NONLINEARITY MEASURES OF RANDOM BOOLEAN FUNCTIONS

KAI-UWE SCHMIDT

Abstract. The r-th order nonlinearity of a Boolean function is the minimum number of elements that have to be changed in its truth table to arrive at a Boolean function of degree at most r. It is shown that the (suitably normalised) r-th order nonlinearity of a random Boolean function converges strongly for all $r \geq 1$. This extends results by Rodier for $r = 1$ and by Dib for $r = 2$. The methods in the present paper are mostly of elementary combinatorial nature and also lead to simpler proofs in the cases that $r = 1$ or 2.

1. Introduction and Results

Let $\mathbb{F}_2$ be a field with two elements. A *Boolean function* f is a mapping from $\mathbb{F}_2^n$ to $\mathbb{F}_2$ and its *truth table* is the list of values $f(x)$ as x ranges over $\mathbb{F}_2^n$ in some fixed order. Let $\mathfrak{B}_n$ be the space of Boolean functions on $\mathbb{F}_2^n$. Every $f \in \mathfrak{B}_n$ can be written uniquely in the form

$$f(x_1, \dots, x_n) = \sum_{k_1,\dots,k_n \in \{0,1\}} a_{k_1,\dots,k_n}\, x_1^{k_1} \cdots x_n^{k_n},$$

where $a_{k_1,\dots,k_n} \in \mathbb{F}_2$. The *degree* of f is defined to be the algebraic degree of this polynomial.

The *r-th order nonlinearity* $N_r(f)$ of a Boolean function f is the minimum number of elements that have to be changed in its truth table to arrive at the truth table of a Boolean function of degree at most r. We state this definition more formally as follows. Let $\mathrm{RM}(r, n)$ be the set of Boolean functions in $\mathfrak{B}_n$ of degree at most r (which is known as the *Reed-Muller code* of length 2^n and order r; see [9, Chapters 13–15], for example) and define the *Hamming distance* between $f, g \in \mathfrak{B}_n$ to be

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

Then the r-th order nonlinearity of f is

$$N_r(f) = \min_{g \in \mathrm{RM}(r,n)} d(f, g).$$

The nonlinearity of Boolean functions is of significant relevance in cryptography since it measures the resistance of a Boolean function against low-degree

Date: 14 August 2013 (revised 07 October 2015).

2010 *Mathematics Subject Classification.* 60B10, 06E30, 11T71.

Key words and phrases. Boolean function, convergence, nonlinearity, Reed-Muller code.

approximation attacks (see [7], for example, and [2] for more background on the role of Boolean functions in cryptography and error-correcting codes).

Our interest is the distribution of the nonlinearity of Boolean functions. To this end, let Ω be the set of infinite sequences of elements from $\mathbb{F}_2$ and let $\mathfrak{B}$ be the space of functions from Ω to $\mathbb{F}_2$. For $f \in \mathfrak{B}$, we denote the function given by $f(x_1, \dots, x_n, 0, 0, \dots)$ by f_n, which is in $\mathfrak{B}_n$. We endow $\mathfrak{B}$ with a probability measure defined by

$$\Pr\big[f \in \mathfrak{B} : f_n = g\big] = 2^{-2^n} \quad \text{for all } g \in \mathfrak{B}_n \text{ and all } n \in \mathbb{N}. \tag{1}$$

A basic probabilistic method can be used to show that, if f is drawn from $\mathfrak{B}$, equipped with the probability measure defined by (1), then

$$\limsup_{n\to\infty} \frac{2^{n-1} - N_r(f_n)}{\sqrt{2^{n-1}\binom{n}{r}\log 2}} \le 1 \quad \text{almost surely.} \tag{2}$$

This was essentially proved by Carlet [1, Theorem 1]. The aim of this note is to prove strong convergence of the normalised r-th order nonlinearity, which shows that the bound (2) is best possible.

