

## Testing Boolean Functions Properties

**Xie Zhengwei**

*School of Computer Science and Engineering  
Sun Yat-sen University, Guangzhou 510006, China  
weixzh2010@163.com*

**Cai Guangya**

*School of Computer Science and Engineering  
Sun Yat-sen University, Guangzhou 510006, China  
caigya@mail2.sysu.edu.cn*

**Paulo Mateus**

*Instituto de Telecomunicações  
Dept. de Matemática, Instituto Superior Técnico  
Av. Rovisco Pais 1049-001 Lisbon, Portugal  
pmat@math.ist.utl.pt*

**Qiu Daowen<sup>†\*</sup>**

*School of Computer Science and Engineering  
Sun Yat-sen Univ., Guangzhou 510006, China  
issqdw@mail.sysu.edu.cn*

**Jozef Gruska**

*Faculty of Informatics  
Masaryk University, Brno, Czech Republic  
gruska@fi.muni.cz*

**Abstract.** The goal in the area of functions property testing is to determine whether a given black-box Boolean function has a particular given property or is  $\varepsilon$ -far from having that property. We investigate here several types of properties testing for Boolean functions (identity, correlations and balancedness) using the Deutsch-Jozsa algorithm (for the Deutsch-Jozsa (D-J) problem) and also the amplitude amplification technique.

At first, we study here a particular testing problem: namely whether a given Boolean function  $f$ , of  $n$  variables, is identical with a given function  $g$  or is  $\varepsilon$ -far from  $g$ , where  $\varepsilon$  is the parameter. We present a one-sided error quantum algorithm to deal with this problem that has the query complexity  $O(\frac{1}{\sqrt{\varepsilon}})$ . Moreover, we show that our quantum algorithm is optimal. Afterwards we show that the classical randomized query complexity of this problem is  $\Theta(\frac{1}{\varepsilon})$ . Secondly, we consider the D-J problem from the perspective of functional correlations and let  $C(f, g)$  denote the

\*Address for correspondence: School of Computer Sci. and Eng., Sun Yat-sen University, Guangzhou 510006, China.

<sup>†</sup> Also works: Instituto de Telecomunicações, Dept. de Matemática, Instituto Superior Técnico, Lisbon, Portugal.

This work is partly supported by the National Natural Science Foundation of China (Nos. 61572532, 61876195), the Natural Science Foundation of Guangdong Province of China (No. 2017B030311011).

correlation of  $f$  and  $g$ . We propose an exact quantum algorithm for making distinction between  $|C(f, g)| = \varepsilon$  and  $|C(f, g)| = 1$  using six queries, while the classical deterministic query complexity for this problem is  $\Theta(2^n)$  queries. Finally, we propose a one-sided error quantum query algorithm for testing whether one Boolean function is balanced versus  $\varepsilon$ -far balanced using  $O(\frac{1}{\varepsilon})$  queries. We also prove here that our quantum algorithm for balancedness testing is optimal. At the same time, for this balancedness testing problem we present a classical randomized algorithm with query complexity of  $O(1/\varepsilon^2)$ . Also this randomized algorithm is optimal. Besides, we link the problems considered here together and generalize them to the general case.

**Keywords:** Deutsch-Jozsa Algorithm, Quantum amplitude amplification, Identity testing, Correlation testing, Balancedness testing

## 1. Introduction

We deal here with the function isomorphism testing problem in a new way. This problem was at first explored by Fischer *et al.* [1]. Function isomorphism testing is to determine whether a given Boolean function  $f$  is isomorphic to a given specific function  $g$  (This means that it is equal to  $g$  up to a permutation of its input variables). So far only a few simple testing problems have been already well understood. For example, it has been shown in [2] that partially symmetric functions are isomorphism testable using a constant number of queries.

Moreover, the following upper bound on the number of queries needed for isomorphism testing has been shown: for any integer  $k$ , if  $g$  is a so-called  $k$ -junta (that is if  $g$  depends on at most  $k$  variables), then it is possible to test  $g$ -isomorphism with  $\text{poly}(k/\varepsilon)$  number of queries. This bound was recently improved by Alon *et al.* [3]. They showed that  $O(k \log k)$  queries are sufficient for such testing. Recently, Chen *et al.* [4] proved that any non-adaptive algorithm that tests whether an unknown Boolean function is a  $k$ -junta, or it is  $\varepsilon$ -far from every  $k$ -junta, must make  $\tilde{\Omega}(k^{3/2}/\varepsilon)$  queries for a wide range of parameters  $k$  and  $\varepsilon$ . A more generalize tolerant testing of  $k$ -junta was investigated in the reference [5]. Blais and O'Donnell [6], Alon *et al.* [3] proved more general lower bounds for the isomorphism testing. In particular, Blais *et al.* [2] showed that all partially symmetric functions were efficiently isomorphism testable (*i.e.* - with a constant number of queries). In addition, there have been already also a lot of results on other types of testing. For example, on the property testing such as linearity testing [7], monotonicity testing [8, 9], group testing [10] and so on.

Quantum query complexity [11, 12] is a well-known black-box model for quantum computation, in which the resource measured is the number of queries needed to compute a function. In particular, one provides a "black-box access" to a function  $f$ , meaning that the quantum algorithm can apply some unitary transformation  $U_f$  that maps basis states of the form  $|x\rangle|y\rangle$  to basis states of the form  $|x\rangle|y \oplus f(x)\rangle$  (or  $|x\rangle$  to  $(-1)^{f(x)}|x\rangle$ , where  $f$  is a Boolean function). This complexity model has been behind very first great algorithmic successes of quantum computing like in the search algorithm of Grover [13] and in the algorithm for period finding which is a subroutine of Shor's quantum factoring algorithm [14].

A natural goal when such a model is considered is to minimize the number of queries to the oracle that are needed to solve the given problem. This minimum is the query complexity of the problem.

However, quantum algorithms for property testing were much less studied than classical algorithms for this problem in the past. However, quantum property testing has been receiving increasing attentions in the last few years, both for testing properties of the classical objects and also for testing properties of quantum objects. For example, Buhrman *et al.* [15] showed that there exist languages with efficient quantum property testers but without so efficient classical testers. Hillery and Andersson [16] presented two quantum algorithms for testing the linearity and the permutation invariance of Boolean functions. Moreover, Bravyi *et al.* [17] described quantum algorithms for testing properties of distributions. More about property testing can be found in [18].

A more general concept, and also very interesting and important, than functions isomorphism is the affine equivalence of Boolean functions. Affine equivalence classification of Boolean functions has significant applications in logic synthesis and in cryptography. The affine equivalence classification of the cosets has been studied, for example, by Hou [19] and Zhang [20]. It is easy to see that if  $f$  and  $g$  are affine equivalent, then they have the same weight and nonlinearity [21]. Canright *et al.* [22] gave a simple necessary and sufficient condition for deciding whether two monomial rotation symmetric Boolean functions are affine equivalent. Fuller [23] made a detailed analysis of affine equivalent Boolean functions for cryptography.

Of a special interest are also the following results. Batu *et al.* [24] considered the problem of testing distribution identity. Diakonikolas *et al.* [25] studied the problem of identity testing for structured distributions. Moreover, in this paper, we focus on the identity testing for Boolean functions (observe that the identity can be seen as a special case of the affine equivalence and isomorphism).

Moreover, Cai *et al.* [26] computed the correlation between any two symmetric Boolean functions. They showed that every symmetric Boolean function having an odd period has an exponentially small correlation with the parity function. Castro *et al.* [27] computed the asymptotic behavior of symmetric Boolean functions and derived a formula that allows to determine if a symmetric Boolean function is asymptotically not balanced. For example, for any integers  $k_1, k_2$ ,  $\lim_{n \rightarrow \infty} C(\sigma_{n,k_1}, \sigma_{n,k_2}) = 0$ , but  $C(\sigma_{n,k_1}, \sigma_{n,k_2}) \neq 0$ , where each  $\sigma_{n,k}$  is an elementary symmetric polynomial of  $n$  variables and of the degree  $k$  and  $C(\sigma_{n,k_1}, \sigma_{n,k_2})$  denotes the correlation between these two functions. Castro *et al.* [28] investigated the asymptotic behavior of  $C(F_1, F_2)$ , where  $F_1 = \sigma_{n,k_1} + \dots + \sigma_{n,k_s}$ ,  $F_2 = F(X_1, \dots, X_j)$  is a Boolean function in the first  $j$ -variables ( $j$  fixed). Recently, Castro *et al.* [29] showed that the generalized Walsh transforms of symmetric and also rotation symmetric Boolean functions are linearly recurrent. This subject can be studied from the point of view of complexity theory or from the algebraic point of view. In this paper, we study such correlations from the perspective of property testing.

Balancedness is one of other important cryptographic properties of Boolean functions. There have been already many results, see for example [20, 30] concerning the balancedness of Boolean functions in cryptography. They also inspired us to further study balancedness from the perspective of the property testing. In particular, we consider here quantum and also classical randomized algorithms for balancedness testing.

