4},$$

which meets our bound. We start by proving the lower bound which came directly from the estimations we derived in the previous chapter.

3.2 Decoding a large number of errors

We consider $\mathbb{F}_2^{2^n}$ as a metric space with the Hamming distance, weight and balls introduced in the previous chapter. Let $f \in \mathcal{B}_\delta[0]$. If f is an uncorrectable error, then either f is not a minimum weight vector in $f+\mathcal{RM}(1,n)$, or is a minimum weight vector but not the coset leader. In both cases,

$$\exists g \in \mathcal{RM}(1,n) \setminus \{0\}, \quad w_H(f) \geq w_H(f+g)$$

$\Leftrightarrow$

$$\exists g \in \mathcal{RM}(1,n) \setminus \{0\}, \quad \delta \geq d_H(f,0) \geq d_H(f,g)$$

$\Leftrightarrow$

$$\exists g \in \mathcal{RM}(1,n) \setminus \{0\}, \quad f \in \mathcal{B}_\delta[0] \cap \mathcal{B}_\delta[g].$$

Hence, the number of uncorrectable errors of weight smaller than δ is less than

$$\Big| \bigcup_{g \in \mathrm{RM}(1,n)\setminus\{0\}} (\mathcal{B}_\delta[0] \cap \mathcal{B}_\delta[g]) \Big|,$$

which is smaller than

$$\sum_{g \in \mathrm{RM}(1,n)\setminus\{0\}} |\mathcal{B}_\delta[0] \cap \mathcal{B}_\delta[g]| .$$

We show in the following theorem that the latter quantity is negligible with respect to the number of vectors in $\mathcal{B}_\delta[0]$, if $c > 1$.

Theorem 3.2.1. *Let $\delta = 2^{n-1} - c\sqrt{2^{n-1} n \log 2}$ where c is a positive real number. The fraction of all errors of weight smaller than δ that are uncorrectable is less than*

$$\frac{2^{n+1}}{\sqrt{\pi}} \cdot \mathcal{H}(-c\sqrt{n \log 2}) \cdot (1 + o(1))$$

and, if $c > 1$, it tends to 0 when n tends to ∞.

Proof. The fraction of all errors of weight smaller than δ that are uncorrectable is less than

$$\sum_{g \in \mathrm{RM}(1,n)\setminus\{0\}} \frac{|\mathcal{B}_\delta[0] \cap \mathcal{B}_\delta[g]|}{|\mathcal{B}_\delta[0]|} = (2^{n+1} - 2) \cdot \frac{1}{\sqrt{\pi}} \cdot \mathcal{H}\left(-c\sqrt{n \log 2}\right) \cdot (1 + o(1))$$

$$= O\left(2^{n(1-c^2)}\right),$$

using (2.17) and proposition 2.3.1. This tends to 0 when n tends to ∞, if $c > 1$. □

3.3 An asymptotic decoding upper bound

We prove in this section that the bound we gave in the previous section is optimal. Indeed, when the number of errors exceeds δ with $c < 1$, maximum-likelihood decoding fails to correct. For a given vector g and an integer ρ, we introduce the Hamming sphere $\mathcal{S}_\rho[g]$ with centre g and radius ρ defined by

$$\mathcal{S}_\rho[0] = \{f \in \mathbb{F}_2^{2^n} \mid d_H(f, 0) = \rho\}.$$

Let us fix the number of errors to be d such that

$$d = 2^{n-1} - \alpha(n)\sqrt{2^{n-1}}$$

where $\alpha(n)$ tends to infinity and $d \geq \delta$. Let $f \in \mathcal{S}_d[0]$. If f is a correctable error, then f is definitely a minimum weight vector in $f + \mathcal{RM}(1, n)$, which gives

$$\forall g \in \mathcal{RM}(1, n) \setminus \{0\}, \quad w_H(f, 0) \leq w(f + g)$$

$\Leftrightarrow$

$$\forall g \in \mathcal{RM}(1, n) \setminus \{0\}, \quad d = d_H(f, 0) \leq d_H(f, g)$$

$\Leftrightarrow$

$$\forall g \in \mathcal{RM}(1, n) \setminus \{0\}, \quad f \notin \mathcal{B}_{d-1}[g].$$

Hence

$$f \text{ is a correctable error of weight } d \implies \sum_{g \in \mathcal{RM}(1,n)\setminus\{0\}} 1_{\mathcal{B}_{d-1}[g]}(f) = 0$$

where $1_{\mathcal{B}_{d-1}[g]}$ is the indicator function of the subset $\mathcal{B}_{d-1}[g]$. Let θ be an integer-valued function on $\mathcal{S}_d[0]$ defined by

$$\theta(f) = \sum_{g \in \mathcal{RM}(1,n)\setminus\{0\}} 1_{\mathcal{B}_{d-1}[g]}(f).$$

Therefore, the fraction of all errors of weight d that are correctable is less than $P(\theta = 0)$. Using the second moment method

$$P(\theta = 0) \leq \frac{E(\theta^2) - E^2(\theta)}{E^2(\theta)}$$

as we did in the previous chapter, we prove that it tends to 0 when n tends to infinity, if $c < 1$.

3.3.1 Expected value of θ

Because of the linearity of E, we have

$$\begin{aligned} E(\theta) &= \sum_{g \in \mathcal{RM}(1,n)\setminus\{0\}} E\left(1_{\mathcal{B}_{d-1}[g]} \mid \mathcal{S}_d[0]\right) \\ &= \sum_{g \in \mathcal{RM}(1,n)\setminus\{0\}} P\left(\mathcal{B}_{d-1}[g] \mid \mathcal{S}_d[0]\right) \\ &= \sum_{g \in \mathcal{RM}(1,n)\setminus\{0\}} \frac{P\left(\mathcal{S}_d[0] \cap \mathcal{B}_{d-1}[g]\right)}{P\left(\mathcal{S}_d[0]\right)}. \end{aligned}$$

Since the size of intersection between $\mathcal{S}_d[0]$ and $\mathcal{B}_{d-1}[g]$ depends only on the weight of g, and that it is null if $g = 1\,(d < 2^{n-1})$, we get

$$E(\theta) = (2^{1+n} - 2) \cdot \frac{P\left(\mathcal{S}_d[0] \cap \mathcal{B}_{d-1}[g]\right)}{P\left(\mathcal{S}_d[0]\right)}$$

where g is any codeword of weight 2^{n-1}.

Proposition 3.3.1. *Let $d = 2^{n-1} - \alpha(n)\sqrt{2^{n-1}}$ where $\alpha(n)$ tends to infinity and $d \geq \delta$. The fraction of errors of weight d in $\mathbb{F}_2^{2^n}$ is*

$$P(\mathcal{S}_d[0]) = \frac{\exp(-\alpha^2(n))}{\sqrt{\pi\, 2^{n-1}}}\,(1 + o(1)).$$

Proof. The number of vectors of weight d is $\binom{2^n}{d}$ and the result is obtained by (2.4). □

Lemma 3.3.1. *Let $d = 2^{n-1} - \alpha(n)\sqrt{2^{n-1}}$ where $\alpha(n)$ tends to infinity and $d \geq \delta$. Let g be a vector of weight 2^{n-1}. We have*

$$P(\mathcal{S}_d[0] \cap \mathcal{B}_{d-1}[g]) = \frac{\exp(-\alpha^2(n))}{\pi\,\sqrt{2^{n-1}}} \cdot \mathcal{H}(-\alpha(n)) \cdot (1 + o(1)).$$

Proof. Due to proposition 2.4.1, we have

$$P(\mathcal{S}_d[0] \cap \mathcal{B}_{d-1}[g]) = 2^{-2^n} \sum_{0 \leq d_1 \leq d-1} \binom{2^{n-1}}{\frac{d-d_1+2^{n-1}}{2}} \binom{2^{n-1}}{\frac{d+d_1-2^{n-1}}{2}}.$$

Whenever the binomial coefficient is not null, d_1 has the same parity as d. We divide the sum into two sums and estimate each one separatly. Let

$$\sum\nolimits_1 = \sum_{2^{n-1}-2^{(n-1)\frac{5}{8}} \leq d_1 \leq d-1} \binom{2^{n-1}}{\frac{d-d_1+2^{n-1}}{2}} \binom{2^{n-1}}{\frac{d+d_1-2^{n-1}}{2}},$$

and

$$\sum\nolimits_2 = \sum_{0 \le d_1 < 2^{n-1} - 2^{(n-1)\frac{5}{8}}} \binom{2^{n-1}}{\frac{d-d_1+2^{n-1}}{2}} \binom{2^{n-1}}{\frac{d+d_1-2^{n-1}}{2}}.$$

By (2.4), we have

$$\sum\nolimits_1 = \frac{2^{2^n}}{\pi\, 2^{n-2}} \cdot (1 + o(1)) \cdot \sum_{\substack{2^{n-1} - 2^{(n-1)\frac{5}{8}} \le d_1 \le d-1 \\ d_1 = d \,(\mathrm{mod}\, 2)}} \exp\left(- \frac{\left(\frac{d-d_1}{2}\right)^2}{2^{n-2}} \right) \cdot \exp\left(- \frac{\left(\frac{d+d_1-2^{n-1}}{2} - 2^{n-2}\right)^2}{2^{n-2}} \right).$$