Theorem 1. *Let f be drawn at random from $\mathfrak{B}$, equipped with the probability measure defined by* (1)*. Then for all fixed $r \ge 1$, as $n \to \infty$,*

$$\frac{2^{n-1} - N_r(f_n)}{\sqrt{2^{n-1}\binom{n}{r}\log 2}} \to 1 \quad \textit{almost surely} \tag{3}$$

and

$$\frac{2^{n-1} - \mathrm{E}[N_r(f_n)]}{\sqrt{2^{n-1}\binom{n}{r}\log 2}} \to 1. \tag{4}$$

Using rather subtle Fourier analytic methods due to Halász [5], Rodier [13] proved (3) for $r = 1$ (see also [11] and [12] for prior results). More precise estimates on the rate of convergence in this case were given by Litsyn and Shpunt [8], using different methods. Dib [3] used a more combinatorial approach to essentially prove (3) for $r = 2$. The methods in this paper are mostly of elementary combinatorial nature and also lead to simpler proofs of (3) in the cases that $r = 1$ or 2.

A brief outline of the proof of Theorem 1 is given next. With the notation as in Theorem 1, write $Y_{n,g} = 2^n - 2d(f_n, g)$ for $g \in \mathfrak{B}_n$ and

$$Y_n = \max_{g \in \mathrm{RM}(r,n)} Y_{n,g},$$

so that $Y_n = 2^n - 2N_r(f_n)$. We make repeated use of the inequality

$$\Pr\Big[\big|Y_n - \mathrm{E}[Y_n]\big| \ge \theta\Big] \le 2\exp\left(-\frac{\theta^2}{2^{n+1}}\right) \quad \text{for } \theta \ge 0, \tag{5}$$

which follows from standard results on concentration of probability measures (see McDiarmid [10, Lemma 1.2], for example). This shows that Y_n

is concentrated around its expectation. Therefore, the main difficulty is to prove (4). We do this by proving upper and lower bounds for $\mathrm{E}[Y_n]$. The upper bound is easy, but for the lower bound we need to work harder. The strategy is as follows. In Section 2, we use a theorem on the weight distribution of Reed-Muller codes due to Kaufman, Lovett, and Porat [6] to show that most pairs of functions in $\mathrm{RM}(r,n)$ have Hamming distance close to 2^{n-1}. Combining this with some large deviation estimates in Section 3 then shows that the events

$$Y_{n,g} \geq \sqrt{2^{n+1}\binom{n}{r}\log 2}$$

are pairwise nearly independent for all g from a large subset of $\mathrm{RM}(r,n)$. This will be the key ingredient to obtain our lower bound for $\mathrm{E}[Y_n]$. We shall complete the proof of Theorem 1 in Section 4.

2. Some results on Reed-Muller codes

In this section, we show that most pairs of functions in $\mathrm{RM}(r,n)$ have Hamming distance close to 2^{n-1}.

The *weight* of a Boolean function f, denoted by $\mathrm{wt}(f)$, is defined to be its Hamming distance to the zero function. For real x, write

$$A_{r,n}(x) = \big|\{g \in \mathrm{RM}(r,n) : \mathrm{wt}(g) \leq 2^n x\}\big|.$$

Our starting point is the following asymptotic characterisation of $A_{r,n}(x)$, which is a special case of a result due to Kaufman, Lovett, and Porat [6].

Lemma 2 ([6, Theorem 3.1])**.** *For all $r \geq 1$, there exists a constant K_r such that*

$$A_{r,n}\left(\frac{1-\delta}{2}\right) \leq \left(\frac{1}{\delta}\right)^{K_r n^{r-1}}$$

for all real δ satisfying $0 < \delta \leq 1/2$.

It should be noted that the case $r = 1$ is not covered in [6, Theorem 3.1]. Lemma 2 however holds trivially in this case, since all but two functions in $\mathrm{RM}(1,n)$ have weight 2^{n-1}.

We now apply Lemma 2 to prove the main result of this section.