The remainder of this paper is organized as follows. In Section 2, we present some of the notations and definitions that will be used in the rest of the paper, as well as a review of the Deutsch-Jozsa algorithm and the quantum amplitude amplification theorem. Then, in Section 3, a quantum algorithm for testing Boolean functions identity is given and explored. Moreover, we provide the lower

bound of testing Boolean functions identity. We further prove that the randomized query complexity of the problem is  $\Theta(\frac{1}{\varepsilon})$ . Also, we compare the complexity of testing identity of distributions with the complexity of testing identity of Boolean functions. In Section 4 and Section 5, we consider correlation testing problems and balancedness testing problems, respectively. Finally, conclusions and open problems are summarized in Section 6.

## 2. Preliminaries

In this section, we first introduce some basic notions related to Boolean functions and quantum computation. For more details, see literature [11, 23].

Let  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  be a Boolean function of  $n$  variables, and let  $B_n$  denote the set of all Boolean functions of  $n$  variables.

**Definition 2.1. (Property tester, [8])** Let  $P_n$  be a set of Boolean functions of  $n$  variables with a certain property. An  $\varepsilon$ -tester,  $\varepsilon > 0$ , for  $P_n$  is a randomized or quantum algorithm which can query an unknown Boolean function  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  on a small number of inputs and

- (1) accepts  $f$  with the probability at least  $\frac{2}{3}$  when  $f \in P_n$ ,
- (2) rejects  $f$  with the probability at least  $\frac{2}{3}$  when  $f$  is  $\varepsilon$ -far from any function in  $P_n$ ,

where “ $f$  is  $\varepsilon$ -far from  $P_n$ ” means that  $\text{Ham}(f, g) = |\{x \in \{0, 1\}^n : f(x) \neq g(x)\}| \geq \varepsilon 2^n$  holds for every  $g \in P_n$ . We also denote by  $\text{Dist}(f, g) = \text{Ham}(f, g)/2^n$ .

**Definition 2.2.** When  $x$  is a string, let  $|x|$  denote the Hamming weight of the string  $x$ , i.e., the number of ones in the string  $x$ .

**Definition 2.3.** Let  $x = (x_1, \dots, x_n)$  and  $\omega = (\omega_1, \dots, \omega_n)$  both belong to  $\{0, 1\}^n$ . The inner product of  $x$  and  $\omega$  is defined as  $x \cdot \omega = x_1\omega_1 \oplus \dots \oplus x_n\omega_n$ , where  $\oplus$  is addition modulo 2.

**Definition 2.4.** Suppose  $f \in B_n$ . The Walsh transform of  $f$ , is defined as

$$W_f(\omega) = \frac{1}{2^n} \sum_{x \in \{0, 1\}^n} (-1)^{f(x) + \omega \cdot x}, \text{ where } \omega \in \{0, 1\}^n.$$

**Definition 2.5. ([28])** The correlation  $C(f, g)$  between two Boolean functions  $f$  and  $g$  of  $n$  variables is defined as the number of arguments for which functions  $f$  and  $g$  have the same value, minus the number of arguments on which they disagree, all divided by  $2^n$ , i.e.,

$$C(f, g) = \frac{1}{2^n} \sum_{x \in \{0, 1\}^n} (-1)^{f(x) + g(x)}.$$

For the case of one function  $f$ , let's denote  $C(f) = \frac{1}{2^n} \sum_{x \in \{0, 1\}^n} (-1)^{f(x)}$ . A Boolean function  $f(x)$  is called balanced if  $C(f) = 0$ .

**Remark 2.6.** According to Definition 2.4 and 2.5, we have  $C(f) = W_f(0)$ .

**Definition 2.7.** ([6]) Two Boolean functions  $f, g \in B_n$  are said to be isomorphic if they are identical up to any reordering of input variables. More precisely, we say that  $f, g \in B_n$  are isomorphic (to each other) if there is a permutation  $\sigma$  on  $[n]$  such that for every  $x = (x_1, x_2, \dots, x_n) \in \{0, 1\}^n$ ,  $f(x_1, x_2, \dots, x_n) = g(x_{\sigma_1}, x_{\sigma_2}, \dots, x_{\sigma_n})$ , where  $\sigma(1, 2, \dots, n) = (\sigma_1, \sigma_2, \dots, \sigma_n)$ .

**Definition 2.8.** ([31]) Two Boolean functions  $f(x_1, x_2, \dots, x_n)$  and  $h(x_1, x_2, \dots, x_n)$  are said to be affine equivalent if  $h(x_1, x_2, \dots, x_n)$  can be written as  $h(x_1, x_2, \dots, x_n) = h(X) = f(AX + b)$ , where  $A$  is an  $n \times n$  nonsingular matrix over the Boolean field  $F_2$ ,  $X$  is a column vector whose transpose is  $X^T = [x_1, x_2, \dots, x_n]$ , and  $b$  is an  $n$ -dimensional vector over  $F_2$ . The addition  $AX + b$  is also over  $F_2$ .

Affine equivalent functions have many similar cryptographic properties, such as correlation immunity, resiliency and propagation characteristics (For example, see [31]). Moreover, Zhang et al.[31] presented a new method for computing affine equivalence classes of Boolean functions due to a group isomorphism. In the following proposition, we establish a connection between the isomorphism relation and the affine equivalence relation of two Boolean functions.

**Proposition 2.9.** Let  $f, g \in B_n$ . Then  $f$  is isomorphic to  $g$  if and only if there exists  $A = \begin{pmatrix} e_{\sigma(1)} \\ e_{\sigma(2)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix}$ ,  $b = \mathbf{0}$  such that  $f(X) = g(AX + b)$ , where  $e_1 = (1, 0, \dots, 0), \dots, e_n = (0, 0, \dots, 1)$  and  $\sigma$  is a permutation on  $[n]$ .

**Proof:**

This follows easily from Definition 2.7 and Definition 2.8. □

**Definition 2.10.** ([32]) Let  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  be a partial Boolean function, and  $D \subseteq \{0, 1\}^n$  be its domain of definition. If for any  $x \in D$  and any  $y \in \{0, 1\}^n$  with  $|x| = |y|$ , it holds that  $y \in D$  and  $f(x) = f(y)$ , then  $f$  is called a *symmetric partial Boolean function*. When  $D = \{0, 1\}^n$ ,  $f$  is a symmetric function.

**Definition 2.11.** ([12, 32]) Let  $f$  be a partial Boolean function with a domain of definition  $D \subseteq \{0, 1\}^n$ . We say a real multilinear polynomial  $p$  approximates  $f$  with the error  $1/3$  if:

- (1)  $|p(x) - f(x)| \leq 1/3$  for all  $x \in D$ ;
- (2)  $0 \leq p(x) \leq 1$  for all  $x \in \{0, 1\}^n$ .

The approximate degree of  $f$  with the error  $1/3$ , denoted by  $\widetilde{\text{deg}}(f)$ , is the minimum degree among all real multilinear polynomials that approximate  $f$  with the error  $1/3$ .

**Lemma 2.12.** ([12, 32]) Let  $f$  be a symmetric partial Boolean function over  $\{0, 1\}^n$  with the domain of definition  $D$ , and  $\widetilde{\deg}(f) = d$ . Then there exists a real multilinear polynomial  $q$  that approximates  $f$  with the error  $1/3$  and  $q$  can be written as

$$q(x) = c_0 + c_1 V_1 + c_2 V_2 + \cdots + c_d V_d, \quad (1)$$

where  $c_i \in \mathbb{R}$ ,  $V_j$  denotes the sum of all products of  $j$  different variables, i.e.,  $V_1 = x_1 + \cdots + x_n$ ,  $V_2 = x_1 x_2 + x_1 x_3 + \cdots + x_{n-1} x_n$ , etc.

**Remark 2.13.** Note that  $V_j$  assumes value  $\binom{|x|}{j} = |x|(|x| - 1)(|x| - 2) \cdots (|x| - j + 1)/j!$  for any string  $x$ , and  $\binom{|x|}{j}$  is a polynomial of degree  $j$  of  $|x|$ . Hence  $\phi(|x|) = c_0 + c_1 \binom{|x|}{1} + c_2 \binom{|x|}{2} + \cdots + c_d \binom{|x|}{d} = q(x)$ . Therefore we can obtain a single-variate polynomial  $\phi(|x|)$  which approximates  $f$  with the error  $1/3$  such that  $\deg(\phi) = \widetilde{\deg}(f)$ .

**Lemma 2.14.** ([33, 34]) Let  $p$  be a polynomial with the following properties:

1. For any integer  $0 \leq i \leq n$ , we have  $b_1 \leq p(i) \leq b_2$  for some fixed  $b_1$  and  $b_2$ .
2. For some real  $0 \leq x \leq n$ , the derivative  $p'(x)$  of  $p$  satisfies  $|p'(x)| \geq c$  for some  $c$ .