Rearranging gives

$$\begin{aligned} \sum\nolimits_1 &= \frac{2^{2^n}}{\pi\, 2^{n-2}} \cdot (1 + o(1)) \cdot \sum_{\substack{2^{n-1} - 2^{(n-1)\frac{5}{8}} \le d_1 \le d-1 \\ d_1 = d \,(\mathrm{mod}\, 2)}} \exp\left(- \frac{\left(d - 2^{n-1}\right)^2 + \left(d_1 - 2^{n-1}\right)^2}{2^{n-1}} \right) \\ &= \frac{2^{2^n} \cdot \exp(-\alpha^2(n))}{\pi\, 2^{n-2}} \cdot (1 + o(1)) \cdot \sum_{\substack{2^{n-1} - 2^{(n-1)\frac{5}{8}} \le d_1 \le d-1 \\ d_1 = d \,(\mathrm{mod}\, 2)}} \exp\left(- \frac{\left(d_1 - 2^{n-1}\right)^2}{2^{n-1}} \right). \end{aligned}$$

Using (2.11), we get

$$\begin{aligned} \sum\nolimits_1 &= \frac{2^{2^n} \cdot \exp(-\alpha^2(n))}{\pi\, 2^{n-2}} \cdot \frac{\sqrt{2^{n-1}}}{2} \cdot \mathcal{H}(-\alpha(n)) \cdot (1 + o(1)) \\ &= \frac{2^{2^n} \cdot \exp(-\alpha^2(n))}{\pi \sqrt{2^{n-1}}} \cdot \mathcal{H}(-\alpha(n)) \cdot (1 + o(1)) \\ &\ge \frac{2^{2^n} \cdot \exp(-\alpha^2(n))}{\pi \sqrt{2^{n-1}}} \cdot \mathcal{H}(-c\sqrt{n \log 2}) \cdot (1 + o(1)), \end{aligned}$$

since $\mathcal{H}(y)$ is an increasing function and $d \ge \delta$

$$2^{n-1} - \alpha(n)\sqrt{2^{n-1}} \ge 2^{n-1} - c\sqrt{2^{n-1} n \log 2}. \tag{3.1}$$

To estimate $\sum_2$, we use (2.3) to bound each binomial coefficient and proceed

as before

$$\begin{aligned}\sum\nolimits_2 &< 2^{2^n}\cdot\exp(-\alpha^2(n))\cdot\sum_{\substack{d_1\le 2^{n-1}-2^{(n-1)\frac{5}{8}}\\ d_1=d\,(\mathrm{mod}\,2)}}\exp\left(-\frac{(d_1-2^{n-1})^2}{2^{n-1}}\right)\\ &< 2^{2^n}\cdot\exp(-\alpha^2(n))\cdot\frac{\sqrt{2^{n-1}}}{2}\cdot\mathcal{H}\left(-2^{(n-1)\frac{1}{8}}\right)\cdot(1+o(1))\\ &= 2^{2^n}\cdot\exp(-\alpha^2(n))\cdot O\left(2^{-2^{n/4}+n/2}\right),\end{aligned}$$

by (2.11), which is much less then $\sum_1$. □

3.3.2 Expected value of θ^2

We have

$$\begin{aligned}E(\theta^2) &= \sum_{\substack{g_1,g_2\in\mathcal{RM}(1,n)\setminus\{0\}\\ g_1\neq g_2}} E\left(1_{\mathcal{B}_{d-1}[g_1]\cap\mathcal{B}_{d-1}[g_2]}\mid\mathcal{S}_d[0]\right)+E(\theta)\\ &= \sum_{\substack{g_1,g_2\in\mathcal{RM}(1,n)\setminus\{0\}\\ g_1\neq g_2}}\frac{P(\mathcal{S}_d[0]\cap\mathcal{B}_{d-1}[g_1]\cap\mathcal{B}_{d-1}[g_2])}{P(\mathcal{S}_d[0])}+E(\theta).\end{aligned}$$

The intersection is null if one of g_i equals 1 or $g_1+g_2=1$. Thus

$$\begin{aligned}E(\theta^2) &= (2^{n+1}-2)(2^{n+1}-4)\cdot\frac{P(\mathcal{S}_d[0]\cap\mathcal{B}_{d-1}[g_1]\cap\mathcal{B}_{d-1}[g_2])}{P(\mathcal{S}_d[0])}+E(\theta)\\ &\le (2^{n+1}-2)^2\cdot\frac{P(\mathcal{S}_d[0]\cap\mathcal{B}_{d-1}[g_1]\cap\mathcal{B}_{d-1}[g_2])}{P(\mathcal{S}_d[0])}+E(\theta),\end{aligned}$$

where g_1 and g_2 are any codewords of weights 2^{n-1} such that $g_1+g_2\neq 1$.

Proposition 3.3.2. *Let Δ, Δ_1 and Δ_2 be positive integers. Let $g_1,g_2\in\mathbb{F}_2^{2^n}$ be two vectors of weight 2^{n-1} and with Hamming distance between them 2^{n-1}. Then the size of the set $(\mathcal{S}_\Delta[0]\cap\mathcal{S}_{\Delta_1}[g_1]\cap\mathcal{S}_{\Delta_2}[g_2])$ equals*

$$\begin{aligned}\sum_{0\le a_{1,1}\le 2^{n-2}}&\binom{2^{n-2}}{a_{1,1}+\frac{-2^n+\Delta_1+\Delta_2}{2}}\times\binom{2^{n-2}}{-a_{1,1}+\frac{2^{n-1}+\Delta-\Delta_2}{2}}\\ &\times\binom{2^{n-2}}{-a_{1,1}+\frac{2^{n-1}+\Delta-\Delta_1}{2}}\times\binom{2^{n-2}}{a_{1,1}}.\end{aligned}$$

Proof. A vector f belongs to $(\mathcal{S}_\Delta[0]\cap\mathcal{S}_{\Delta_1}[g_1]\cap\mathcal{S}_{\Delta_2}[g_2])$ if and only if

$$\begin{cases} d_H(f,0)=\Delta\\ d_H(f,g_1)=\Delta_1\\ d_H(f,g_2)=\Delta_2.\end{cases}\tag{3.2}$$

Now, in order to count the number of vectors satisfying these conditions, we look at $A = \{1 \leq i \leq 2^n\}$ as the union of the following disjoint sets

$$A_{j,k} = \{i \in A \mid g_{1,i} = j,\ g_{2,i} = k\}, \qquad j, k = 0 \text{ or } 1$$

and find the number of possible ways to choose the Δ nonzero coordinates of f from these sets, while respecting its distance to g_1 (resp. g_2), that is Δ_1 (resp. Δ_2). We have

$$|A_{0,0}| = |A_{0,1}| = |A_{1,0}| = |A_{1,1}| = 2^{n-2},$$

since

$$|A_{0,1}| + |A_{1,1}| = |A_{1,0}| + |A_{1,1}| = |A_{0,1}| + |A_{1,0}| = |A_{0,0}| + |A_{1,1}| = 2^{n-1},$$

because g_1 and g_2 have weights 2^{n-1} and agree (resp. disagree) 2^{n-1} times. Now, we consider

$$a_{j,k} = |\{i \in A_{j,k} \mid f_i = 1\}| \qquad j, k = 0 \text{ or } 1.$$

Then (3.2) becomes

$$\begin{cases} \Delta = a_{0,0} + a_{0,1} + a_{1,0} + a_{1,1} \\ \Delta_1 = a_{0,0} + a_{0,1} + |A_{1,0}| - a_{1,0} + |A_{1,1}| - a_{1,1} \\ \Delta_2 = a_{0,0} + a_{1,0} + |A_{0,1}| - a_{0,1} + |A_{1,1}| - a_{1,1}. \end{cases}$$

And the solution is

$$\begin{cases} a_{0,0} = a_{1,1} + \dfrac{-2^n + \Delta_1 + \Delta_2}{2} \\ a_{0,1} = -a_{1,1} + \dfrac{2^{n-1} + \Delta - \Delta_2}{2} \\ a_{1,0} = -a_{1,1} + \dfrac{2^{n-1} + \Delta - \Delta_1}{2}. \end{cases} \tag{3.3}$$

Thus, the number of vectors in $(\mathcal{S}_\Delta[0] \cap \mathcal{S}_{\Delta_1}[g_1] \cap \mathcal{S}_{\Delta_2}[g_2])$ is

$$\sum_{0 \leq a_{1,1} \leq 2^{n-2}} \binom{2^{n-2}}{a_{0,0}} \binom{2^{n-2}}{a_{0,1}} \binom{2^{n-2}}{a_{1,0}} \binom{2^{n-2}}{a_{1,1}}.$$

By definition, $a_{i,j}$ is a non-negative integer less than 2^{n-2}. Hence, for values of Δ, Δ_1 and Δ_2 that do not satisfy these conditions, the sum is null. □

Proposition 3.3.3. *Let $d = 2^{n-1} - \alpha(n)\sqrt{2^{n-1}}$ where $\alpha(n)$ tends to infinity and $d \geq \delta$. We have*

$$P(\mathcal{S}_d[0] \cap \mathcal{B}_{d-1}[g_1] \cap \mathcal{B}_{d-1}[g_2]) < \frac{\exp(-\alpha^2(n))}{\pi^{3/2} \cdot \sqrt{2^{n-1}}} \cdot \mathcal{H}^2\left(-\alpha(n)\right) \cdot (1 + o(1)).$$