Lemma 3. *Let $\epsilon > 0$ be real and let $r \geq 1$ be integral. Then, for all sufficiently large n, there exists a subset $S \subset \mathrm{RM}(r,n)$ of cardinality at least $2^{(1-\epsilon)\binom{n}{r}}$ such that*

$$\big|d(g,h) - 2^{n-1}\big| \leq 2^{n-1}/\tbinom{n}{r} \quad \textit{for all } g,h \in S \textit{ with } g \neq h. \tag{6}$$

Proof. Let $B_{r,n}$ be the number of functions g in $\mathrm{RM}(r,n)$ satisfying

$$\big|\,\mathrm{wt}(g) - 2^{n-1}\big| \geq 2^{n-1}/\tbinom{n}{r}.$$

Since $\mathrm{RM}(r,n)$ contains the nonzero constant function, there is a bijection between the functions in $\mathrm{RM}(r,n)$ of weight w and the functions in $\mathrm{RM}(r,n)$

of weight $2^n - w$. Therefore,

$$B_{r,n} = 2A_{r,n}\left(\frac{1-1/\binom{n}{r}}{2}\right)$$

and so by Lemma 2,

$$\log_2\left(\frac{B_{r,n}}{2}\right) \le K_r n^{r-1}\log_2\binom{n}{r} \le K_r\binom{n}{r}\frac{r^r}{n}\log_2\binom{n}{r},$$

where K_r is the same constant as in Lemma 2. Therefore,

$$B_{r,n} \le 2^{\epsilon\binom{n}{r}} \tag{7}$$

for all sufficiently large n.

Next we construct the set S iteratively as follows. We take n large enough, so that the bound (7) for $B_{r,n}$ holds. Choose a $g \in \mathrm{RM}(r,n)$ to be in S and delete all $u \in \mathrm{RM}(r,n)$ satisfying

$$\left|d(g,u) - 2^{n-1}\right| \ge 2^{n-1}/\tbinom{n}{r}.$$

From (7) it is readily verified that the number of deleted functions is at most $2^{\epsilon\binom{n}{r}}$. We can continue in this way to choose functions of $\mathrm{RM}(r,n)$ to be in S, while maintaining the property (6), as long as the number of chosen functions times $1 + 2^{\epsilon\binom{n}{r}}$ is less than the cardinality of $\mathrm{RM}(r,n)$, namely $2^{1+\binom{n}{1}+\dots+\binom{n}{r}}$. We can therefore obtain a set S satisfying (6) and

$$|S| \ge \frac{2^{1+\binom{n}{1}+\dots+\binom{n}{r}}}{1+2^{\epsilon\binom{n}{r}}} \ge \frac{2^{\binom{n}{r}}}{2^{\epsilon\binom{n}{r}}}$$

for all sufficiently large n. □

3. Some large deviation estimates

In this section, we give some estimates for tail probabilities of sums of independent identically distributed random variables. For $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$, we denote their scalar product by $\langle \mathbf{a}, \mathbf{b}\rangle$.

Lemma 4. *Let $\mathbf{g}$ and $\mathbf{h}$ be elements of $\{-1,1\}^N$ and let X be drawn at random from $\{-1,1\}^N$, equipped with the uniform probability measure. Write $Y_g = \langle X, \mathbf{g}\rangle$ and $Y_h = \langle X, \mathbf{h}\rangle$. Then, for all $t_1, t_2 \in \mathbb{R}$,*

$$\mathrm{E}\left[\exp(t_1 Y_g + t_2 Y_h)\right] \le \exp\left(\tfrac{1}{2}N\left(t_1^2 + t_2^2\right) + t_1 t_2\langle \mathbf{g}, \mathbf{h}\rangle\right).$$

