On the computation of q-transform of Boolean functions

Zhixiong Chen[*], Ting Gu[*][**], Andrew Klapper[***], and Yu Zhou[****]

[*]Key Laboratory of Applied Mathematics of Fujian Province University, Putian University, P. R. China
[**]Department of Mathematics and Computer Science, College of the Holy Cross, USA
[***]Department of Computer Science, University of Kentucky, USA
[****]Science and Technology on Communication Security Laboratory, Chengdu, P. R. China

Abstract

The *q-transform* of a Boolean function f, introduced by A. Klapper, measures the cross-correlation between f and the functions obtainable from a function q by nonsingular linear change of basis. It is helpful for investigating the higher-order nonlinearity of Boolean functions.

In this work, we employ the cross-correlation theory of Boolean functions to give a relation between the q-transforms and the Walsh-Hadamard transforms. As an application, we illustrate the relation with quadratic functions q which have been well-studied. We prove a (tight) lower bound on the second-order nonlinearity for bent functions.

Preliminary. Let n be a positive integer, let $\mathbb{F}_2^n = \{0,1\}^n$, treated as row vectors, and let $\mathcal{B}_n = \{f : \mathbb{F}_2^n \to \{0,1\}\}$, the set of Boolean functions of dimension n. We refer the reader to Carlet's book chapter [1] and Cusick and Stănică's monograph [5] for background on Boolean functions.

For $f, g \in \mathcal{B}_n$, the *Hamming distance* $d(f,g)$ from f to g, defined by

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

is an important parameter for f and g, where $wt(h)$ is the number of $x \in \mathbb{F}_2^n$ such that $h(x) = 1$. It measures how 'far' from f to g. The equation

$$W(f,g) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x)+g(x)} \in \mathbb{Z}$$

defines the cross-correlation between f and g. It is useful for defining several measures used to analyze Boolean functions. For example, the *nonlinearity* of a function f is defined by

$$nl(f) = \min_{w \in \mathbb{F}_2^n} d(f, w \cdot x).$$

where $w \cdot x = w_1x_1 + w_2x_2 + \ldots + w_nx_n$ is a linear function or zero.

The *Walsh-Hadamard transform coefficient* of f at $w \in \mathbb{F}_2^n$ is defined by

$$S_f(w) = W(f, w \cdot x), \quad w \in \mathbb{F}_2^n.$$

It is a powerful tool for analyzing Boolean functions. For instance, the nonlinearity can be written as

$$nl(f) = 2^{n-1} - \frac{1}{2} \max_{w \in \mathbb{F}_2^n} |S_f(w)|.$$

It is well known that $|S_f(w)| = 2^{n/2}$ for all $w \in \mathbb{F}_2^n$ if and only if f is bent.

[*]corresponding author: Ting Gu, tgu@holycross.edu

Let $deg(f)$ be the algebraic degree[1] of f. More generally the *rth-order (higher order) nonlinearity* of f, denoted by $nl_r(f)$, is defined as

$$nl_r(f) = \min_{\deg(g)\leq r} d(f,g).$$

We have

$$nl_r(f) = 2^{n-1} - \frac{1}{2}\max_{\deg(g)\leq r} |W(f,g)|. \tag{1}$$

Thus $nl_r(f)$ should be as large as possible (at least for small r) for the requirements of cryptographic security. However, there is no known efficient algorithm that computes $nl_r(f)$ even if $r = 2$. Quoting Carlet [2, pp.1262-1263],

> Computing the r-th order nonlinearity of a given function with algebraic degree strictly greater than r is a hard task for $r > 1$. In the case of the first order, much is known in theory and also algorithmically since the nonlinearity is related to the Walsh transform, which can be computed by the algorithm of the fast Fourier Transform (FFT).$\cdots$ But for $r > 1$, very little is known. Even the second order nonlinearity is known only for a few peculiar functions and for functions in small numbers of variables.

Luckily, we find that the *q-transform* for Boolean functions, introduced by Klapper [7], might be useful for computing the r-th order nonlinearity. This leads to the main contribution of this work.

Let $GL_n = GL_n(\mathbb{F}_2)$ be the group of all $n \times n$ invertible matrices over $\mathbb{F}_2$. Let $N = |GL_n| = (2^n - 1)(2^n - 2)(2^n - 4)\cdots(2^n - 2^{n-1})$, the cardinality of GL_n. Let q_A denote the function $q_A(x) = q(xA)$ for $q \in \mathcal{B}_n$ and $A \in GL_n$. Let $\mathbf{0}$ denote the $n \times n$ zero matrix.