Proof. The size of the set $\mathcal{S}_d[0] \cap \mathcal{B}_{d-1}[g_1] \cap \mathcal{B}_{d-1}[g_2]$ is

$$\sum_{\substack{0\le d_1,\, d_2\le d-1\\ 0\le a_{1,1}\le 2^{n-2}}} \binom{2^{n-2}}{a_{0,0}}\binom{2^{n-2}}{a_{0,1}}\binom{2^{n-2}}{a_{1,0}}\binom{2^{n-2}}{a_{1,1}},$$

where the $a_{i,j}$ are given in (3.3). Whenever the binomial coefficient is not null, d_1 and d_2 have the same parity as d. We can use (2.4) to bound the binomial coefficients when $d_i \ge 2^{n-1} - \dfrac{2^{(n-3)\frac{5}{8}}}{2}$ and $|a_{1,1} - 2^{n-3}| \le \dfrac{2^{(n-3)\frac{5}{8}}}{2}$. Let us consider D_1 to be the domain covering these values. Then

$$\begin{aligned}
&\sum_{\mathrm{D}_1} \binom{2^{n-2}}{a_{0,0}}\binom{2^{n-2}}{a_{0,1}}\binom{2^{n-2}}{a_{1,0}}\binom{2^{n-2}}{a_{1,1}}\\
&= \left(\frac{2^{2^{n-2}}}{\sqrt{\pi 2^{n-3}}}\right)^4 \cdot (1+o(1)) \cdot \sum_{\mathrm{D}_1} \exp\Big(-\frac{1}{2^{n-3}}\big((a_{0,0}-2^{n-3})^2\\
&\qquad + (a_{0,1}-2^{n-3})^2 + (a_{1,0}-2^{n-3})^2 + (a_{1,1}-2^{n-3})^2\big)\Big).
\end{aligned}$$

And for any values of d_1, d_2 and $a_{1,1}$, we have

$$\begin{aligned}
&\sum_{\substack{d_1,\, d_2\\ a_{1,1}}} \exp\Big(-\frac{1}{2^{n-3}}\big((a_{0,0}-2^{n-3})^2 + (a_{0,1}-2^{n-3})^2 + (a_{1,0}-2^{n-3})^2\\
&\qquad + (a_{1,1}-2^{n-3})^2\big)\Big)\\
&= \exp(-\alpha^2(n)) \sum_{\substack{d_1,\, d_2\\ a_{1,1}}} \exp\left(-\frac{1}{2^{n-1}}\left((d_1-2^{n-1})^2 + (d_2-2^{n-1})^2\right)\right)\\
&\qquad \times \exp\left(-\frac{1}{2^{n-5}}\left(a_{1,1} + \frac{-d+d_1+d_2-2^n}{4}\right)^2\right)\\
&= \exp(-\alpha^2(n)) \sum_{d_1,\, d_2} \exp\left(-\frac{1}{2^{n-1}}\left((d_1-2^{n-1})^2 + (d_2-2^{n-1})^2\right)\right)\\
&\qquad \times \sum_{a_{1,1}} \exp\left(-\frac{1}{2^{n-5}}\left(a_{1,1} + \mathrm{cte}/4\right)^2\right), \qquad (3.4)
\end{aligned}$$

where the constant is independent of $a_{1,1}$. Now,

$$\begin{aligned}
\sum_{a_{1,1}} \exp\left(-\frac{1}{2^{n-5}}\left(a_{1,1}+\mathrm{cte}/4\right)^2\right) &< \sum_{x\in\mathbb{Z}+\mathrm{cte}/4} \exp\left(-\frac{x^2}{2^{n-5}}\right)\\
&< \int_{-\infty}^{\infty} \exp\left(-\frac{y^2}{2^{n-5}}\right)\mathrm{d}y + 2\\
&= \sqrt{2^{n-5}\pi} + 2.
\end{aligned}$$

Putting all together and using (2.11), we obtain

$$\begin{aligned}
&\sum_{\mathrm{D}_1} \binom{2^{n-2}}{a_{0,0}} \binom{2^{n-2}}{a_{0,1}} \binom{2^{n-2}}{a_{1,0}} \binom{2^{n-2}}{a_{1,1}} \\
&< \left(\frac{2^{2^{n-2}}}{\sqrt{\pi 2^{n-3}}} \right)^4 \times \exp(-\alpha^2(n)) \times \left(\sqrt{2^{n-5}\pi} + 2 \right) \times (1 + o(1)) \qquad (3.5) \\
&\quad \times \sum_{\substack{2^{n-1} - \frac{2^{(n-3)\frac{5}{8}}}{2} \le d_i \le d-1 \\ d_i = d \,(\mathrm{mod}\, 2)}} \exp\left(-\frac{1}{2^{n-1}} \left((d_1 - 2^{n-1})^2 + (d_2 - 2^{n-1})^2 \right) \right) \\
&< \frac{2^{2^n + (-3n+7)/2}}{\pi^{3/2}} \times \exp(-\alpha^2(n)) \times (1 + o(1)) \\
&\quad \times \left(\sum_{\substack{2^{n-1} - \frac{2^{(n-3)\frac{5}{8}}}{2} \le d_1 \le d-1 \\ d_1 = d \,(\mathrm{mod}\, 2)}} \exp\left(-\frac{\left(d_1 - 2^{n-1}\right)^2}{2^{n-1}} \right) \right)^2 \\
&= \frac{2^{2^n + (-3n+7)/2}}{\pi^{3/2}} \cdot \exp(-\alpha^2(n)) \cdot \left(\frac{\sqrt{2^{n-1}}}{2} \right)^2 \cdot \mathcal{H}^2\left(-\alpha(n)\right) \cdot (1 + o(1)) \\
&= \frac{2^{2^n} \cdot \exp(-\alpha^2(n))}{\pi^{3/2} \times \sqrt{2^{n-1}}} \times \mathcal{H}^2\left(-\alpha(n)\right) \times (1 + o(1)). \qquad (3.6)
\end{aligned}$$

Outside D_1, there are two cases to consider for which we use (2.3) to bound each binomial coefficient, which adds the factor $\left(\sqrt{\pi\, 2^{n-3}}\right)^4$ to (3.5). Then, either one of d_i is smaller than $2^{n-1} - \dfrac{2^{(n-3)\frac{5}{8}}}{2}$, and that replaces one $\mathcal{H}\left(-\alpha(n)\right)$ with $\mathcal{H}\left(-\dfrac{2^{(n-3)\frac{5}{8}}}{2\sqrt{2^{n-1}}} \right)$, which yields a much smaller bound than (3.6) due to (3.1), or $|a_{1,1} - 2^{n-3}|$ is greater than $\dfrac{2^{(n-3)\frac{5}{8}}}{2}$ and in this case, we go a little backward and give a more precise estimation of (3.4). One can check that if

$$2^{n-1} - \frac{2^{(n-3)\frac{5}{8}}}{2} \le d_i \le d \quad \text{and} \quad |a_{1,1} - 2^{n-3}| > \frac{2^{(n-3)\frac{5}{8}}}{2},$$

Then

$$\left| a_{1,1} + \frac{-d + d_1 + d_2 - 2^n}{4} \right| > \frac{2^{(n-3)\frac{5}{8}}}{4}.$$

Hence

$$\begin{aligned}
&\sum_{a_{1,1}} \exp\left(-\frac{1}{2^{n-5}}\left(a_{1,1}+\mathrm{cte}/4\right)^2\right)\\
&\quad=\sum_{a_{1,1}} \exp\left(-\frac{1}{2^{n-5}}\left(a_{1,1}+\frac{-d+d_1+d_2-2^n}{4}\right)^2\right)\\
&\quad\leq \sum_{x<-\frac{2^{(n-3)\frac{5}{8}}}{4}} \exp\left(-\frac{x^2}{2^{n-5}}\right)\\
&\quad=\sqrt{2^{n-5}}\cdot\mathcal{H}\left(-\frac{2^{(n-3)\frac{5}{8}}}{4\sqrt{2^{n-5}}}\right),
\end{aligned}$$

which replaces the factor $\left(\sqrt{2^{n-5}\pi}+2\right)$ in (3.5), and once again yields a much smaller bound than (3.6) due to (3.1). And that ends the proof.

□

Putting all together, we proved that

$$\begin{aligned}
E(\theta) &= (2^{n+1}-2)\cdot\frac{1}{\sqrt{\pi}}\cdot\mathcal{H}\left(-\alpha(n)\right)\cdot(1+o(1))\\
E(\theta^2) &< E(\theta)+(2^{n+1}-2)^2\cdot\frac{1}{\pi}\cdot\mathcal{H}^2\left(-\alpha(n)\right)\cdot(1+o(1)).
\end{aligned}$$

Thus

$$E(\theta^2)\leq E(\theta)+E^2(\theta)(1+o(1)). \tag{3.7}$$

Now, we can state our result.