Then,  $\deg(p) \geq \sqrt{cn/(c + b_2 - b_1)}$ .

**Lemma 2.15.** ([32]) For any partial Boolean function  $f$ ,  $\widetilde{Q}(f) \geq \frac{1}{2} \widetilde{\deg}(f)$ , where  $\widetilde{Q}(f)$  denotes the quantum query complexity for  $f$  with the bounded-error  $1/3$ .

We review now briefly the Deutsch-Jozsa algorithm [35] task and the quantum amplitude amplification [36]. The Deutsch-Jozsa algorithm is to distinguish for a given Boolean function  $f$  one of two cases with certainty using only one query, under the promise that the function  $f$  is either constant or balanced. (Here, we do not describe details of the D-J algorithm - they will not be of importance in the following.)

Let  $\mathcal{A}$  be any quantum algorithm (a unitary operator) that acts on a Hilbert space  $\mathcal{H}$  and uses no measurement. Let  $|\Psi\rangle = \mathcal{A}|0\rangle$  denote the state obtained by applying the unitary  $\mathcal{A}$  to the initial zero state. The amplification process is realized by repeating an application of the following unitary operator on the state  $|\Psi\rangle$ ,

$$Q = -\mathcal{A}S_0\mathcal{A}^{-1}S_\chi, \quad (2)$$

where the operator  $S_\chi$  conditionally changes the phase of so called good states,

$$|x\rangle \rightarrow \begin{cases} -|x\rangle & \text{if } \chi(x) = 1, \\ |x\rangle & \text{if } \chi(x) = 0, \end{cases}$$

while the operator  $S_0$  changes the sign of the amplitudes if and only if the state is the zero state  $|0\rangle$ . Write  $|\Psi\rangle = |\Psi_1\rangle + |\Psi_0\rangle$  is a superposition of so called good and so called bad components (basis substates) of  $|\Psi\rangle$ . We have that

$$Q|\Psi_1\rangle = (1 - 2a)|\Psi_1\rangle - 2a|\Psi_0\rangle, \quad (3)$$

$$Q|\Psi_0\rangle = 2(1 - a)|\Psi_1\rangle + (1 - 2a)|\Psi_0\rangle, \quad (4)$$

$$Q^j|\Psi\rangle = \frac{1}{\sqrt{a}} \sin((2j + 1)\theta_a)|\Psi_1\rangle + \frac{1}{\sqrt{1 - a}} \cos((2j + 1)\theta_a)|\Psi_0\rangle, \quad (5)$$

where  $\sin^2(\theta_a) = a = \langle\Psi_1|\Psi_1\rangle$ .

**Lemma 2.16. (Quadratic speedup [36])** Let  $\mathcal{A}$  be any quantum algorithm that uses no measurements, and let  $\chi : Z \rightarrow \{0, 1\}$  be any Boolean function. Let  $a$  be the initial success probability of  $\mathcal{A}$ . Suppose  $a > 0$ , and set  $m = \lfloor \pi/4\theta_a \rfloor$ , where  $\theta_a$  is defined such that  $\sin^2(\theta_a) = a$  and  $0 < \theta_a \leq \pi/2$ . Then, if we compute  $Q^m\mathcal{A}|0\rangle$  and measure the system, the outcome is good with probability at least  $\max(1 - a, a)$ .

When the success probability  $a$  of quantum algorithm  $\mathcal{A}$  is known, we redefine the unitary operator  $Q$  to find a good solution with certainty. Let

$$Q = Q(\mathcal{A}, \phi, \varphi, \chi) = -\mathcal{A}S_0(\phi)\mathcal{A}^{-1}S_\chi(\varphi). \quad (6)$$

Here the operator  $S_\chi(\varphi)$  conditionally changes the phase of good states,

$$|x\rangle \rightarrow \begin{cases} e^{i\varphi}|x\rangle & \text{if } \chi(x) = 1 \\ |x\rangle & \text{if } \chi(x) = 0, \end{cases}$$

while the operator  $S_0(\phi)$  multiplies the amplitude by a factor of  $e^{i\phi}$  iff the initial state is the zero state  $|0\rangle$ .

**Lemma 2.17.** [36] Let  $Q = Q(\mathcal{A}, \phi, \varphi, \chi)$  be defined as above, then

$$Q|\Psi_1\rangle = e^{i\varphi}((1 - e^{i\phi})a - 1)|\Psi_1\rangle + e^{i\varphi}(1 - e^{i\phi})a|\Psi_0\rangle \quad (7)$$

$$Q|\Psi_0\rangle = (1 - e^{i\phi})(1 - a)|\Psi_1\rangle - ((1 - e^{i\phi})a + e^{i\phi})|\Psi_0\rangle, \quad (8)$$

where  $a = \langle\Psi_1|\Psi_1\rangle$ .

### 3. Testing identity of two unknown Boolean functions

Testing of the function isomorphism with classical or quantum algorithms has already been investigated in the literature [6, 2]. In spite of the fact that the affine equivalence has been already studied in cryptography, there are only few results concerning quantum algorithms for affine equivalence testing. Proposition 1 establishes the relation between the affine equivalence and the isomorphism,

and this may contribute to the further study of the testing of the affine equivalence. Identity can be seen as a special case of isomorphism and affine equivalence. Motivated by the isomorphism testing, we consider the identity testing problem. In this section, we give a quantum algorithm for testing Boolean functions identity. Moreover, our algorithm will be shown to be optimal. Then we present a lower bound on classical algorithms for the same problem. We also show that such lower bound is reachable.

### 3.1. Quantum lower and upper bounds for testing Boolean functions identity

**Problem 1** The Identity Testing Problem: Given two unknown Boolean functions  $f, g$ , and an  $\varepsilon > 0$ , the testing problem is to determine whether  $f$  is identical with  $g$  or is  $\varepsilon$ -far from  $g$ , under the promise that one of them holds.

In the black box model, one provides a “black-box access” to a function  $f$ , meaning that the quantum algorithm can apply some unitary transformation  $U_f : |x\rangle|y\rangle \rightarrow |x\rangle|y \oplus f(x)\rangle$  (or  $|x\rangle$  to  $(-1)^{f(x)}|x\rangle$ , where  $f$  is a Boolean function). We present now a quantum algorithm for the above testing problem: Algorithm 1.

---

#### Algorithm 1: Quantum Algorithm for Identity Testing

---

**Input:** Black-boxes for  $f$  and  $g$

**Output:**  $f$  is identical with  $g$  iff  $z = 0$

$|\psi_0\rangle \leftarrow |0\rangle^{\otimes n} \otimes |1\rangle$ ;

Apply  $I^{\otimes n} \otimes H$  to  $|\psi_0\rangle$  and get  $|\psi_1\rangle = |0\rangle^{\otimes n} \otimes |-\rangle$ ;

Apply  $H^{\otimes n} \otimes I$  to  $|\psi_1\rangle$  and get  $|\psi_2\rangle = \sum_{x \in \{0,1\}^n} \frac{1}{\sqrt{2^n}} |x\rangle |-\rangle$ ;

Apply  $U_f$  to  $|\psi_2\rangle$  and get  $|\psi_3\rangle = \sum_{x \in \{0,1\}^n} \frac{(-1)^{f(x)}}{\sqrt{2^n}} |x\rangle |-\rangle$ ;

Apply  $U_g$  to  $|\psi_3\rangle$  and get  $|\psi_4\rangle = \sum_{x \in \{0,1\}^n} \frac{(-1)^{f(x)+g(x)}}{\sqrt{2^n}} |x\rangle |-\rangle$ ;

Apply  $H^{\otimes n} \otimes I$  to  $|\psi_4\rangle$  and get  $|\psi_5\rangle = \sum_{z \in \{0,1\}^n} W_h(z) |z\rangle |-\rangle$ , where  $h = f + g$ ;

Apply  $Q^m$  to  $|\psi_5\rangle$  and get  $|\psi_{final}\rangle$ , where  $Q = -\mathcal{A}S_0\mathcal{A}^{-1}S_\chi$ ,  $m = O(\frac{1}{\sqrt{\varepsilon}})$ ,

$\mathcal{A}|0\rangle^{\otimes n}|-\rangle = \sum_{z \in \{0,1\}^n} W_h(z) |z\rangle |-\rangle$ ,  $\mathcal{A} = (H^{\otimes n} \otimes I)(U_g)(U_f)(H^{\otimes n} \otimes I)$  and

$|z\rangle |-\rangle (z \neq 0)$  are good states;

Measurement the first  $n$  qubits and get  $z$ ;

Return  $z$ ;

---

**Theorem 3.1.** Algorithm 1 solves Identity Testing Problem 1 with one-sided error  $\varepsilon > 0$  using  $O(\frac{1}{\sqrt{\varepsilon}})$  queries. If  $f$  is identical with  $g$ , Algorithm 1 outputs “ $f$  is identical with  $g$ ” with certainty. If  $f$  is  $\varepsilon$ -far from  $g$ , Algorithm 1 outputs “ $f$  is  $\varepsilon$ -far from  $g$ ” with the probability at least  $\frac{3}{4}$ .