Definition 0.1 *([7]) Let $q \in \mathcal{B}_n$. If $f \in \mathcal{B}_n$ and $A \in GL_n$, then the* q-transform coefficient of f at A *is $W_q(f)(A) = W(f, q_A)$. Also let*

$$W_q(f)(\mathbf{0}) = \sum_{x\in\mathbb{F}_2^n} (-1)^{f(x)} \triangleq I_f,$$

where I_f is referred to as the imbalance *of f. Then q-transform is the list of $W_q(f)(A)$ for all $A \in GL_n$.*

If q is linear, then the q-transform is essentially the Walsh-Hadamard transform.

From eqn. (1), the rth-order nonlinearity can be written as

$$nl_r(f) = 2^{n-1} - \frac{1}{2}\max_{\deg(q)\leq r} \max_{A\in GL_n\cup\{\mathbf{0}\}} |W_q(f)(A)|. \tag{2}$$

Klapper considered the statistical behavior of the q-transform with respect to two probability distributions [7]. For a random variable X on $GL_n \cup \{\mathbf{0}\}$, let $E'[X]$ denote the expected value of X with respect to the uniform distribution on GL_n. Let $E[X]$ denote the expected value of X on $GL_n \cup \{\mathbf{0}\}$ with respect to ω, which is the probability distribution on $GL_n \cup \{\mathbf{0}\}$, defined by

$$\omega(A) = \frac{1}{N + N/(2^n - 1)} = \frac{2^n - 1}{2^n N}$$

for $A \in GL_n$ and

$$\omega(\mathbf{0}) = \frac{N/(2^n - 1)}{N + N/(2^n - 1)} = \frac{1}{2^n},$$

[1]The *degree* of Boolean functions, is defined as the number of variables in the largest product term of the functions' algebraic normal form (ANF) having a non-zero coefficient.

where $N = |GL_n|$. The choice of ω comes from the case when q is linear, so that we can consider the q-transform $W_q(f)$ as a generalization of the Walsh-Hadamard transform $S_f(w)$.

Then, for balanced $q \in \mathcal{B}_n$, by some computations involving ω that we shall omit [7, p. 2801], we get

$$E'[W_q(f)(A)^2] = (2^{2n} - I_f^2)/(2^n - 1), \tag{3}$$

and so

$$E[W_q(f)(A)^2] = 2^n. \tag{4}$$

This is a generalization of Parseval's equation.

Certain basic topics involving q-transforms have appeared in the literature. For example, we computed the auto-correlation of q-bent functions (if they exist) [8]. We answered the question as to whether q-bent functions exist [3]. We investigated q-nearly bent functions in [4] and the q-correlation immune functions in [6]. However, many open questions remain, many described in our previous work [3, 4, 6, 7, 8]. For example, how can we model other measures of cryptographic security using q-transforms? These questions still remain interesting when restricted to simple q such as x_1x_2 or $x_1x_2 + x_3$.

A computational formula of q-transform coefficients. The theorem below indicates a relation between the q-transform and the Walsh-Hadamard transform, which is a direct result coming from the cross-correlation of two functions. The proof is omitted due to space limits.

Theorem 0.2 *Let n be an integer and $q \in \mathcal{B}_n$. Then the q-transform coefficient of $f \in \mathcal{B}_n$ at $A \in GL_n$ is*

$$W_q(f)(A) = 2^{-n} \sum_{w \in \mathbb{F}_2^n} S_f(w) S_{q_A}(w).$$

Theorem 0.2 shows that the computation of the q-transform can be based on the computation of the Walsh-Hadamard transform. A fast method for computing the Walsh-Hadamard transform is the divide-and-conquer butterfly algorithm, known as the *Fast Fourier Transform* (FFT) [10, Theorem 5, p. 422], [1, Sect. 2.2]. So the overall time complexity is $O(n2^n)$ for computing $W_q(f)(A)$ for any $A \in GL_n$. This gives us a feasible method to calculate the q-transform (for fixed f and q). And thus it might be helpful to compute $nl_r(f)$ (in particular $nl_2(f)$) for some special f.