Theorem 3.3.1. *Let $d=2^{n-1}-\alpha(n)\sqrt{2^{n-1}}$ where $\alpha(n)$ tends to infinity and $d\geq\delta$. The fraction of correctable erros of weight d tends to 0 when n tends to infinity, if $c<1$.*

Proof. The proof is the same as in section 2.6, but we repeat it for the sake of clarity. If f is a correctable error, then $\theta(f)=0$. Therefore, the fraction of all errors of weight d that are correctable is less than $P(\theta=0)$. And

$$P(\theta=0)\leq\frac{E(\theta^2)-E^2(\theta)}{E^2(\theta)}.$$

And by (3.7)

$$\frac{E(\theta^2)-E^2(\theta)}{E^2(\theta)}\leq\frac{1}{E(\theta)}+o(1).$$

Since $\mathcal{H}(y)$ is an increasing function and having (3.1), we get

$$\begin{aligned} E(\theta) &\geq (2^{n+1} - 2) \cdot \frac{1}{\sqrt{\pi}} \cdot \mathcal{H}\left(-c\sqrt{n \log 2}\right) \cdot (1 + o(1)) \\ &= 2^{1+n-c^2 n} \times \frac{(1 + o(1))}{2c\sqrt{n\pi \log 2}} \end{aligned}$$

which tends to infinity with n, if $c < 1$. □

Chapter 4

Asymptotic Nonlinearity of Vectorial Boolean Functions

4.1 Introduction

Let m and n be two positive integers. Functions from the vectorspace $V_m = \mathbb{F}_2^m$ to the vectorspace $V_n = \mathbb{F}_2^n$ are called (m, n)-functions or more generally, vectorial (i.e. multi-output) Boolean functions. These functions include the (single-output) Boolean functions ($n = 1$). For a cryptographic use, such functions need to fulfill many criteria in order to ensure the robustness of the cryptosystems in which they are involved [Car10b]. Among these criteria, a very important notion is the nonlinearity of these functions that must be as high as possible in order to resist to the linear attack. A (m, n)-function is affine if and only if it is a $\mathbb{F}_2$-linear map plus a constant. The nonlinearity $\mathcal{NL}(f)$ of a (m, n)-function f equals the minimum Hamming distance between all the component (Boolean) functions of f, that is $v \cdot f$ where $v \in V_n^* = V_n \setminus \{0\}$, and all affine $(m, 1)$-functions. This notion was introduced and studied by Nyberg [Nyb93] and later by Chabaud and Vaudenay [CV95]. It can be computed through the Walsh transform of these components. For a given $v \in V_n^*$, the Walsh transform of $v \cdot f$ is the Fourier transform of $\chi_{v\cdot f}(x) = (-1)^{(v\cdot f)(x)}$ the ± 1-representation of $v \cdot f$. Let us denote by $\widehat{V_m}$ the set of characters of V_m. For every $\mu \in \widehat{V_m}$, we have

$$\widehat{\chi_{v\cdot f}}(\mu) = \sum_{x \in V_m} (-1)^{(v\cdot f)(x)} \mu(x),$$

where $\mu(x) = (-1)^{x\cdot y}$ for some y in V_m. And the nonlinearity of f is

$$\mathcal{NL}(f) = 2^{m-1} - \frac{1}{2} \max_{\substack{v \in V_n^* \\ \mu \in \widehat{V_m}}} \left| \widehat{\chi_{v\cdot f}}(\mu) \right|. \tag{4.1}$$

Hence a function has high nonlinearity if the Walsh values of all its components have low magnitudes. The covering radius bound being valid for every

m-variable Boolean function, is a fortiori, valid for every (m, n)-function

$$\mathcal{NL}(f) \leq 2^{m-1} - 2^{m/2-1}. \tag{4.2}$$

Nyberg proved that this bound can be achieved with equality only if m is even and $n \leq m/2$ by the so-called bent functions [Nyb91]. Bent functions being not balanced (i.e. they do not take every value of $\mathbb{F}_2^n$ the same number 2^{m-n} of times), they are not suitable for cryptotsystems. This is why it is important to study functions with large but not optimal nonlinearities, among which some balanced functions exist. When $n \geq m-1$, we have the Sidelnikov-Chabaud-Vaudenay bound [CV95]

$$\mathcal{NL}(f) \leq 2^{m-1} - \frac{1}{2}\sqrt{3 \times 2^m - 2 - 2\frac{(2^m-1)(2^{m-1}-1)}{(2^n-1)}},$$

which coincides with the covering radius bound for $n = m-1$ and improves it for $n \geq m$. This bound is achieved when $n = m$, with n odd. The (m, m)-functions which achieve this bound $2^{m-1} - 2^{\frac{m-1}{2}}$ are called almost bent functions. Other bounds have been obtained in [CD07] when n is sufficiently greater than m. Finding better bounds than (4.2) in the other cases remains an open problem.

While several papers [Car02, LS09, OS98, Rod06, Sch13] studied the distribution of the asymptotic nonlinearity of Boolean functions when $n = 1$ as we presented in the first chapter, less was done for $n > 1$. It was proven that almost all m-variable Boolean functions have nonlinearities close to $2^{m-1} - 2^{m/2-1}\sqrt{2m \log 2}$ when m is large enough. In [Dib13], we extended this fact to (m, n)-functions under some restrictions. We prove in theorem 4.3.1 that almost all (m, n)-functions are highly nonlinear. Theorem 4.2.1 provides an upper bound for the nonlinearity of almost all (m, n)-functions, yielding together with theorem 4.3.1, a concentration of the nonlinearity of almost all these functions close to $2^{m-1} - 2^{\frac{m-1}{2}}\sqrt{(m+n)\log 2}$ when $n \leq m$ and m large enough. Namely, we prove the following theorem:

Theorem 4.1.1. *Let m and n be two positive integers such that $n \leq m$, and $0 < \beta < \frac{1}{4}$. The probability that*

$$2^{\frac{m+1}{2}}\sqrt{(m+n)\log 2}\,(1-\beta) \leq \max_{\substack{v \in V_n^* \\ \mu \in \widehat{V}_m}} |\widehat{\chi_{v \cdot f}}(\mu)| \leq 2^{\frac{m+1}{2}}\sqrt{(m+n)\log 2}\,(1+\beta)$$

tends to 1 when m tends to infinity.

When $n = 1$ and using (4.1), we find in particular the point of concentration that is already established for the nonlinearity of random Boolean functions.

We begin by proving the lower bound of $\max_{v \in V_n^*; \mu \in \widehat{V}_m} |\widehat{\chi_{v \cdot f}}(\mu)|$ which is more difficult.

4.2 An asymptotic lower bound on $\max_{v\in V_n^*;\mu\in\widehat{V}_m} |\widehat{\chi_{v\cdot f}}(\mu)|$

To obtain this bound, we use G. Halász method in [Hal73] concerning random trigonometric polynomials. This work inspired F. Rodier [Rod06] to prove a lower asymptotic bound of the above term for Boolean functions ($n = 1$). We use the same scheme of proof, however, it was necessary to take more precise approximations in the case of (m, n)-functions. In this section, we take $n \leq m$, the other case need even more precise calculations.

Let $u(x)$, that will be completely constructed in section 4.2.3, be a function on $\mathbb{R}$ satisfying

$$0 \leq u(x) \leq 1 \quad \forall x \in \mathbb{R}, \quad u(x) = \begin{cases} 0 & \text{for} \quad |x| \leq M \\ 1 & \text{for} \quad |x| \geq M + \Delta, \end{cases}$$

where $M = 2^{\frac{m+1}{2}}\sqrt{(m+n)\log 2}\,(1-\beta)$ with $0 < \beta < 1/4$ and $\Delta = \sqrt{\dfrac{2^m}{\log 2^m}}$. We consider the random variable η on the space of (m, n)-functions

$$\eta(f) = \int_{V_n^*}\int_{\widehat{V}_m} u\left(\widehat{\chi_{v\cdot f}}(\mu)\right) d\mu dv,$$

where $d\mu$ (resp. dv) is a uniform measure over $\widehat{V}_m$ (resp. V_n^*) of total mass 1. The function $u(x)$ is the real Fourier transform of a measure U on $\mathbb{R}$

$$u(x) = \int_{\mathbb{R}} \exp\left(-2\pi i t x\right) dU(t).$$

Hence

$$\eta(f) = \int_{V_n^*}\int_{\widehat{V}_m}\int_{\mathbb{R}} \exp\left(-2\pi i t\, \widehat{\chi_{v\cdot f}}(\mu)\right) dU(t) d\mu dv.$$

For a given f, $\eta(f) = 0$ is equivalent to $\max_{v\in V_n^*;\,\mu\in\widehat{V}_m} |\widehat{\chi_{v\cdot f}}(\mu)| \leq M$. When $n \leq m$, we shall prove by applying Chebyshev's inequality

$$P(\eta = 0) \leq \frac{E(\eta^2) - E(\eta)^2}{E(\eta)^2}$$

that this occurs with probability tending to 0 for large enough m. Before evaluating the first and second moment of η, some estimations are necessary but we choose to give the proof later. The following proposition is given in [Hal73] and [Rod06] but we repeat it for the reader's convenience.

Proposition 4.2.1. *When m tends to infinity, we have the following estimations:*

$$\int_{\mathbb{R}} |dU(t)| = O(m) \tag{4.3}$$

$$\int_{\mathbb{R}} |t|^p |dU(t)| = O\left(\frac{m}{2^m}\right)^{p/2} \quad \textit{for} \quad 1 \le p \le 32 \tag{4.4}$$

$$\left| \int_{\mathbb{R}} \exp\left(-2^{m+1}\pi^2 t^2\right) t^p dU(t) \right| = O\left(2^{-m\frac{p}{2}-(m+n)(1-2\beta)} m^{p/2-1/2}\right), \tag{4.5}$$

for $0 \le p \le 24$.

Proof. See section 4.2.3.