**Proof:**

According to Definition 2.4, we obtain

$$W_h(0) = \sum_{x \in \{0,1\}^n} \frac{(-1)^{f(x)+g(x)}}{2^n} \quad (9)$$

$$= \frac{1}{2^n} \sum_{x:f(x)=g(x)} (-1)^{f(x)+g(x)} + \frac{1}{2^n} \sum_{x:f(x) \neq g(x)} (-1)^{f(x)+g(x)} \quad (10)$$

$$= \frac{1}{2^n} [2^n - 2 \cdot \text{Ham}(f, g)] \quad (11)$$

$$= 1 - 2 \cdot \text{Dist}(f, g). \quad (12)$$

Without loss of generality, we assume  $\varepsilon \leq \text{Dist}(f, g) \leq 1/15$ . Otherwise, Problem 1 is easy to deal with classically.

Case 1:  $f$  is identical with  $g$ . Then the final state  $|\psi_{final}\rangle$  in Algorithm 1 is  $|0\rangle^{\otimes n} \otimes |-\rangle$ . Therefore, Algorithm 1 outputs “ $f$  is identical with  $g$ ” with a certainty.

Case 2:  $f$  is  $\varepsilon$ -far from  $g$ ,  $0 < W_h(0) = 1 - 2 \cdot \text{Dist}(f, g) \leq 1 - 2\varepsilon$ . Then the final state  $|\psi_{final}\rangle$  in Algorithm 1 is:

$$\sum_{z \in \{0,1\}^n} W_h(z) |z\rangle |-\rangle = |\phi_1\rangle |-\rangle + |\phi_0\rangle |-\rangle, \text{ where} \quad (13)$$

$$|\phi_1\rangle = \sum_{z \neq 0} W_h(z) |z\rangle, \quad (14)$$

$$|\phi_0\rangle = \sum_{z=0} W_h(z) |z\rangle, \quad (15)$$

and therefore

$$a = \sin^2 \theta_a = \langle \phi_1 | \phi_1 \rangle \quad (16)$$

$$= \sum_{z \neq 0} W_h^2(z) = 1 - \sum_{z=0} W_h^2(z) \quad (17)$$

$$\geq 1 - (1 - 2\varepsilon)^2 = 4\varepsilon - 4\varepsilon^2 \geq \frac{56}{15}\varepsilon. \quad (18)$$

Therefore, we obtain  $\theta_a \geq \arcsin \sqrt{\frac{56}{15}\varepsilon}$ . To prepare the final state  $|\psi_{final}\rangle$  in Algorithm 1, we need to call black box two times, i.e., that is to use two queries. So it takes 4 queries (i.e.,  $\mathcal{A}$  uses two queries,  $\mathcal{A}^{-1}$  uses two queries too) when Algorithm 1 runs  $Q$  once time, where  $Q = -\mathcal{A}S_0\mathcal{A}^{-1}S_\chi$ . According to Lemma 2.16, Algorithm 1 runs  $Q$   $m$  times, where  $m = \lfloor \pi/4\theta_a \rfloor$ . Finally, we know that the total number of queries of Algorithm 1 is  $4m + 2 = O(\frac{1}{\sqrt{\varepsilon}})$ . Hence the query complexity of Algorithm 1 is  $O(\frac{1}{\sqrt{\varepsilon}})$ . According to the hypothesis  $\varepsilon \leq \text{Dist}(f, g) \leq 1/15$ , we have  $(\frac{13}{15})^2 \leq W_h^2(0) \leq (1 - 2\varepsilon)^2$ , i.e.,  $a \leq 56/225$ . Again, by Lemma 2.16, Algorithm 1 outputs “ $f$  is  $\varepsilon$ -far from  $g$ ” with the probability  $p \geq \max(1 - 56/225, 56/225) > \frac{3}{4}$ .  $\square$

**Remark 3.2.** For an unknown function  $f$ , we can test  $f$  using different types of  $g$  so as to obtain an information concerning  $f$ . Also we can test two unknown functions whether they are identical or they are  $\varepsilon$ -far from each other. Besides, in the following quantum algorithms  $|-\rangle$  also stands for  $\frac{|0\rangle - |1\rangle}{\sqrt{2}}$ .

As the next we show that the quantum lower bound for Problem 1 is also  $\Omega(\frac{1}{\sqrt{\varepsilon}})$ . That means that Algorithm 1 is optimal.

**Theorem 3.3.** Any quantum query algorithm for Identity Testing Problem 1 requires  $\Omega(\frac{1}{\sqrt{\varepsilon}})$  queries.

**Proof:**

Let  $X = (x_0, x_1, \dots, x_{N-1})$ ,  $Y = (y_0, y_1, \dots, y_{N-1})$ . Let  $f(i) = x_i$ ,  $g(i) = y_i$ , for  $i = 0, 1, \dots, N-1$ ,  $x_i, y_i \in \{0, 1\}$  and  $N = 2^n$ . Values of any  $n$ -bit Boolean function can now be expressed by an  $N$ -bit string. The task to solve Problem 1 can now be reduced to that of computing the Hamming weight of an  $N$ -bit string  $Z = (x_0 \oplus y_0, x_1 \oplus y_1, \dots, x_{N-1} \oplus y_{N-1})$ .

In order to do that, we consider any  $N$ -bit symmetric partial Boolean function  $H : \{D \subset \{0, 1\}^N\} \rightarrow \{0, 1\}$  defined as follows:

$$H(Z) = \begin{cases} 0 & \text{if } |Z| = 0, \\ 1 & \text{if } |Z| \geq \varepsilon N, \\ \text{undefined} & \text{otherwise.} \end{cases} \quad (19)$$

Solving Problem 1 is now equivalent to computing symmetric partial Boolean function  $H(Z)$ . By Lemma 2.12, there exists an univariate polynomial  $\phi(t)$  approximating  $H$  with an error  $1/3$  and  $\widetilde{\deg}(H) = \deg(\phi)$ . According to the Lagrange's mean value theorem on the interval  $[0, \varepsilon N]$ , we have

$$\phi(\varepsilon N) - \phi(0) = \phi'(\xi)(\varepsilon N), 0 < \xi < \varepsilon N. \quad (20)$$

Since the polynomial  $\phi(t)$  approximates  $H$  with the error  $1/3$ , we have

$$0 \leq \phi(0) \leq 1/3, \quad 2/3 \leq \phi(\varepsilon N) \leq 1, \quad (21)$$

$$\phi(\varepsilon N) - \phi(0) \geq 1/3 > 0. \quad (22)$$

Therefore,  $\phi'(\xi) = c_0/\varepsilon N$ , where  $c_0 = \phi(\varepsilon N) - \phi(0)$ . In other words,  $\phi(t)$  has the following properties:

1.  $0 \leq \phi(i) \leq 1$  for any integer  $0 \leq i \leq N$ .
2. There exists  $\xi$  such that  $|\phi'(\xi)| \geq c_0/\varepsilon N$ .

By Lemma 2.14, we obtain

$$\widetilde{\deg}(H) = \deg(\phi) \quad (23)$$

$$\geq \sqrt{(c_0 N / \varepsilon N) / (c_0 / \varepsilon N + 1 - 0)} \quad (24)$$

$$= \sqrt{c_0 / (c_0 / N + \varepsilon)} \quad (25)$$

$$= \Omega\left(\frac{1}{\sqrt{\varepsilon}}\right). \quad (26)$$

Finally, by Lemma 2.15, we get  $\widetilde{Q}(H) = \Omega(\frac{1}{\sqrt{\varepsilon}})$ . □

**Remark 3.4.** According to Theorem 3.1 and Theorem 3.3, the quantum query complexity of Problem 1 is  $\Theta(\frac{1}{\sqrt{\varepsilon}})$ .

### 3.2. Classical lower and upper bounds for testing Boolean functions identity

It is now natural to ask whether our quantum algorithm has any advantage comparing to its classical counterparts. In order to do that, we present in this subsection a lower bound for randomized classical algorithms for testing Boolean functions identity.

In order to prove such a lower bound, we will use the following definition.

**Definition 3.5.** We say that inputs  $X_{i_1}, X_{i_2}, \dots, X_{i_k}$  are good for some functions  $f$  and  $g$  if there exists  $x_{i_j}$  such that  $f(X_{i_j}) \neq g(X_{i_j})$ , where  $X_{i_l} \in \{0, 1\}^n$  for  $l = 1, 2, \dots, k$ . Otherwise, the inputs  $X_{i_1}, X_{i_2}, \dots, X_{i_k}$  will be said to be bad. We will denote the event

$$B_k = \{X_{i_1}, X_{i_2}, \dots, X_{i_k} \mid \text{inputs } X_{i_j} \text{ are bad, } 1 \leq j \leq k\}.$$

**Theorem 3.6.** Any classical randomized algorithm for Identity Testing Problem 1 requires  $\Omega(\frac{1}{\varepsilon})$  queries.