Proof. Write $X = (X_1, \dots, X_N)$, $\mathbf{g} = (g_1, \dots, g_N)$, and $\mathbf{h} = (h_1, \dots, h_N)$. Then

$$\begin{aligned}\mathrm{E}\left[\exp(t_1 Y_g + t_2 Y_h)\right] &= \mathrm{E}\left[\prod_{j=1}^N \exp\left(X_j(t_1 g_j + t_2 h_j)\right)\right]\\ &= \prod_{j=1}^N \mathrm{E}\left[\exp\left(X_j(t_1 g_j + t_2 h_j)\right)\right]\end{aligned}$$

using that the X_j's are independent. Since the X_j's take on each of the values 1 and -1 with probability $1/2$, we see that

$$\mathrm{E}\left[\exp(t_1Y_g + t_2Y_h)\right] = \prod_{j=1}^{N} \cosh(t_1g_j + t_2h_j).$$

By comparing the Maclaurin series of $\cosh(x)$ and $\exp(x^2/2)$, we find that $\cosh(x) \le \exp(x^2/2)$. Thus

$$\begin{aligned}\mathrm{E}\left[\exp(t_1Y_g + t_2Y_h)\right] &\le \prod_{j=1}^{N} \exp\left(\tfrac{1}{2}(t_1g_j + t_2h_j)^2\right)\\ &= \exp\left(\frac{1}{2}\sum_{j=1}^{N}(t_1g_j + t_2h_j)^2\right),\end{aligned}$$

from which the desired bound easily follows. □

We next apply Lemma 4 to vectors $\mathbf{g}$ and $\mathbf{h}$ whose scalar product is sufficiently small.

Lemma 5. *Let $r \ge 0$ be an integer and let $\mathbf{g}$ and $\mathbf{h}$ be elements of $\{-1,1\}^{2^n}$ satisfying $|\langle \mathbf{g}, \mathbf{h}\rangle| \le 2^n/\binom{n}{r}$. Let X be drawn at random from $\{-1,1\}^{2^n}$, equipped with the uniform probability measure. Write $Y_g = \langle X, \mathbf{g}\rangle$, $Y_h = \langle X, \mathbf{h}\rangle$, and*

$$\lambda = \sqrt{2^{n+1}\tbinom{n}{r}\log 2}.$$

Then

$$\Pr\left[Y_g \ge \lambda \cap Y_h \ge \lambda\right] \le 4/4^{\binom{n}{r}}.$$

Proof. Writing $s = \lambda/2^n$, an application of Markov's inequality gives

$$\begin{aligned}\Pr\left[Y_g \ge \lambda \cap Y_h \ge \lambda\right] &= \Pr\left[\exp(sY_g) \ge \exp(s\lambda) \cap \exp(sY_h) \ge \exp(s\lambda)\right]\\ &\le \frac{\mathrm{E}\left[\exp(sY_g)\exp(sY_h))\right]}{[\exp(s\lambda)]^2}\\ &\le \frac{\exp(2^ns^2(1 + 1/\binom{n}{r}))}{[\exp(s\lambda)]^2}\end{aligned}$$

by Lemma 4. This last expression equals $4/4^{\binom{n}{r}}$, as required. □

We also need the following estimate. (Here and in what follows, we use $o(1)$ to denote a suitable nonnegative function of n whose limit equals zero.)

Lemma 6. *Let $X_1, \dots, X_{2^n}$ be independent random variables taking on each of -1 and 1 with probability $1/2$. Then, for all $r \ge 1$, we have, as $n \to \infty$,*

$$\Pr\left[X_1 + \cdots + X_{2^n} \ge \sqrt{2^{n+1}\tbinom{n}{r}\log 2}\right] \ge \frac{1 - o(1)}{2^{\binom{n}{r}}\sqrt{4\pi\binom{n}{r}\log 2}}.$$

Proof. This is a special case of a normal tail approximation of the distribution of $X_1 + \cdots + X_{2^n}$ (see Feller [4, Chapter VII, (6.7)], for example). $\square$

4. Proof of Theorem 1

Recall from the introduction that $Y_{n,g} = 2^n - 2d(f_n, g)$ for $g \in \mathfrak{B}_n$ and

$$Y_n = \max_{g \in \mathrm{RM}(r,n)} Y_{n,g},$$

so that $Y_n = 2^n - 2N_r(f_n)$. Notice that

$$Y_{n,g} = \sum_{x \in \mathbb{F}_2^n} (-1)^{f_n(x)+g(x)}, \tag{8}$$

from which we see that $Y_{n,g}$ is a sum of 2^n random variables, each taking each of the values -1 and 1 with probability $1/2$.