4.2.1 Expected value of η

Lemma 4.2.1.

$$E(\eta) = \int_{\mathbb{R}} \exp\left(-2^{m+1}\pi^2 t^2\right) dU(t) + O\left(2^{-2m-n+2\beta(m+n)} m^{3/2}\right) + O\left(2^{-2m} m^4\right) \tag{4.6}$$

Proof. We have

$$E(\eta) = \int_{V_n^*} \int_{\widehat{V}_m} \int_{\mathbb{R}} E\left(\exp\left(-2\pi i t\, \widehat{\chi_{v\cdot f}}(\mu)\right)\right) dU(t) d\mu dv.$$

The random variables $\chi_{v\cdot f}(x)$ indexed by x are independent, and then also the random variables $\chi_{v\cdot f}(x)\mu(x)$ which take values $+1$ and -1 with probability $1/2$. Thus

$$\begin{aligned} E\left(\exp\left(-2\pi i t\, \widehat{\chi_{v\cdot f}}(\mu)\right)\right) &= E\Big(\prod_{x\in V_m} \exp\left(-2\pi i t \chi_{v\cdot f}(x)\mu(x)\right)\Big) \\ &= \prod_{x\in V_m} E\left(\exp\left(-2\pi i t \chi_{v\cdot f}(x)\mu(x)\right)\right) \\ &= \cos^{2^m}(2\pi t) \qquad (4.7) \\ &= \exp\Big(-2^{m+1}\pi^2 t^2 - \frac{4}{3}\pi^4 2^m t^4\Big) + O(2^m t^6) \end{aligned}$$

for $|t| \le \frac{1}{3\pi}$, by applying on (4.7)

$$\log\cos y = -\frac{y^2}{2} - \frac{y^4}{12} + O\left(y^6\right) \quad \text{for} \quad |y| \le 1$$

and

$$\exp(-a) = \exp(-b) + O(b-a) \quad \text{for} \quad a, b \ge 0. \tag{4.8}$$

For $|t| > \frac{1}{3\pi}$, we use the trivial bound 1 for the integrand. This gives

$$\begin{aligned}
&\int_{\mathbb{R}} E\left(\exp\left(-2\pi i t\, \widehat{\chi_{v\cdot f}}(\mu)\right)\right) dU(t) \\
&= \int_{-\frac{1}{3\pi}}^{\frac{1}{3\pi}} \exp\Big(-2^{m+1}\pi^2 t^2 - \frac{4}{3}\pi^4 2^m t^4\Big) dU(t) + O\Big(2^m \int_{\mathbb{R}} t^6 |dU(t)|\Big) \\
&\qquad + O\Big(\int_{|t|\geq\frac{1}{3\pi}} |dU(t)|\Big).
\end{aligned}$$

We extend the first integral over the real line making the same error as the third term, that can be included in the second one. This yields

$$\int_{\mathbb{R}} \exp\Big(-2^{m+1}\pi^2 t^2 - \frac{4}{3}\pi^4 2^m t^4\Big) dU(t) + O\Big(2^m \int_{\mathbb{R}} t^6 |dU(t)|\Big).$$

By (4.4), the remainder equals $O\left(2^{-2m} m^3\right)$. As for the main term, we use

$$\exp(-a) = 1 - a + O(a^2) \quad \text{for} \quad a > 0,$$

in addition to (4.4) and (4.5) as follows

$$\begin{aligned}
&\int_{\mathbb{R}} \exp\Big(-2^{m+1}\pi^2 t^2 - \frac{4}{3}\pi^4 2^m t^4\Big) dU(t) \\
&= \int_{\mathbb{R}} \exp\left(-2^{m+1}\pi^2 t^2\right) \left(1 - \frac{4}{3}\pi^4 2^m t^4 + O\left(2^{2m} t^8\right)\right) dU(t) \\
&= \int_{\mathbb{R}} \exp\left(-2^{m+1}\pi^2 t^2\right) dU(t) + O\left(2^{-2m-n+2\beta(m+n)} m^{3/2}\right) \\
&\quad + O\Big(2^{2m} \int_{\mathbb{R}} t^8 |dU(t)|\Big) \\
&= \int_{\mathbb{R}} \exp\left(-2^{m+1}\pi^2 t^2\right) dU(t) + O\left(2^{-2m-n+2\beta(m+n)} m^{3/2}\right) + O\left(2^{-2m} m^4\right).
\end{aligned}$$

The proof is complete recalling that the total mass over V_n^* and $\widehat{V}_m$ is 1. □

4.2.2 The second moment

The random variable $\eta^2(f)$ consists of three sums

$$\begin{aligned}
\eta^2(f) = &\int_{\substack{V_n^{*2}\times\widehat{V}_m^2 \\ (v,\mu)=(v',\mu')}} u\left(\widehat{\chi_{v\cdot f}}(\mu)\right) u\left(\widehat{\chi_{v'\cdot f}}(\mu')\right) d\mu dv d\mu' dv' \\
&+ \int_{\substack{V_n^{*2}\times\widehat{V}_m^2 \\ v=v' \\ \mu\neq\mu'}} u\left(\widehat{\chi_{v\cdot f}}(\mu)\right) u\left(\widehat{\chi_{v'\cdot f}}(\mu')\right) d\mu dv d\mu' dv' \\
&+ \int_{\substack{V_n^{*2}\times\widehat{V}_m^2 \\ v\neq v'}} u\left(\widehat{\chi_{v\cdot f}}(\mu)\right) u\left(\widehat{\chi_{v'\cdot f}}(\mu')\right) d\mu dv d\mu' dv',
\end{aligned}$$

which we denote respectively by $\eta_1^2(f)$, $\eta_2^2(f)$ et $\eta_3^2(f)$.

Lemma 4.2.2.

$$E(\eta_1^2) \leq \frac{1}{2^m(2^n-1)}\, E(\eta).$$

Proof. We have

$$\begin{aligned}
\eta_1^2(f) &= \int_{\substack{V_n^{*2}\times \widehat{V}_m^2 \\ (v,\mu)=(v',\mu')}} u\left(\widehat{\chi_{v\cdot f}}(\mu)\right) u\left(\widehat{\chi_{v'\cdot f}}(\mu')\right) d\mu dv d\mu' dv' \\
&= \frac{1}{2^m(2^n-1)} \int_{V_n^*\times \widehat{V}_m} u^2\left(\widehat{\chi_{v\cdot f}}(\mu)\right) d\mu dv \\
&\leq \frac{1}{2^m(2^n-1)} \int_{V_n^*\times \widehat{V}_m} u\left(\widehat{\chi_{v\cdot f}}(\mu)\right) d\mu dv = \frac{1}{2^m(2^n-1)}\, \eta(f)
\end{aligned}$$

noting that $0 \leq u(x) \leq 1$, $\forall x \in \mathbb{R}$.

□

For $E(\eta_2^2)$ (resp. $E(\eta_3^2)$), we use the representation of u as a Fourier transform of U

$$\begin{aligned}
E(\eta_2^2) &= E\Bigg(\int_{\substack{V_n^{*2}\times \widehat{V}_m^2 \\ v=v' \\ \mu\neq\mu'}} u\left(\widehat{\chi_{v\cdot f}}(\mu)\right) u\left(\widehat{\chi_{v\cdot f}}(\mu')\right) d\mu dv d\mu' dv' \Bigg) \\
&= \int_{\substack{V_n^{*2}\times \widehat{V}_m^2 \\ v=v' \\ \mu\neq\mu'}} \int_{\mathbb{R}^2} E\left(\exp\left(-2\pi i t\, \widehat{\chi_{v\cdot f}}(\mu) - 2\pi i t'\, \widehat{\chi_{v\cdot f}}(\mu')\right)\right) dU(t) dU(t') d\mu dv d\mu' dv'.
\end{aligned}$$

We evaluate the integrand in the following lemma.

Lemma 4.2.3. *Let $v \in V_n^*$ and $\mu,\ \mu' \in \widehat{V}_m$ such that $\mu \neq \mu'$. For t and t' of absolute value smaller than $\frac{1}{3\pi}$, we have*

$$\begin{aligned}
E\left(\exp\left(-2\pi i t\, \widehat{\chi_{v\cdot f}}(\mu) - 2\pi i t'\, \widehat{\chi_{v\cdot f}}(\mu')\right)\right) = \exp\Bigg(&- 2^m \sum_{i=1}^{4} \sum_{j=0}^{i} c_{i,j} t^{2i-2j} t'^{2j} \Bigg) \\
&+ 2^m O(|t|+|t'|)^{10},
\end{aligned}$$

where $c_{i,j}$ are positive reals.

Proof. The random variables $\chi_{v\cdot f}(x)$ indexed by x are independent, and then also the random variables $\chi_{v\cdot f}(x)\left(t\mu(x) + t'\mu'(x)\right)$ which take values $(t\mu(x) + t'\mu'(x))$ and $-\left(t\mu(x) + t'\mu'(x)\right)$ with probability $1/2$. Thus

$$\begin{aligned}
&E\left(\exp\left(-2\pi i t\, \widehat{\chi_{v\cdot f}}(\mu) - 2\pi i t'\, \widehat{\chi_{v\cdot f}}(\mu')\right)\right) \\
&= E\Big(\prod_{x\in V_m} \exp\left(-2\pi i \chi_{v\cdot f}(x)\left(t\mu(x) + t'\mu'(x)\right)\right) \Big) \\
&= \prod_{x\in V_m} E\left(\exp\left(-2\pi i \chi_{v\cdot f}(x)\left(t\mu(x) + t'\mu'(x)\right)\right)\right) \\
&= \prod_{x\in V_m} \cos(2\pi(t\mu(x) + t'\mu'(x))).
\end{aligned}$$

Since $\mu \neq \mu'$, they agree 2^{m-1} times

$$\begin{aligned}
&\prod_{x\in V_m} \cos(2\pi(t\mu(x)+t'\mu'(x)))\\
&= \cos^{2^{m-1}}(2\pi(t+t'))\cos^{2^{m-1}}(2\pi(t-t'))\\
&= \exp\Big(-2^m\Big(\sum_{i=1}^{4} c_i(t+t')^{2i} + O(t+t')^{10} + \sum_{i=1}^{4} c_i(t-t')^{2i} + O(t-t')^{10}\Big)\Big),
\end{aligned}$$

for $|t| \le \frac{1}{3\pi}$, $|t'| \le \frac{1}{3\pi}$, by applying

$$\log\cos y = -\frac{y^2}{2} - \frac{y^4}{12} - \frac{y^6}{45} - \frac{17y^8}{2520} + O\left(y^{10}\right) \quad \text{for} \quad |y| \le 1.$$

Simplifying and using (4.8) give the result. □

Lemma 4.2.4.