**Proof:**

Let  $\mathcal{A}$  be a randomized algorithm for Problem 1 and  $N = 2^n$ . Let us also suppose that  $\mathcal{A}$  makes some random adaptive queries for  $m$  different inputs  $X_{i_1}, X_{i_2}, \dots, X_{i_m}$ . If  $f$  is  $\varepsilon$ -far from  $g$ , according to the conditional probability formula and the fact that  $B_m \subseteq B_{m-1} \subseteq \dots \subseteq B_1$ , we get that

$$P(B_m) = P(B_m | B_{m-1})P(B_{m-1}) \tag{27}$$

$$= P(B_m | B_{m-1})P(B_{m-1} | B_{m-2})P(B_{m-2}) \tag{28}$$

$$= P(B_1) \prod_{k=2}^m P(B_k | B_{k-1}) \tag{29}$$

$$= (1 - \varepsilon) \frac{(1 - \varepsilon)N - 1}{N - 1} \frac{(1 - \varepsilon)N - 2}{N - 2} \dots \frac{(1 - \varepsilon)N - (m - 1)}{N - (m - 1)} \tag{30}$$

$$= (1 - \varepsilon) \left[ 1 - \frac{\varepsilon N}{N - 1} \right] \left[ 1 - \frac{\varepsilon N}{N - 2} \right] \dots \left[ 1 - \frac{\varepsilon N}{N - (m - 1)} \right] \tag{31}$$

$$\geq (1 - \varepsilon) \left[ 1 - \sum_{k=1}^{m-1} \frac{\varepsilon N}{N - k} \right] \tag{32}$$

$$\geq (1 - \varepsilon) \left[ 1 - \frac{\varepsilon(m - 1)N}{N - (m - 1)} \right] \tag{33}$$

$$\geq (1 - \varepsilon) \left[ 1 - \frac{\varepsilon m N}{N - (m - 1)} \right]. \tag{34}$$

If  $m = o(\frac{1}{\varepsilon})$  and  $\varepsilon$  is sufficiently small, then  $(1 - \varepsilon) \left[ 1 - \frac{\varepsilon m N}{N - (m - 1)} \right] \approx 1$ . That is  $P(B_m) \approx 1$ . Therefore, the algorithm cannot determine whether  $f$  is identical with  $g$  or is  $\varepsilon$ -far from  $g$  with a high success probability.  $\square$

The above lower bound is tight. Indeed, we give now an algorithm to reach the lower bound.

---

**Algorithm 2:** Classical Randomized Algorithm for Identity Testing

---

**Input:** Black-boxes for  $f$  and  $g$ , parameter  $\varepsilon > 0$   
**Output:**  $f$  is  $\varepsilon$ -far from  $g$  iff  $f(x) \neq g(x)$  for some  $x$   
 $r \leftarrow 0$ ;  
**while**  $r \leq \frac{\ln 3}{\varepsilon}$  **do**  
    Take an element  $x \in \{0, 1\}^n$  uniformly at random;  
    **if**  $f(x) \neq g(x)$  **then**  
        | outputs “ $f$  is  $\varepsilon$ -far from  $g$ ” and halts;  
    **end**  
     $r \leftarrow r + 1$ ;  
**end**

---

**Theorem 3.7.** Algorithm 2 solves Problem 1 with one-sided error using  $O(\frac{1}{\varepsilon})$  queries. If  $f$  is identical with  $g$ , our Algorithm 2 outputs “ $f$  is identical with  $g$ ” with certainty. If  $f$  is  $\varepsilon$ -far from  $g$ , Algorithm 2 outputs “ $f$  is  $\varepsilon$ -far from  $g$ ” with the probability at least  $\frac{2}{3}$ .

**Proof:**

Case 1: Let  $f$  be identical with  $g$ . Algorithm 2 then clearly outputs “ $f$  is identical with  $g$ ” with certainty.

Case 2: Let  $f$  be  $\varepsilon$ -far from  $g$ . If Algorithm 2 runs  $r$  loops, Algorithm 2 reports  $f$  is identical with  $g$  with an error probability  $p_r \leq (1 - \varepsilon)^r < \frac{1}{3}$ . The above inequality then holds when we take  $r > \frac{\ln 3}{\varepsilon}$ , i.e.,  $r = O(\frac{1}{\varepsilon})$ .  $\square$

**Remark 3.8.** According to Theorem 3.6 and Theorem 3.7, the randomized query complexity for our Problem 1 is  $\Theta(\frac{1}{\varepsilon})$ . Observe also that Batu *et al.* [24] and Diakonikolas *et al.* [25] studied the question of the identity testing for all kinds of distributions. The query complexity of these algorithms is not only related to the metric distance parameter  $\varepsilon$ , but also to the distribution size. The query complexity of our algorithm depends on the value of  $\varepsilon$  and not on input size of Boolean functions. The reasons for this need to be further discussed.

## 4. Testing correlations between two unknown Boolean functions

Various properties and correlations of some and between symmetric Boolean functions have been already explored extensively including some asymptotic properties. For example, the Deutsch-Jozsa quantum algorithm solved Deutsch’s problem using one query: it is the problem (as already disussed) to decide whether a given function  $f$  is constant or balanced under the promise that it has one of these two properties. This problem is equivalent to the problem to distinguish  $|x\rangle = 2^{n-1}$  from  $|y\rangle = 0, 2^n$ , where  $|y\rangle$  denotes the Hamming weight of string  $y$ . Gruska, Qiu and Zheng [37] generalized the distributed Deutsch-Jozsa promise problem. Recently, Qiu and Zheng [32] also generalized the D-J problem and gave an exact quantum query algorithm to distinguish such strings. Now, we consider

the D-J problem from the perspective of functional correlation. In fact, it is a problem equivalent to distinguishing  $|C(f, \mathbf{0})| = 0$  from  $|C(f, \mathbf{0})| = 1$ , where  $\mathbf{0}$  denotes that  $f(x)$  is 0 for all values of  $x$ .<sup>1</sup> Based on the above facts, we generalize the idea of Deutsch-Jozsa quantum algorithm to distinguish the correlation  $|C(f, g)| = \varepsilon$  from  $|C(f, g)| = 1$  between two unknown functions.

### 4.1. Exact quantum query algorithms for testing correlation

**Problem 2** Correlation Testing Problem: Given two Boolean functions  $f, g$  as black boxes, the testing problem is to determine whether  $|C(f, g)| = \varepsilon$  or  $|C(f, g)| = 1$ , where  $0 \leq \varepsilon \leq \frac{\sqrt{3}}{2}$ , under the promise that one of them holds.

---

**Algorithm 3:** Quantum Algorithm for Correlation Testing

---

**Input:** Black-boxes for  $f$  and  $g$

**Output:**  $|C(f, g)| = 1$  iff  $z = 0$

$|\psi_0\rangle \leftarrow |0\rangle^{\otimes n} \otimes |1\rangle$ ;

Apply  $I^{\otimes n} \otimes H$  to  $|\psi_0\rangle$  and get  $|\psi_1\rangle = |0\rangle^{\otimes n} \otimes |-\rangle$ ;

Apply  $H^{\otimes n} \otimes I$  to  $|\psi_1\rangle$  and get  $|\psi_2\rangle = \sum_{x \in \{0,1\}^n} \frac{1}{\sqrt{2^n}} |x\rangle |-\rangle$ ;

Apply  $U_f$  to  $|\psi_2\rangle$  and get  $|\psi_3\rangle = \sum_{x \in \{0,1\}^n} \frac{(-1)^{f(x)}}{\sqrt{2^n}} |x\rangle |-\rangle$ ;

Apply  $U_g$  to  $|\psi_3\rangle$  and get  $|\psi_4\rangle = \sum_{x \in \{0,1\}^n} \frac{(-1)^{f(x)+g(x)}}{\sqrt{2^n}} |x\rangle |-\rangle$ ;

Apply  $H^{\otimes n} \otimes I$  to  $|\psi_4\rangle$  and get  $|\psi_5\rangle = \sum_{z \in \{0,1\}^n} W_h(z) |z\rangle |-\rangle = |\Psi\rangle |-\rangle = |\Psi_1\rangle |-\rangle + |\Psi_0\rangle |-\rangle$ ,

where  $|\Psi_1\rangle = \sum_{z \neq 0} W_h(z) |z\rangle$ ,  $|\Psi_0\rangle = \sum_{z=0} W_h(z) |z\rangle$  and  $h = f + g$ ;

Apply  $Q$  to  $|\psi_5\rangle$  and get  $|\psi_{final}\rangle$ , where  $Q = -\mathcal{A}S_0(\phi)\mathcal{A}^{-1}S_z(\varphi)$ ,  $\mathcal{A}|0\rangle^{\otimes n}|-\rangle = |\Psi\rangle |-\rangle$ ,