We can efficiently compute the q-transform coefficients $W_q(f)(A)$ if both f and q have few non-zero Walsh-Hadamard transform coefficients. We introduce the parameter

$$\Delta_h =: |\{w \in \mathbb{F}_2^n : S_h(w) \neq 0\}|.$$

Theorem 1 in [9] told us that $\Delta_h \geq 4$ for all non-affine functions h and no h has $\Delta_h \in \{2, 3, 5, 6, 7\}$. When $\Delta_h = 4$, we see that the four non-zero Walsh-Hadamard transform coefficients of h are

$$(2^{n-1}, 2^{n-1}, 2^{n-1}, -2^{n-1}) \text{ or } (-2^{n-1}, -2^{n-1}, -2^{n-1}, 2^{n-1}).$$

In particular, we find $\Delta_q = 4$ for $q(x) = x_1x_2$ or $q(x) = x_1x_2 + x_3$.

Next we compute the q-transform coefficients of a bent function for those $q \in \mathcal{B}_n$ with $\Delta_q = 4$.

Theorem 0.3 *Let $n \geq 4$ be even and $q \in \mathcal{B}_n$ with $\Delta_q = 4$. If f is a bent function, then the q-transform coefficient of f satisfies*

$$W_q(f)(A) \in \{0, \pm 2^{n/2}, \pm 2^{n/2+1}\}, \quad A \in GL_n.$$

Example 0.4 *Let $n = 4$, let $A \in GL_n$, and let $f(x) = x_1x_2 + x_3x_4$. Then f is bent.*

If $q = x_1x_2$, then f has a four-valued q-transform spectrum with $W_q(f)(A) \in \{0, \pm 4, 8\}$.

If $q = x_1x_2 + x_3$, then f has five-valued q-transform spectrum with $W_q(f)(A) \in \{0, \pm 4, \pm 8\}$.

Similar to Theorem 0.3, we have the following for plateaued functions[2] (including semi-bent functions).

Theorem 0.5 *Let $n \geq 4$ and $q \in \mathcal{B}_n$ with $\Delta_q = 4$. If f is a plateaued function with $S_f(w) \in \{0, \pm\lambda\}$ for $w \in \mathbb{F}_2^n$, then the q-transform coefficient of f satisfies*

$$W_q(f)(A) \in \left\{0, \pm\frac{1}{2}\lambda, \pm\lambda, \pm\frac{3}{2}\lambda, \pm 2\lambda\right\}, \quad A \in GL_n.$$

Example 0.6 *Let $n = 4$ and $f(x) = x_1 + x_2 + x_3x_4$ be plateaued (with $\lambda = 8$). Then f has a five-valued q-transform spectrum with $W_q(f)(A) \in \{0, \pm 4, \pm 8\}$ for $q = x_1x_2$ and $q = x_1x_2 + x_3$, respectively.*

Second-order nonlinearity of bent functions. Proving general upper or lower bounds on the higher-order nonlinearity of functions is also a difficult task, even for second-order nonlinearity[1]. Here we use Theorem 0.2 to prove a general lower bound on the second-order nonlinearity of bent functions. Based on some examples, we show that the bound is tight.

The computation of the second-order nonlinearity of functions relates to quadratic functions. For quadratic $h \in \mathcal{B}_n$, we define

$$Ker(h) = \{x \in \mathbb{F}_2^n : h(x + y) + h(x) + h(y) = 0 \text{ for any } y \in \mathbb{F}_2^n\},$$

the *kernel* of h. $Ker(h)$ is a linear subspace of $\mathbb{F}_2^n$. For each $0 \leq k < n$ with $k \equiv n \pmod 2$, there always exists a quadratic h such that the dimension of $Ker(h)$ equals k.

Theorem 0.7 *Let $n \geq 6$ be even and $q \in \mathcal{B}_n$ be a quadratic function with kernel $Ker(q)$ of dimension k. If $f \in \mathcal{B}_n$ is a bent function of degree $d \geq 3$, then the q-transform coefficient of f satisfies*

$$|W_q(f)(A)| \leq \begin{cases} 2^{n-k/2}, & \text{if } k > 0, \\ 2^n - 2^{n-d+1}, & \text{if } k = 0, \end{cases}$$

for all $A \in GL_n$.

Theorem 0.7 leads to a lower bound on the second-order nonlinearity of bent functions.