We shall first prove the second part (4) of the theorem by establishing lower and upper bounds for $\mathrm{E}[Y_n]$. The first part (3) will then easily follow from the second part and (5).

To obtain an upper bound for $\mathrm{E}[Y_n]$, let $s \in \mathbb{R}$ and invoke Jensen's inequality to find that

$$\begin{aligned}
\exp(s\,\mathrm{E}[Y_n]) &\le \mathrm{E}\left[\exp(sY_n)\right] \\
&= \mathrm{E}\Big[\max_{g \in \mathrm{RM}(r,n)} \exp(sY_{n,g})\Big] \\
&\le \sum_{g \in \mathrm{RM}(r,n)} \mathrm{E}\left[\exp(sY_{n,g})\right] \\
&\le 2^{1+\binom{n}{1}+\cdots+\binom{n}{r}} \exp(2^{n-1}s^2)
\end{aligned}$$

by Lemma 4 with $t_1 = s$ and $t_2 = 0$ using (8). Hence

$$\mathrm{E}[Y_n] \le \frac{1}{s}\big(1 + \tbinom{n}{1} + \cdots + \tbinom{n}{r}\big) \log 2 + 2^{n-1}s.$$

Now choose s such that both summands are equal. This gives

$$\mathrm{E}[Y_n] \le \sqrt{2^{n+1}\big(1 + \tbinom{n}{1} + \cdots + \tbinom{n}{r}\big) \log 2}. \tag{9}$$

Next we derive a lower bound for $\mathrm{E}[Y_n]$. From Lemma 3 we see that there exists a subset $S_n \subset \mathrm{RM}(r,n)$ satisfying

$$|S_n| = 2^{(1-o(1))\binom{n}{r}}, \tag{10}$$

(where $o(1)$ is a suitable nonnegative function of n tending to zero) such that

$$\Bigg|\sum_{x \in \mathbb{F}_2^n} (-1)^{g(x)+h(x)}\Bigg| \le 2^n/\tbinom{n}{r} \quad \text{for all } g,h \in S_n \text{ with } g \ne h. \tag{11}$$

Writing

$$\lambda_n=\sqrt{2^{n+1}\tbinom{n}{r}\log 2}, \tag{12}$$

we have

$$\begin{aligned}\Pr\big[Y_n\ge\lambda_n\big]&\ge\Pr\big[\max_{g\in S_n}Y_{n,g}\ge\lambda_n\big]\\&\ge\sum_{g\in S_n}\Pr\big[Y_{n,g}\ge\lambda_n\big]-\frac{1}{2}\sum_{\substack{g,h\in S_n\\ g\ne h}}\Pr\big[Y_{n,g}\ge\lambda_n\cap Y_{n,h}\ge\lambda_n\big]\end{aligned}$$

by the Bonferroni inequality. Lemma 6 gives a lower bound for the probabilities in the first sum and, using (8) and (11), Lemma 5 gives an upper bound for the probabilities in the second sum. Applying these bounds gives

$$\Pr\big[Y_n\ge\lambda_n\big]\ge|S_n|\cdot\frac{1-o(1)}{2^{\binom{n}{r}}\sqrt{4\pi\binom{n}{r}\log 2}}-\frac{|S_n|^2}{2}\cdot\frac{4}{4^{\binom{n}{r}}}.$$

Using (10) and observing that the first term dominates the second term, we obtain

$$\Pr\big[Y_n\ge\lambda_n\big]\ge\exp\big(-o(1)\tbinom{n}{r}\big). \tag{13}$$

On the other hand, we find from (5) with $\theta=\lambda_n-\mathrm{E}[Y_n]$ that

$$\Pr\big[Y_n\ge\lambda_n\big]\le 2\exp\left(-\frac{(\lambda_n-\mathrm{E}[Y_n])^2}{2^{n+1}}\right)$$

whenever $\mathrm{E}[Y_n]\le\lambda_n$. Comparison with (13) gives $\mathrm{E}[Y_n]/\lambda_n\ge 1-o(1)$ and combination with (9) gives

$$\lim_{n\to\infty}\mathrm{E}[Y_n]/\lambda_n=1, \tag{14}$$

which proves the second part (4) of the theorem.