$\phi = 2 \arcsin \frac{1}{2\sqrt{1-\varepsilon^2}}$ ,  $\varphi = 2 \arcsin \frac{1}{2\sqrt{1-\varepsilon^2}}$ ,  $\mathcal{A} = (H^{\otimes n} \otimes I)(U_g)(U_f)(H^{\otimes n} \otimes I)$  and  $S_z(\varphi)$  is defined such that

$$|z\rangle |-\rangle \rightarrow \begin{cases} e^{i\varphi} |z\rangle |-\rangle & \text{if } z \neq 0, \\ |z\rangle |-\rangle & \text{if } z = 0. \end{cases}$$

Measurement the first  $n$  qubits and get  $z$ ;

Return  $z$ ;

---

**Theorem 4.1.** Algorithm 3 solves Problem 2 with certainty using 6 queries only.

**Proof:**

Case 1:  $|C(f, g)| = 1$ . We obtain  $\mathcal{A}|0\rangle^{\otimes n}|-\rangle = |\Psi\rangle |-\rangle = |\Psi_0\rangle |-\rangle$ . Hence

$$\begin{aligned} Q|\Psi\rangle |-\rangle &= Q|\Psi_0\rangle |-\rangle \\ &= -\mathcal{A}S_0(\phi)\mathcal{A}^{-1}S_z(\varphi)|\Psi_0\rangle |-\rangle \\ &= -\mathcal{A}(I - (1 - e^{i\phi})|0\rangle^{\otimes n}\langle 0|^{\otimes n} \otimes |-\rangle\langle -|)\mathcal{A}^{-1}I|\Psi_0\rangle |-\rangle \\ &= -|\Psi_0\rangle |-\rangle + (1 - e^{i\phi})|\Psi_0\rangle |-\rangle \\ &= -e^{i\phi}|\Psi_0\rangle |-\rangle. \end{aligned} \tag{35}$$

---

<sup>1</sup>For the sake of beauty, here we abuse the absolute value notation.

Therefore, the result of measurement is  $z = 0$ , and Algorithm 3 outputs “ $|C(f)| = 1$ ” with certainty.

Case 2a:  $|C(f, g)| = \varepsilon = 0$ . We obtain  $\mathcal{A}|0\rangle^{\otimes n}|-\rangle = |\Psi\rangle|-\rangle = |\Psi_1\rangle|-\rangle$ . Hence

$$\begin{aligned} Q|\Psi\rangle|-\rangle &= Q|\Psi_1\rangle|-\rangle \\ &= -\mathcal{A}S_0(\phi)\mathcal{A}^{-1}S_z(\varphi)|\Psi_1\rangle|-\rangle \\ &= -\mathcal{A}(I - (1 - e^{i\phi})|0\rangle^{\otimes n}\langle 0|^{\otimes n} \otimes |-\rangle\langle -|)\mathcal{A}^{-1}e^{i\varphi}|\Psi_1\rangle|-\rangle \\ &= e^{i\varphi}(-|\Psi_1\rangle + (1 - e^{i\phi})|\Psi_1\rangle)|-\rangle \\ &= -e^{i(\varphi+\phi)}|\Psi_1\rangle|-\rangle. \end{aligned} \quad (36)$$

Therefore, the result of measurement is  $z \neq 0$ , and Algorithm 3 outputs “ $|C(f, g)| = \varepsilon$ ” with certainty.

Case 2b:  $|C(f, g)| = \varepsilon \in (0, \frac{\sqrt{3}}{2})$ . This is equivalent to  $|W_h(0)| = \varepsilon$ , so  $\sin^2(\theta_a) = a = \langle \Psi_1|\Psi_1\rangle = 1 - \varepsilon^2$ . Finally, we have  $0 < \pi/4\theta_a - \frac{1}{2} < 1$  and  $m = \lfloor \pi/4\theta_a - \frac{1}{2} \rfloor = 0$ . Therefore, we only need to compute the parameters value  $\varphi$  and  $\phi$  using  $Q$  once in order to achieve a precise distinction.

Let us now choose  $\varphi$  and  $\phi$  such that

$$e^{i\varphi}(1 - e^{i\phi})\sqrt{a}\sin(\theta_a) = ((1 - e^{i\phi})a + e^{i\phi})\frac{1}{\sqrt{1-a}}\cos(\theta_a). \quad (37)$$

This equation is equivalent to

$$(1 - \varepsilon^2)(1 - e^{i\phi})(e^{i\varphi} - 1) = e^{i\phi}. \quad (38)$$

and by solving this equation we obtain  $\phi = \varphi = 2 \arcsin \frac{1}{2\sqrt{1-\varepsilon^2}}$ . By Lemma 2.17, we get

$$|\Psi_{final}\rangle = Q|\Psi\rangle|-\rangle = \left[ a(e^{i\varphi} - 1)(1 - e^{i\phi}) - (e^{i\phi} + e^{i(\varphi)}) + 1 \right] |\Psi_1\rangle|-\rangle. \quad (39)$$

Therefore, the result of the measurement provides  $z \neq 0$ , i.e., Algorithm 3 outputs “ $|C(f, g)| = \varepsilon$ ” with certainty.

Case 2c:  $|C(f, g)| = \varepsilon = \frac{\sqrt{3}}{2}$ . This is equivalent to  $|W_h(0)| = \frac{\sqrt{3}}{2}$ , so  $\sin^2(\theta_a) = a = \langle \Psi_1|\Psi_1\rangle = \frac{1}{4}$ . We have  $\phi = \varphi = 2 \arcsin \frac{1}{2\sqrt{1-\varepsilon^2}} = \pi$ . By Lemma 2.17, the amplitude of state  $|\Psi_0\rangle|-\rangle$  is

$$e^{i\varphi}(1 - e^{i\phi})a - ((1 - e^{i\phi})a + e^{i\phi}) \quad (40)$$

$$= -2a(1 - e^{i\phi}) - e^{i\phi} = 0. \quad (41)$$

That is, equation 39 holds. Therefore, the result of the measurement provides  $z \neq 0$ , i.e., Algorithm 3 outputs “ $|C(f, g)| = \varepsilon$ ” with certainty.  $\square$

**Theorem 4.2.** For correlation testing Problem 2, the classical deterministic query complexity is  $\Theta(N)$ . More specifically, solving Problem 2 with zero error requires  $(1 - \varepsilon)N + 2$  or  $(1 + \varepsilon)N + 2$  queries.

**Proof:**

If the first  $(\frac{1-\varepsilon}{2})N$  input queries return  $f(x_i) \neq g(x_i)$  or  $(\frac{1+\varepsilon}{2})N$  input queries return  $f(x_i) = g(x_i)$ , then we will need to make another query as well. So the total number of queries is  $(1 - \varepsilon)N + 2$  or  $(1 + \varepsilon)N + 2$ . Therefore the theorem holds.  $\square$

## 5. Testing balancedness of Boolean functions

Balancedness is also one of the properties of Boolean functions important in cryptography, see for example [20, 30]. Of interest are also functions that are not balanced, but  $\varepsilon$ -far balanced. To this end, we introduce the following concept. We study the balance problem from the perspective of function property testing. In the following we study balancedness and  $\varepsilon$ -far balancedness of Boolean functions.

**Definition 5.1.** A Boolean function  $f(x)$  is called  $\varepsilon$ -far balanced if  $|C(f)| \geq \varepsilon$ .

**Problem 3** Balancedness Testing Problem: Given an unknown Boolean functions  $f$ , balancedness testing problem is to determine whether  $f$  is balanced or is  $\varepsilon$ -far balanced, under the promise that one of them holds.

---

### Algorithm 4: Quantum Algorithm for Balancedness Testing

---

**Input:** Black-boxes for  $f$  and parameter  $\varepsilon > 0$

**Output:**  $f$  is balanced iff  $z \neq 0$

$|\psi_0\rangle \leftarrow |0\rangle^{\otimes n} \otimes |1\rangle$ ;

Apply  $I^{\otimes n} \otimes H$  to  $|\psi_0\rangle$  and get  $|\psi_1\rangle = |0\rangle^{\otimes n} \otimes |-\rangle$ ;

Apply  $H^{\otimes n} \otimes I$  to  $|\psi_1\rangle$  and get  $|\psi_2\rangle = \sum_{x \in \{0,1\}^n} \frac{1}{\sqrt{2^n}} |x\rangle |-\rangle$ ;

Apply  $U_f$  to  $|\psi_2\rangle$  and get  $|\psi_3\rangle = \sum_{x \in \{0,1\}^n} \frac{(-1)^{f(x)}}{\sqrt{2^n}} |x\rangle |-\rangle$ ;

Apply  $H^{\otimes n} \otimes I$  to  $|\psi_3\rangle$  and get  $|\psi_4\rangle = \sum_{z \in \{0,1\}^n} W_f(z) |z\rangle |-\rangle = |\Psi\rangle |-\rangle = |\Psi_1\rangle |-\rangle + |\Psi_0\rangle |-\rangle$ ,  
where  $|\Psi_1\rangle = \sum_{z=0} W_f(z) |z\rangle$ ,  $|\Psi_0\rangle = \sum_{z \neq 0} W_f(z) |z\rangle$ ;

Apply  $Q^m$  to  $|\psi_4\rangle$  and get  $|\psi_{final}\rangle$ , where  $Q = -\mathcal{A}S_0\mathcal{A}^{-1}S_\chi$ ,  $m = O(\frac{1}{\varepsilon})$ ,  $\mathcal{A}|0\rangle^{\otimes n}|-\rangle = |\Psi\rangle |-\rangle$ ,

$\mathcal{A} = (H^{\otimes n} \otimes I)(U_f)(H^{\otimes n} \otimes I)$  and  $|0\rangle |-\rangle$  is the only good state;

Measurement the first  $n$  qubits and get  $z$ ;

Return  $z$ ;

---

**Theorem 5.2.** Algorithm 4 solves Balancedness Testing Problem 3 with one-sided error using  $O(\frac{1}{\varepsilon})$  queries. If  $f$  is balanced, Algorithm 4 outputs “ $f$  is balanced” with certainty. If  $f$  is  $\varepsilon$ -far balanced, Algorithm 4 outputs that “ $f$  is  $\varepsilon$ -far balanced” with probability at least  $\frac{3}{4}$ .