$$\begin{aligned}
E(\eta_2^2) = \frac{1}{2^n-1}\Big(\int_{\mathbb{R}} \exp\left(-2^{m+1}\pi^2 t^2\right) dU(t)\Big)^2 + O\left(2^{-3m-3n+4\beta(m+n)}m\right)\\
+ O\left(2^{-4m-n}m^9\right).
\end{aligned}$$

Proof. Using the previous lemma together with the trivial bound 1 for the integrand outside the square $|t| \ge \frac{1}{3\pi}$, $|t'| \ge \frac{1}{3\pi}$ give

$$\begin{aligned}
&\int_{\mathbb{R}^2} E\left(\exp\left(-2\pi i t\,\widehat{\chi_{v\cdot f}}(\mu) - 2\pi i t'\,\widehat{\chi_{v\cdot f}}(\mu')\right)\right) dU(t)dU(t')\\
&= \int_{-\frac{1}{3\pi}}^{\frac{1}{3\pi}}\int_{-\frac{1}{3\pi}}^{\frac{1}{3\pi}} \exp\Big(-2^m\sum_{i=1}^{4}\sum_{j=0}^{i} c_{i,j}t^{2i-2j}t'^{2j}\Big)dU(t)dU(t')\\
&\quad + O\Big(2^m\int_{\mathbb{R}^2}(|t|+|t'|)^{10}|dU(t)||dU(t')|\Big) \qquad (4.9)\\
&\quad + O\Big(\int_{|t|\ge\frac{1}{3\pi}}\int_{\mathbb{R}}|dU(t)||dU(t')|\Big).
\end{aligned}$$

We extend integration in the first term over $\mathbb{R}^2$ making the same error as the third term, that is smaller than the second one.
Noting that $c_{1,0} = c_{1,1} = 2\pi^2$, and applying

$$\exp(-a) = 1 - a + \frac{a^2}{2} - \frac{a^3}{6} + O(a^4) \quad \text{for} \quad a > 0,$$

the first term then becomes

$$\int_{\mathbb{R}^2} \exp(-2^{m+1}\pi^2 t^2)\exp(-2^{m+1}\pi^2 t'^2)\Big(1 - 2^m \sum_{i=2}^{4}\sum_{j=0}^{i} c_{i,j} t^{2i-2j} t'^{2j} + 2^{2m}\sum_{i=4}^{8}\sum_{j=0}^{i} l_{i,j} t^{2i-2j} t'^{2j} - 2^{3m}\sum_{i=6}^{12}\sum_{j=0}^{i} p_{i,j} t^{2i-2j} t'^{2j} + 2^{4m} O\Big(\sum_{i=8}^{16}\sum_{j=0}^{i} r_{i,j} t^{2i-2j} t'^{2j}\Big)\Big) dU(t) dU(t')$$

and by (4.5), we get

$$\begin{aligned}\Big(\int_{\mathbb{R}} \exp(-2^{m+1}\pi^2 t^2) dU(t)\Big)^2 &- 2^m \sum_{i=2}^{4}\sum_{j=0}^{i} O\left(2^{-mi-2(m+n)(1-2\beta)} m^{i-1}\right) \\ &+ 2^{2m}\sum_{i=4}^{8}\sum_{j=0}^{i} O\left(2^{-mi-2(m+n)(1-2\beta)} m^{i-1}\right) \\ &- 2^{3m}\sum_{i=6}^{12}\sum_{j=0}^{i} O\left(2^{-mi-2(m+n)(1-2\beta)} m^{i-1}\right) \\ &+ 2^{4m} O\left(\sum_{i=8}^{16}\sum_{j=0}^{i} r_{i,j} \int_{\mathbb{R}} t^{2i-2j}|dU(t)| \int_{\mathbb{R}} t'^{2j}|dU(t')|\right). \end{aligned} \tag{4.10}$$

Terms in (4.10) with $i = j$ or $j = 0$ are equal to $2^{4m} O(m) O\left(\frac{m}{2^m}\right)^8$ by (4.3) and (4.4). The other terms are equal to $2^{4m} O\left(\frac{m}{2^m}\right)^8$ by (4.4). This gives

$$\left(\int_{\mathbb{R}} \exp\left(-2^{m+1}\pi^2 t^2\right) dU(t)\right)^2 + O\left(2^{-3m-2n+4\beta(m+n)} m\right) + O\left(2^{-4m} m^9\right).$$

As for (4.9), it can be estimated just like (4.10) using (4.3) and (4.4), yielding $O\left(2^{-4m} m^6\right)$. We end the calculations by integrating over the other variables. □

Lemma 4.2.5.

$$E(\eta_3^2) = \Big(1 - \frac{1}{2^n - 1}\Big) E^2(\eta).$$

Proof. $E(\eta_3^2)$ equals

$$\int_{\substack{V_n^{*2} \times \widehat{V}_m^2 \\ v \neq v'}} \int_{\mathbb{R}^2} E\left(\exp\left(-2\pi i t\, \widehat{\chi_{v \cdot f}}(\mu) - 2\pi i t'\, \widehat{\chi_{v' \cdot f}}(\mu')\right)\right) dU(t) dU(t') d\mu dv d\mu' dv'.$$

Since $v \neq v'$, the random variables $\widehat{\chi_{v\cdot f}}(\mu)$ and $\widehat{\chi_{v'\cdot f}}(\mu')$ are independent. Thus

$$\begin{aligned}
&E\left(\exp\left(-2\pi it\,\widehat{\chi_{v\cdot f}}(\mu) - 2\pi it'\,\widehat{\chi_{v'\cdot f}}(\mu')\right)\right)\\
&= E\left(\exp\left(-2\pi it\,\widehat{\chi_{v\cdot f}}(\mu)\right)\right) E\left(\exp\left(-2\pi it'\,\widehat{\chi_{v'\cdot f}}(\mu')\right)\right)\\
&= \cos^{2^m}(2\pi t)\cos^{2^m}(2\pi t'),
\end{aligned}$$

as shown previously in (4.7). And,

$$\begin{aligned}
E(\eta_3^2) &= \int_{\substack{V_n^{*2}\times\widehat{V}_m^2\\ v\neq v'}} d\mu dv d\mu' dv' \left(\int_{\mathbb{R}} \cos^{2^m}(2\pi t) dU(t)\right)^2\\
&= \Big(1 - \frac{1}{2^n - 1}\Big) E^2(\eta).
\end{aligned}$$

□

Lemma 4.2.6.

$$\frac{1}{E(\eta)} = O\left(2^{(m+n)(1-\beta)^2}\sqrt{m}\right). \tag{4.11}$$

Proof. We have

$$E(\eta) = \int_{\mathbb{R}} \exp\left(-2^{m+1}\pi^2 t^2\right) dU(t) + O\left(2^{-2m-n+2\beta(m+n)} m^{3/2}\right) + O\left(2^{-2m} m^4\right).$$

The Fourier transform of $\exp\left(-2^{m+1}\pi^2 t^2\right)$ is $\dfrac{1}{\sqrt{2^{m+1}\pi}} \exp\left(-\dfrac{x^2}{2^{m+1}}\right)$. Hence, by Plancherel's theorem, we have

$$\begin{aligned}
\int_{\mathbb{R}} \exp\left(-2^{m+1}\pi^2 t^2\right) dU(t) &= \frac{1}{\sqrt{2^{m+1}\pi}} \int_{\mathbb{R}} \exp\left(-\frac{x^2}{2^{m+1}}\right) u(x) dx\\
&\geq \frac{1}{\sqrt{2^{m+1}\pi}} \int_{|x|\geq M+\Delta} \exp\left(-\frac{x^2}{2^{m+1}}\right) dx.
\end{aligned}$$

To evaluate the integral, we use (2.5). Thus, we obtain

$$\int_{\mathbb{R}} \exp\left(-2^{m+1}\pi^2 t^2\right) dU(t) \geq C_1\, 2^{-(m+n)(1-\beta)^2} m^{-1/2}.$$

Adding the fact that

$$O\left(2^{-2m-n+2\beta(m+n)} m^{3/2}\right) + O\left(2^{-2m} m^4\right) = o\left(2^{-(m+n)(1-\beta)^2} m^{-1/2}\right),$$

proves the result. □

Theorem 4.2.1. *Let $0 < \beta < \frac{1}{4}$ and γ any positive real. When m tends to infinity and $n \leq m$, we have*

$$P\Big(\max_{\substack{v\in V_n^*\\ \mu\in\widehat{V}_m}} |\widehat{\chi_{v\cdot f}}(\mu)| \leq 2^{\frac{m+1}{2}} \sqrt{(m+n)\log 2}\,(1-\beta)\Big) = P(\eta = 0) = O\left(m^{-\gamma}\right).$$

Proof. We have

$$\begin{aligned} &E(\eta^2) - E^2(\eta) \\ &\quad \leq \frac{E(\eta)}{2^m(2^n-1)} + \frac{1}{2^n-1}\left(\left(\int_{\mathbb{R}} \exp\left(-2^{m+1}\pi^2 t^2\right) dU(t) \right)^2 - E^2(\eta) \right) \\ &\qquad + O\left(2^{-3m-3n+4\beta(m+n)} m\right) + O\left(2^{-4m-n} m^9\right), \end{aligned}$$

and by (4.6), we get

$$\begin{aligned} &E(\eta^2) - E^2(\eta) \\ &\quad \leq \frac{E(\eta)}{2^m(2^n-1)} + \frac{1}{2^n-1}\left(\left(\int_{\mathbb{R}} \exp\left(-2^{m+1}\pi^2 t^2\right) dU(t) \right) \right. \\ &\qquad \left. O\left(2^{-2m-n+2\beta(m+n)} m^{3/2}\right) + O\left(2^{-4m} m^8\right) + E(\eta) O\left(2^{-2m} m^4\right) \right) \\ &\qquad + O\left(2^{-3m-3n+4\beta(m+n)} m\right) + O\left(2^{-4m-n} m^9\right). \end{aligned}$$

When divided by $E^2(\eta)$, we can check using (4.11) and (4.5) that every term is smaller than $O\left(m^{-\gamma}\right)$. □

4.2.3 Some estimations

Before giving the proof, we first complete the construction of u. Let us fix a 34 times continuously differentiable function α on $[0,1]$, which takes 0 at 0, 1 at 1, takes values between 0 and 1, and with vanishing derivatives up to the 18th order at 0 and 1. By choosing $u(x)$ to be equal to $\alpha\left(\frac{|x|-M}{\Delta}\right)$ for $M \leq |x| \leq M+\Delta$, $u(x)$ is then a 34 times differentiable function on $\mathbb{R}$ with $|u^r(x)| \leq \dfrac{\text{constant}}{\Delta^r}$, for $r = 0, 1, ..., 34$.