To prove the first part (3), we let $\epsilon>0$ and invoke the triangle inequality to obtain

$$\Pr\big[|Y_n/\lambda_n-1|>\epsilon\big]\le\Pr\big[|Y_n-\mathrm{E}[Y_n]|/\lambda_n>\tfrac{1}{2}\epsilon\big]+\Pr\big[|\mathrm{E}[Y_n]/\lambda_n-1|>\tfrac{1}{2}\epsilon\big].$$

By (14), the second probability on the right hand side equals zero for all sufficiently large n, and by (5), the first probability on the right hand side is at most $2\cdot 2^{-(\epsilon^2/4)\binom{n}{r}}$. Hence,

$$\sum_{n=1}^{\infty}\Pr\big[|Y_n/\lambda_n-1|>\epsilon\big]<\infty,$$

from which and the Borel-Cantelli Lemma we conclude that

$$\lim_{n\to\infty}Y_n/\lambda_n=1\quad\text{almost surely.}$$

This proves (3) and completes the proof of the theorem. □

Acknowledgement

I thank Claude Carlet for some careful comments on a draft of this paper.

References

[1] C. Carlet, *The complexity of Boolean functions from cryptographic viewpoint*, Complexity of Boolean Functions (Dagstuhl, Germany), Dagstuhl Seminar Proceedings, no. 06111, 2006.

[2] ———, *Boolean functions for cryptography and error-correcting codes.*, Boolean models and methods in mathematics, computer science, and engineering (Y. Crama and P. L. Hammer, eds.), Cambridge University Press, 2010, pp. 257–397.

[3] S. Dib, *Distribution of Boolean functions according to the second-order nonlinearity*, Arithmetic of finite fields, Lecture Notes in Comput. Sci., vol. 6087, Springer, Berlin, 2010, pp. 86–96.

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

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

[6] T. Kaufman, S. Lovett, and E. Porat, *Weight distribution and list-decoding size of Reed-Muller codes*, IEEE Trans. Inform. Theory **58** (2012), no. 5, 2689–2696.

[7] L. R. Knudsen and M. J. B. Robshaw, *Non-linear approximations in linear cryptanalysis*, Proceedings Eurocrypt'96, Lecture Notes Comput. Sci., vol. 1070, 1996, pp. 224–236.

[8] S. Litsyn and A. Shpunt, *On the distribution of Boolean function nonlinearity*, SIAM J. Discrete Math. **23** (2008/09), no. 1, 79–95.

[9] F. J. MacWilliams and N. J. A. Sloane, *The theory of error-correcting codes*, Amsterdam, The Netherlands: North Holland, 1977.

[10] C. McDiarmid, *On the method of bounded differences*, Surveys in Combinatorics (J. Siemons, ed.), London Math. Soc. Lectures Notes Ser. 141, Cambridge Univ. Press, Cambridge, 1989, pp. 148–188.

[11] D. Olejár and M. Stanek, *On cryptographic properties of random Boolean functions*, J.UCS **4** (1998), no. 8, 705–717 (electronic).

[12] F. Rodier, *Sur la non-linéarité des fonctions booléennes*, Acta Arith. **115** (2004), no. 1, 1–22.

[13] ———, *Asymptotic nonlinearity of Boolean functions*, Des. Codes Cryptogr. **40** (2006), no. 1, 59–70.

Department of Mathematics, University of Paderborn, Warburger Str. 100, 33098 Paderborn, Germany

E-mail address: kus@math.upb.de