### Proof:

Case 1:  $f$  is balanced, i.e.,  $|C(f)| = 0$ . We obtain  $\mathcal{A}|0\rangle^{\otimes n}|-\rangle = |\Psi\rangle |-\rangle = |\Psi_0\rangle |-\rangle$ . Hence

$$\begin{aligned}
 Q|\Psi\rangle |-\rangle &= Q|\Psi_0\rangle |-\rangle \\
 &= -\mathcal{A}S_0\mathcal{A}^{-1}S_\chi|\Psi_0\rangle |-\rangle \\
 &= -\mathcal{A}(I - 2|0\rangle^{\otimes n}\langle 0|^{\otimes n} \otimes |-\rangle\langle -|)\mathcal{A}^{-1}I|\Psi_0\rangle |-\rangle \\
 &= -|\Psi_0\rangle |-\rangle + 2|\Psi_0\rangle |-\rangle \\
 &= |\Psi_0\rangle |-\rangle = |\Psi\rangle |-\rangle.
 \end{aligned} \tag{42}$$

Therefore, Algorithm 4 outputs “ $f$  is balanced” with certainty.

Case 2:  $f$  is  $\varepsilon$ -far balanced, i.e.,  $|C(f)| \geq \varepsilon$ ,  $|W_f(0)| = |C(f)| \geq \varepsilon$ . In this case the final state  $|\psi_{final}\rangle$  in Algorithm 4 is:

$$\sum_{z \in \{0,1\}^n} W_f(z)|z\rangle|-\rangle = |\Psi_1\rangle|-\rangle + |\Psi_0\rangle|-\rangle, \quad (43)$$

$$|\Psi_1\rangle = \sum_{z=0} W_f(z)|z\rangle, \quad (44)$$

$$|\Psi_0\rangle = \sum_{z \neq 0} W_f(z)|z\rangle, \quad (45)$$

and therefore

$$\begin{aligned} a &= \sin^2 \theta_a = \langle \Psi_1 | \Psi_1 \rangle \\ &= C^2(f) \geq \varepsilon^2. \end{aligned} \quad (46)$$

Hence,  $\theta_a \geq \arcsin \varepsilon$ .

In order to prepare the final state  $|\psi_{final}\rangle$  in Algorithm 4, we need to call the function black-box two times, i.e., to use two queries. So it takes in total 4 queries (indeed,  $\mathcal{A}$  uses two queries,  $\mathcal{A}^{-1}$  uses two queries also) when Algorithm 4 runs  $Q$  once, where  $Q = -\mathcal{A}S_0\mathcal{A}^{-1}S_\chi$ . According to Lemma 2.16, Algorithm 4 runs  $Q$   $m = \lfloor \pi/4\theta_a \rfloor$  times. Since the total number of queries of Algorithm 4 is  $2m + 1 = O(\frac{1}{\varepsilon})$ , i.e., the query complexity of Algorithm 4 is  $O(\frac{1}{\varepsilon})$ . Without loss of generality, let's now assume  $\varepsilon \leq 1/2$ . In such a case we get, using Lemma 2.16, Algorithm 4 outputs “ $f$  is  $\varepsilon$ -far balanced” with the probability  $p \geq \max(1 - 1/4, 1/4) > \frac{3}{4}$ .  $\square$

Finally, we prove that our quantum algorithm for Problem 4 is optimal.

For  $X = (x_1, \dots, x_n) \in \{0, 1\}^n$ . Let  $l, l'$  be integers such that  $0 \leq l, l' \leq n$ . Let us now define the partial Boolean function  $f_{l,l'}$  on  $\{0, 1\}^n$ , as

$$f_{l,l'}(X) = \begin{cases} 1 & \text{if } |X| = l, \\ 0 & \text{if } |X| = l' \\ \text{undefined} & \text{otherwise.} \end{cases} \quad (47)$$

In addition, let  $m \in \{l, l'\}$  be such that  $|\frac{n}{2} - m|$  is maximized, and let  $\Delta_l = |l - l'|$ .

**Lemma 5.3.** ([38]) A lower bound for quantum query complexity of computing the partial function  $f_{l,l'}$ , given the input as an oracle, is  $\Omega\left(\sqrt{n/\Delta_l} + \sqrt{m(n-m)/\Delta_l}\right)$ .

**Theorem 5.4.** Any quantum query algorithm for Balancedness Testing Problem 3 requires  $\Omega(\frac{1}{\varepsilon})$  queries.

**Proof:**

Let  $Y = (y_0, y_1, \dots, y_{N-1})$  where  $N = 2^n$ , for  $i = 0, 1, \dots, N-1$ ,  $y_i \in \{0, 1\}$  and  $f(i) = y_i$ . The value of any  $n$ -bit Boolean function can be expressed by an  $N$ -bit string  $Y$ . The task to solve the Balancedness problem can now be reduced to the problem to distinguish the Hamming weight  $|Y| = \frac{N}{2}$  from  $|Y| \leq \frac{(1-\varepsilon)N}{2}$  or  $|Y| \geq \frac{(1+\varepsilon)N}{2}$ . In view of this, let us consider two arbitrary  $N$ -bit



symmetric partial Boolean functions  $H_1, H_2 : \{D \subset \{0, 1\}^N\} \rightarrow \{0, 1\}$  (for a subset  $D$ ), which are defined as follows:

$$H_1(Y) = \begin{cases} 1 & \text{if } |Y| = \frac{N}{2}, \\ 0 & \text{if } |Y| = \frac{(1-\varepsilon)N}{2} \\ \text{undefined} & \text{otherwise.} \end{cases} \tag{48}$$

$$H_2(Y) = \begin{cases} 1 & \text{if } |Y| = \frac{N}{2}, \\ 0 & \text{if } |Y| = \frac{(1+\varepsilon)N}{2}, \\ \text{undefined} & \text{otherwise.} \end{cases} \tag{49}$$

We consider here only the case of calculating  $H_1(Y)$ . The other case is similar. Let  $l' = \frac{(1-\varepsilon)}{2}N$ ,  $l = \frac{N}{2}$ . We take  $m = \frac{(1-\varepsilon)}{2}N$  such that  $|\frac{N}{2} - m|$  is maximized, and let  $\Delta_l = |l - l'| = \frac{\varepsilon}{2}N$ . Therefore we have

$$\Omega \left( \sqrt{N/\Delta_l} + \sqrt{m(N-m)/\Delta_l} \right) \tag{50}$$

$$= \Omega \left( \sqrt{N/\frac{\varepsilon}{2}N} + \sqrt{\frac{1-\varepsilon}{2}N(N - \frac{1-\varepsilon}{2}N)/\frac{\varepsilon}{2}N} \right) \tag{51}$$

$$= \Omega \left( \sqrt{2/\varepsilon} + \sqrt{1 - \varepsilon^2/\varepsilon} \right) \tag{52}$$

$$= \Omega(1/\varepsilon). \tag{53}$$

According to Lemma 5.3, the query complexity of computing  $H_1(Y)$  (or  $H_2(Y)$ ) is  $\Omega(1/\varepsilon)$ . Therefore, the theorem holds.  $\square$

Therefore, the quantum query complexity of Balancedness Testing Problem 3 is  $\Theta(\frac{1}{\varepsilon})$ . Next, let's consider the classical algorithm for this problem. We first give the de Moivre-Laplace Central Limit Theorem as follows:

**Lemma 5.5.** ([39])(de Moivre-Laplace Central Limit Theorem). If  $X$  is a random variable having the binomial with the parameters  $n$  and  $p$ , the limiting form of the distribution function of the standardized random variable

$$Z = \frac{X - np}{\sqrt{np(1-p)}}$$

as  $n \rightarrow \infty$ , is given by the standard normal distribution

$$F(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt, -\infty < z < \infty.$$