Proof. The measure U, having u as its Fourier transform, can be written as the sum of the Dirac measure at the origin and

$$g(t) = \int_{\mathbb{R}} \exp(-2\pi i t x)(u(x)-1)dx = \int_{-M-\Delta}^{M+\Delta} \exp(-2\pi i t x)(u(x)-1)dx.$$

We have

$$|g(t)| \leq 2(M+\Delta) = O(M). \tag{4.12}$$

And integration by parts gives

$$|t^r g(t)| \leq \int_{-M-\Delta}^{M+\Delta} |u^{(r)}(x)|dx = O\left(\frac{1}{\Delta^{r-1}}\right) \quad \text{for} \quad r = 1, ..., 34. \tag{4.13}$$

To prove (4.3), we use (4.12) for $|t| \leq \dfrac{1}{\Delta}$ and (4.13) with $r = 2$ for $|t| \geq \dfrac{1}{\Delta}$

$$\int_{\mathbb{R}} |dU(t)| = 1 + \int_{\mathbb{R}} |g(t)|dt = O\left(\frac{M}{\Delta}\right) = O(m).$$

To prove (4.4), we use (4.13) with $r=p$ for $|t|\le\frac{1}{\Delta}$ and with $r=p+2$ for $|t|\ge\frac{1}{\Delta}$:

$$\int_{\mathbb{R}}|t^p||dU(t)|=\int_{\mathbb{R}}|t^p||g(t)|dt=O\left(\frac{1}{\Delta^p}\right)=O\left(\frac{m}{2^m}\right)^{p/2}\quad\text{for}\quad p=1,...,32.$$

To prove (4.5), we use the Plancherel's theorem. The Fourier transform of t^pU is $\frac{i^p}{(2\pi)^p}u^{(p)}(x)$ and that of $\exp\left(-2^{m+1}\pi^2t^2\right)$ is $\frac{1}{\sqrt{2^{m+1}\pi}}\exp\left(-\frac{x^2}{2^{m+1}}\right)$, thus

$$\begin{aligned}\left|\int_{\mathbb{R}}\exp\left(-2^{m+1}\pi^2t^2\right)t^pdU(t)\right| &= \frac{1}{\sqrt{2^{m+1}\pi}\,(2\pi)^p}\left|\int_{\mathbb{R}}\exp\left(-\frac{x^2}{2^{m+1}}\right)u^{(p)}(x)dx\right|\\ &= O\left(\frac{1}{\Delta^p\sqrt{2^m}}\right)\int_{|x|\ge M}\exp\left(-\frac{x^2}{2^{m+1}}\right)dx.\\ &= O\left(2^{\left(-m\frac{p}{2}-(m+n)(1-\beta)^2\right)}m^{p/2-1/2}\right),\end{aligned}\tag{4.14}$$

using (2.5). □

4.3 An asymptotic upper bound on $\max_{v\in V_n^*;\mu\in\widehat{V}_m}|\widehat{\chi_{v\cdot f}}(\mu)|$

The result of this section is a generalization of F. Rodier's result [Rod04] on Boolean functions inspired by the work of R. Salem and A. Zygmund [SZ54] on trigonometric series.

Lemma 4.3.1. *Let λ be a real number, $v\in V_n^*$ and $\mu\in\widehat{V}_m$. Then, for f running in the space of (m,n)-functions*

$$E\left(\exp\left(\lambda\,\widehat{\chi_{v\cdot f}}(\mu)\right)\right)\le\exp\left(2^{m-1}\lambda^2\right).$$

Proof. The random variables $\chi_{v\cdot f}(x)$ indexed by x are independent, and then also $\chi_{v\cdot f}(x)\mu(x)$ which take values $+1$ and -1 with probability $1/2$. Thus

$$\begin{aligned}E\left(\exp\left(\lambda\,\widehat{\chi_{v\cdot f}}(\mu)\right)\right) &= E\left(\prod_{x\in V_m}\exp\left(\lambda\,\chi_{v\cdot f}(x)\mu(x)\right)\right)\\ &= \prod_{x\in V_m}E\left(\exp\left(\lambda\,\chi_{v\cdot f}(x)\mu(x)\right)\right)\\ &= \prod_{x\in V_m}\cosh\lambda.\end{aligned}$$

And

$$\cosh\lambda\le\exp\frac{\lambda^2}{2}.$$

□

Theorem 4.3.1. *Let m and n be any positive integers and β any positive real. Then*

$$P\Big(\max_{\substack{v\in V_n^*\\ \mu\in\widehat{V}_m}} |\widehat{\chi_{v\cdot f}}(\mu)| \geq 2^{\frac{m+1}{2}}\sqrt{(m+n)\log 2}\,(1+\beta)\Big) \leq 2^{-(m+n)(2\beta+\beta^2)+1}.$$

Proof. There exists (v_0,μ_0) in $V_n^*\times\widehat{V}_m$ such that $\max\limits_{\substack{v\in V_n^*\\ \mu\in\widehat{V}_m}} |\widehat{\chi_{v\cdot f}}(\mu)| = |\widehat{\chi_{v_0\cdot f}}(\mu_0)|$.

Let λ be a positive real, we have

$$\begin{aligned}
&\exp\Big(\lambda \max_{\substack{v\in V_n^*\\ \mu\in\widehat{V}_m}} |\widehat{\chi_{v\cdot f}}(\mu)|\Big)\\
&\quad\leq \exp\big(\lambda\,\widehat{\chi_{v_0\cdot f}}(\mu_0)\big) + \exp\big(-\lambda\,\widehat{\chi_{v_0\cdot f}}(\mu_0)\big)\\
&\quad\leq 2^{m+n}\int_{V_n^*}\int_{\widehat{V}_m} \left(\exp\left(\lambda\,\widehat{\chi_{v\cdot f}}(\mu)\right) + \exp\left(-\lambda\,\widehat{\chi_{v\cdot f}}(\mu)\right)\right) d\mu dv.
\end{aligned}$$

When f ranges over the space of (m,n)-functions, using lemma 4.3.1 and recalling that the total mass over V_n^* and $\widehat{V}_m$ is 1, we have

$$\begin{aligned}
&E\Big(\exp\Big(\lambda \max_{\substack{v\in V_n^*\\ \mu\in\widehat{V}_m}} |\widehat{\chi_{v\cdot f}}(\mu)|\Big)\Big)\\
&\quad\leq 2^{m+n+1}\exp\left(2^{m-1}\lambda^2\right)\\
&\quad= 2^{-(m+n)(2\beta+\beta^2)+1}\exp\left(2^{m-1}\lambda^2 + (m+n)(1+\beta)^2\log 2\right).
\end{aligned}$$

Thus,

$$E\Big(\exp\Big(\lambda \max_{\substack{v\in V_n^*\\ \mu\in\widehat{V}_m}} |\widehat{\chi_{v\cdot f}}(\mu)| - 2^{m-1}\lambda^2 - (m+n)(1+\beta)^2\log 2\Big)\Big) \leq 2^{-(m+n)(2\beta+\beta^2)+1}.$$

Consequently,

$$P\Big(\exp\Big(\lambda \max_{\substack{v\in V_n^*\\ \mu\in\widehat{V}_m}} |\widehat{\chi_{v\cdot f}}(\mu)| - 2^{m-1}\lambda^2 - (m+n)(1+\beta)^2\log 2\Big) \geq 1\Big)$$

is less than

$$2^{-(m+n)(2\beta+\beta^2)+1}.$$

And finally,

$$P\Big(\max_{\substack{v\in V_n^*\\ \mu\in\widehat{V}_m}} |\widehat{\chi_{v\cdot f}}(\mu)| \geq 2^{m-1}\lambda + \frac{(m+n)(1+\beta)^2\log 2}{\lambda}\Big) \leq 2^{-(m+n)(2\beta+\beta^2)+1}.$$

The best bound is obtained when $\lambda = 2^{\frac{1-m}{2}}\sqrt{(m+n)\log 2}\,(1+\beta)$, which gives the required result. When $(m+n)$ tends to infinity, we obtain then a lower bound of the nonlinearity of almost all (m,n)-functions. □

Bibliography

[AS92] Noga Alon and Joel H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Inc., New York, 1992. With an appendix by Paul Erdős, A Wiley-Interscience Publication.

[BW72] Elwyn R. Berlekamp and Lloyd R. Welch. Weight distributions of the cosets of the (32, 6) reed-muller code. *IEEE Transactions on Information Theory*, 18(1):203–207, 1972.

[Car02] Claude Carlet. On cryptographic complexity of Boolean functions. In *Finite fields with applications to coding theory, cryptography and related areas (Oaxaca, 2001)*, pages 53–69. Springer, Berlin, 2002.