Theorem 0.8 *Let $n \geq 6$ be even and $f \in \mathcal{B}_n$ be bent of degree $d \geq 3$. Then the second-order nonlinearity of f satisfies*

$$nl_2(f) \geq 2^{n-d}.$$

Proof. We use eqn. (1). Since f is bent, we have

$$\max_{\deg(g)\leq 1} |W(f, g)| = \max_{w\in\mathbb{F}_2^n} |S_f(w)| = 2^{n/2}. \tag{5}$$

From Theorem 0.7, we have

$$\max_{\deg(g)=2} |W(f, g)| = \max_{\deg(q)=2} \max_{A\in GL_n} |W_q(f)(A)| \leq 2^n - 2^{n-d+1}. \tag{6}$$

Putting eqns. (1), (5) and (6) together, we get

$$nl_2(f) = 2^{n-1} - \frac{1}{2} \max_{\deg(g)\leq 2} |W(f, g)| \geq 2^{n-1} - (2^{n-1} - 2^{n-d}) = 2^{n-d}.$$

This completes the proof.

We give an example to illustrate that the lower bound in Theorem 0.8 is tight.

[2] A function $f \in \mathcal{B}_n$ is called a plateaued function if $S_f(w) = \{0, \pm\lambda\}$ for $w \in \mathbb{F}_2^n$ where $\lambda \geq 2^{n/2}$. In particular, a plateaued function is semi-bent if $\lambda = 2^{(n+1)/2}$ for odd n or if $\lambda = 2^{(n+2)/2}$ for even n.

Example 0.9 *Let $n = 2m \geq 6$ and $3 \leq d \leq m$. Let*

$$q(x, y) = \sum_{i=1}^{m} x_i y_i$$

and let

$$f(x, y) = y_1 y_2 \cdots y_d + \sum_{i=1}^{m} x_i y_i, \quad 3 \leq d \leq m$$

for $x = (x_1, x_2, \ldots, x_m)$ and $y = (y_1, y_2, \ldots, y_m)$. Then f is bent of degree d and q is quadratic bent. Since $y_1 y_2 \cdots y_d \in \mathcal{B}_n$ has weight $wt(y_1 y_2 \cdots y_d) = 2^{n-d}$, we have

$$W_q(f)(E) = \sum_{(x,y) \in \mathbb{F}_2^n} (-1)^{f(x,y)+q_E(x,y)} = \sum_{(x,y) \in \mathbb{F}_2^n} (-1)^{y_1 y_2 \cdots y_d} = 2^n - 2^{n-d+1},$$

where $E \in GL_n$ is the identity matrix. This means that there is a bent function whose q-transform coefficient is maximal. So the f in this example has second-order nonlinearity $nl_2(f) = 2^{n-d}$.

References

[1] C. Carlet. Boolean Functions for Cryptography and Error Correcting Codes. Available at: http://www.math.univ-paris13.fr/~carlet/pubs.html. (2010)

[2] C. Carlet. Recursive lower bounds on the nonlinearity profile of Boolean functions and their applications. IEEE Trans. Inf. Theory, 2008, 54(3) : 1262-1272.

[3] Z. Chen, T. Gu and A. Klapper. On q-bentness of Boolean functions. Designs, Codes Cryptogr., 2019, 87(1): 163-171.

[4] Z. Chen, A. Klapper. On q-nearly bent Boolean functions. Discrete Appl. Math., 2020, 279: 210-217.

[5] T. Cusick and P. Stănică. Cryptographic Boolean Functions and Applications. Academic Press, Amsterdam, 2009.

[6] T. Gu, Z. Chen and A. Klapper. Correlation immune functions with respect to the q-transform. Cryptogr. Commun., 2018, 10(6): 1063-1073.

[7] A. Klapper. A new transform related to distance from a Boolean function. IEEE Trans. Inf. Theory, 2016, 62(5): 2798-2812.

[8] A. Klapper and Z. Chen. On the nonexistence of q-bent Boolean functions. IEEE Trans. Inf. Theory, 2018, 64(4): 2953-2961.

[9] D. Pei, W. Qin. The correlation of a Boolean function with its variables. International Conference on Cryptology in India, pp. 1-8. Springer, Berlin, Heidelberg, 2000.

[10] F. J. MacWilliams, N. J. A. Sloane. The Theory of Error-correcting Codes. North-Holland Publishing Company, Amsterdam, 1978.