[Car06] Claude Carlet. The complexity of boolean functions from cryptographic viewpoint. In Matthias Krause, Pavel Pudlák, Rüdiger Reischuk, and Dieter van Melkebeek, editors, *Complexity of Boolean Functions*, volume 06111 of *Dagstuhl Seminar Proceedings*. Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI), Schloss Dagstuhl, Germany, 2006.

[Car10a] Claude Carlet. Boolean functions for cryptography and error correcting codes. In *Boolean models and methods in mathematics, computer science, and engineering*, pages 257–397. Cambridge: Cambridge University Press, 2010.

[Car10b] Claude Carlet. Vectorial Boolean functions for cryptography. In *Boolean models and methods in mathematics, computer science, and engineering*, pages 398–469. Cambridge: Cambridge University Press, 2010.

[CD07] Claude Carlet and Cunsheng Ding. Nonlinearities of S-boxes. *Finite Fields Appl.*, 13(1):121–135, 2007.

[CM07] Claude Carlet and Sihem Mesnager. Improving the upper bounds on the covering radii of binary Reed-Muller codes. *IEEE Trans. Inform. Theory*, 53(1):162–173, 2007.

[CT00] Anne Canteaut and Michaël Trabbia. Improved fast correlation attacks using parity-check equations of weight 4 and 5. In *Advances in Cryptology - EUROCRYPT 2000*, volume 1805 of *Lecture Notes in Computer Science*, pages 573–588, 2000.

[CV95] Florent Chabaud and Serge Vaudenay. Links between differential and linear cryptanalysis. In *Advances in cryptology—EUROCRYPT '94 (Perugia)*, volume 950 of *Lecture Notes in Comput. Sci.*, pages 356–365. Springer, Berlin, 1995.

[Dib10] Stéphanie Dib. Distribution of boolean functions according to the second-order nonlinearity. In M. Anwar Hasan and Tor Helleseth, editors, *WAIFI*, volume 6087 of *Lecture Notes in Computer Science*, pages 86–96. Springer, 2010.

[Dib13] Stéphanie Dib. Asymptotic nonlinearity of vectorial boolean functions. *Cryptography and communications*, 2013.

[Dur91] Richard Durrett. *Probability: theory and examples*. Duxbury Press, Belmont, CA, first edition, 1991.

[Fel68] William Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons Inc., New York, 1968.

[Fon98] Caroline Fontaine. *Contribution à la recherche de fonctions booléennes hautement non linéaires, et au marquage d'images en vue de la protection des droits d'auteur*. PhD thesis, Université Paris VI, 1998.

[GKP89] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete mathematics*. Addison-Wesley Publishing Company Advanced Book Program, Reading, MA, 1989. A foundation for computer science.

[Hal73] G. Halász. On a result of Salem and Zygmund concerning random polynomials. *Studia Sci. Math. Hungar.*, 8:369–377, 1973.

[HKL05] Tor Helleseth, Torleiv Kløve, and Vladimir I. Levenshtein. Error-correction capability of binary linear codes. *IEEE Trans. Inform. Theory*, 51(4):1408–1423, 2005.

[KDY10] Selçuk Kavut and Melek Diker Yücel. 9-variable Boolean functions with nonlinearity 242 in the generalized rotation symmetric class. *Inform. and Comput.*, 208(4):341–350, 2010.

[KR96] Lars R. Knudsen and M. J. B. Robshaw. Non-linear approximations in linear cryptanalysis. In *Proceedings of the 15th annual international conference on Theory and application of cryptographic techniques*, EUROCRYPT'96, pages 224–236, Berlin, Heidelberg, 1996. Springer-Verlag.

[LS09] Simon Litsyn and Alexander Shpunt. On the distribution of Boolean function nonlinearity. *SIAM J. Discrete Math.*, 23(1):79–95, 2008/09.

[Mat94] Mitsuru Matsui. Linear cryptanalysis method for des cipher. In Tor Helleseth, editor, *Advances in Cryptology - EUROCRYPT'93*, volume 765 of *Lecture Notes in Computer Science*, pages 386–397. Springer Berlin Heidelberg, 1994.

[MS77] F. J. MacWilliams and N. J. A. Sloane. *The theory of error-correcting codes. I.* North-Holland Publishing Co., Amsterdam, 1977. North-Holland Mathematical Library, Vol. 16.

[Myk80] Johannes J. Mykkeltveit. The covering radius of the (128, 8) Reed-Muller code is 56. *IEEE Trans. Inform. Theory*, 26(3):359–362, 1980.

[Nyb91] Kaisa Nyberg. Perfect nonlinear S-boxes. In *Advances in cryptology—EUROCRYPT '91 (Brighton, 1991)*, volume 547 of *Lecture Notes in Comput. Sci.*, pages 378–386. Springer, Berlin, 1991.

[Nyb93] Kaisa Nyberg. On the construction of highly nonlinear permutations. In *Advances in cryptology—EUROCRYPT '92 (Balatonfüred, 1992)*, volume 658 of *Lecture Notes in Comput. Sci.*, pages 92–98. Springer, Berlin, 1993.

[OS98] Daniel Olejár and Martin Stanek. On cryptographic properties of random Boolean functions. *J. UCS*, 4(8):705–717 (electronic), 1998.

[PW90] Nick J. Patterson and Douglas H. Wiedemann. Correction to: "The covering radius of the $(2^{15}, 16)$ Reed-Muller code is at least 16 276" [IEEE Trans. Inform. Theory **29** (1983), no. 3, 354–356; MR0712397 (85h:94031)]. *IEEE Trans. Inform. Theory*, 36(2):443, 1990.

[Rod03] François Rodier. On the nonlinearity of boolean functions. In *Proceedings of the 2003 International Workshop on Coding and Cryptography*, pages 397 – 405, 2003.

[Rod04] François Rodier. Sur la non-linéarité des fonctions booléennes. *Acta Arith.*, 115(1):1–22, 2004.

[Rod06] François Rodier. Asymptotic nonlinearity of Boolean functions. *Des. Codes Cryptogr.*, 40(1):59–70, 2006.

[Sch13] Kai-Uwe Schmidt. Nonlinearity measures of random boolean functions. arXiv/1308.3112, 2013.

[SK98] Takeshi Shimoyama and Toshinobu Kaneko. Quadratic relation of s-box and its application to the linear attack of full round des. In Hugo Krawczyk, editor, *CRYPTO*, volume 1462 of *Lecture Notes in Computer Science*, pages 200–211. Springer, 1998.

[SZ54] R. Salem and A. Zygmund. Some properties of trigonometric series whose terms have random signs. *Acta Math.*, 91:245–301, 1954.

Résumé : Parmi les différents critères qu'une fonction booléenne doit satisfaire en cryptographie, on s'intéresse à la non-linéarité. Pour une fonction booléenne donnée, cette notion mesure la distance de Hamming qui la sépare des fonctions de degré au plus 1. C'est un critère naturel pour évaluer la complexité d'une fonction cryptographique, celle-ci ne devant pas admettre une approximation qui soit simple, comme par une fonction de degré 1, ou plus généralement une fonction de bas degré. Ainsi, il est important de considérer plus généralement, la non-linéarité d'ordre supérieur, qui pour un ordre donné r, mesure la distance d'une fonction donnée à l'ensemble des fonctions de degré au plus r. Cette notion est également importante pour les fonctions vectorielles, i.e., celles à plusieurs sorties. Quand le nombre de variables est grand, presque toutes les fonctions ont une non-linéarité (d'ordre 1) voisine d'une certaine valeur, assez élevée. Dans [Dib10], on étend ce résultat à l'ordre 2. Cette méthode qui consiste à observer comment les boules de Hamming recouvrent l'hypercube des fonctions booléennes, nous conduit naturellement vers une borne de décodage théorique des codes de Reed-Muller d'ordre 1, coïncidant au même endroit où se concentre la non-linéarité de presque toutes les fonctions ; une approche nouvelle pour un résultat pas entièrement nouveau. On étudie aussi la non-linéarité des fonctions vectorielles. On montre [Dib13] avec une approche différente, que le comportement asymptotique est le même que celui des fonctions booléennes: une concentration de la non-linéarité autour d'une valeur assez élevée.

Mots-clés : Fonctions booléennes, non-linéarité, codes de Reed-Muller, décodage par maximum de vraisemblance.

Abstract : Among the different criteria that a Boolean function must satisfy in symmetric cryptography, we focus on the nonlinearity of these. This notion measures the Hamming distance between a given function and the set of functions with degree at most 1. It is a natural criterion to evaluate the complexity of a cryptographic function that must not have a simple approximation as by a function of degree 1, or more generally, a function of low degree. Hence, it is important to consider the higher order nonlinearity, which for a given order r, measures the distance between a given function and the set of all functions of degree at most r. This notion is equally important for multi-output Boolean functions. When the number of variables is large enough, almost all Boolean functions have nonlinearities lying in a small neighbourhood of a certain high value. In [Dib10], we prove that this fact holds when considering the second-order nonlinearity. Our method which consists in observing how the Hamming balls pack the hypercube of Boolean functions led quite naturally to a theoretical decoding bound for the first-order Reed-Muller code, coinciding with the concentration point of the nonlinearity of almost all functions. This was a new approach for a result which is not entirely new. We also studied the nonlinearity of multi-output

functions. We proved [Dib13] with a different approach, that the asymptotic behaviour of multi-output functions is the same as the single-output ones: a concentration of the nonlinearity around a certain large value.

Keywords : Boolean functions, nonlinearity, Reed-Muller codes, maximum-likelihood decoding.