Using the equation  $P(a < Z \leq b) = F(b) - F(a)$ , the de Moivre-Laplace Central Limit Theorem has the following equivalent forms

$$\lim_{n \rightarrow \infty} P[a \leq \frac{X - np}{\sqrt{np(1-p)}} \leq b] = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt. \tag{54}$$

---

**Algorithm 5:** Classical Randomized Algorithm for Balancedness Testing

---

**Input:** Black-boxes for  $f$  and parameter  $\varepsilon > 0$

**Output:**  $f$  is balanced iff  $|C'(f)| \leq 0.5\varepsilon$

$T \leftarrow O(1/\varepsilon^2)$ ;

Take  $T$  elements  $\{x_1, x_2, \dots, x_T\}$  uniformly at random, where  $x_i \in \{0, 1\}^n$ ;

Compute  $|C'(f)| = |\sum_{i=1}^T (-1)^{f(x_i)}|$ ;

Return  $|C'(f)|$ ;

---

**Theorem 5.6.** Algorithm 5 solves Balancedness Testing Problem 3 using  $O(1/\varepsilon^2)$  queries. If  $f$  is balanced, Algorithm 5 outputs “ $f$  is balanced” with the probability at least  $\frac{2}{3}$ ; If  $f$  is  $\varepsilon$ -far balanced, Algorithm 5 outputs “ $f$  is  $\varepsilon$ -far balanced” with the probability at least  $\frac{2}{3}$ .

**Proof:**

Let  $A_f = \{x : f(x) = 1\}$ . Let us now consider an experiment in which we will choose  $T$  times independently and randomly elements  $x_i \in \{0, 1\}^n$ . Let  $\xi$  be the number of times that those chosen elements occur in  $A_f$  in  $T$  trials, i.e.,  $\xi = |\{x_i : f(x_i) = 1\}|$ . Then  $\xi$  is a random variable. Moreover, we have  $\xi \sim b(T, c)$ , where  $P(A_f) = c$  and  $b(T, c)$  denotes binomial distribution. The distribution

law for  $\xi$  is as follows:  $P\{\xi = k\} = \binom{T}{k} c^k (1 - c)^{T-k}$ ,  $k = 0, 1, \dots, T$ . So the probability is

$$P\{(c - \delta)T \leq \xi \leq (c + \delta)T\} = \sum_{k=(c-\delta)T}^{(c+\delta)T} \binom{T}{k} c^k (1 - c)^{T-k}, \tag{55}$$

where  $\delta$  is an accuracy parameter.

It is difficult to find the exact value for the probability we need. We use The de Moivre-Laplace Central Limit Theorem to find its approximation. Let

$$P\{(c - \delta)T \leq \xi \leq (c + \delta)T\} \tag{56}$$

$$= P\left\{ \frac{(c - \delta)T - Tc}{\sqrt{Tc(1 - c)}} \leq \frac{\xi - Tc}{\sqrt{Tc(1 - c)}} \leq \frac{(c + \delta)T - Tc}{\sqrt{Tc(1 - c)}} \right\} \tag{57}$$

$$\approx \int_{\frac{(c-\delta)T - Tc}{\sqrt{Tc(1-c)}}}^{\frac{(c+\delta)T - Tc}{\sqrt{Tc(1-c)}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \text{ (Using Equation 54 approximation)} \tag{58}$$

$$= \int_{\frac{-\delta T}{\sqrt{Tc(1-c)}}}^{\frac{\delta T}{\sqrt{Tc(1-c)}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \tag{59}$$

$$= F\left(\frac{\delta T}{\sqrt{Tc(1 - c)}}\right) - F\left(\frac{-\delta T}{\sqrt{Tc(1 - c)}}\right) \tag{60}$$

$$= 2F\left(\frac{\delta T}{\sqrt{Tc(1 - c)}}\right) - 1 \geq \frac{2}{3}. \tag{61}$$

By checking the value of the distribution function, we have  $\frac{\delta T}{\sqrt{Tc(1-c)}} \geq 1$ , i.e.  $T = \Omega(1/\delta^2)$ .

Case 1:  $f$  is balanced, i.e.,  $|C(f)| = 0$ . That is  $c = 0.5$ . If the random variable  $\xi$  satisfies the inequality  $(c - \delta)T \leq \xi \leq (c + \delta)T$ , that is  $0.5 - \delta \leq \frac{\xi}{T} \leq 0.5 + \delta$ , we know that  $-2\delta \leq 1 - \frac{2\xi}{T} \leq 2\delta$ . Hence,

$$|C'(f)| = \frac{1}{T} \left| \sum_{x_i} (-1)^{f(x_i)} \right| \tag{62}$$

$$= \frac{1}{T} \left| \sum_{x_i: f(x_i)=1} (-1)^{f(x_i)} + \sum_{x_i: f(x_i)=0} (-1)^{f(x_i)} \right| \tag{63}$$

$$= \frac{1}{T} |-\xi + T - \xi| = \left| 1 - \frac{2\xi}{T} \right| \leq 2\delta. \tag{64}$$

If we now take  $\delta = 0.1\varepsilon$ , then  $P\{(0.5 - 0.1\varepsilon)T \leq \xi \leq (0.5 + 0.1\varepsilon)T\} \geq \frac{2}{3}$  using  $T = O(1/\delta^2) = O(1/\varepsilon^2)$  queries. In other words we get that  $|C'(f)| \leq 0.2\varepsilon$  with the probability at least  $\frac{2}{3}$ . Naturally,  $|C'(f)| \leq 0.5\varepsilon$  with the probability at least  $\frac{2}{3}$ . That is Algorithm 5 outputs  $|C(f)| = 0$  with the probability at least  $\frac{2}{3}$  using  $T = O(1/\delta^2) = O(1/\varepsilon^2)$  queries.

Case 2:  $f$  is  $\varepsilon$ -far balanced, i.e.,  $|C(f)| \geq \varepsilon$ . Let's consider the case of  $|C(f)| = \varepsilon$ . That is  $c = \frac{1+\varepsilon}{2}$ . In order to discuss it we need to consider two cases.

Case 2a:  $c = \frac{1+\varepsilon}{2}$ . If the random variable  $\xi$  satisfies the inequality  $(c - \delta)T \leq \xi \leq (c + \delta)T$ , that is  $\frac{1+\varepsilon}{2} - \delta \leq \frac{\xi}{T} \leq \frac{1+\varepsilon}{2} + \delta$ . Hence, we have  $-\varepsilon - 2\delta \leq 1 - \frac{2\xi}{T} \leq -\varepsilon + 2\delta$ . If we now take  $\delta = 0.1\varepsilon$ , then  $P\{(\frac{1+\varepsilon}{2} - 0.1\varepsilon)T \leq \xi \leq (\frac{1+\varepsilon}{2} + 0.1\varepsilon)T\} \geq \frac{2}{3}$  using  $T = O(1/\delta^2) = O(1/\varepsilon^2)$  queries. In other words, we can draw a conclusion that  $|C'(f)| = |1 - \frac{2\xi}{T}| \geq 0.8\varepsilon > 0.5\varepsilon$  with the probability at least  $\frac{2}{3}$  using  $O(1/\varepsilon^2)$  queries. That means that Algorithm 5 outputs  $|C(f)| \geq \varepsilon$  with probability at least  $\frac{2}{3}$  using  $O(1/\varepsilon^2)$  queries.

Case 2b:  $c = \frac{1-\varepsilon}{2}$ . If the random variable  $\xi$  satisfies the inequality  $(c - \delta)T \leq \xi \leq (c + \delta)T$ , that is  $\frac{1-\varepsilon}{2} - \delta \leq \frac{\xi}{T} \leq \frac{1-\varepsilon}{2} + \delta$ . Hence, we know that  $\varepsilon - 2\delta \leq 1 - \frac{2\xi}{T} \leq \varepsilon + 2\delta$ . If we now take  $\delta = 0.1\varepsilon$ , then  $P\{(\frac{1-\varepsilon}{2} - 0.1\varepsilon)T \leq \xi \leq (\frac{1-\varepsilon}{2} + 0.1\varepsilon)T\} \geq \frac{2}{3}$  using  $T = O(1/\delta^2) = O(1/\varepsilon^2)$  queries. In other words, we get that  $|C'(f)| = |1 - \frac{2\xi}{T}| \geq 0.8\varepsilon > 0.5\varepsilon$  with the probability at least  $\frac{2}{3}$  using  $O(1/\varepsilon^2)$  queries. That is Algorithm 5 outputs in any case  $|C(f)| \geq \varepsilon$  with the probability at least  $\frac{2}{3}$  using  $O(1/\varepsilon^2)$  queries.  $\square$

Next, we prove optimality of our Algorithm 5.

**Lemma 5.7.** ([40]) Let  $X_1, X_2, \dots, X_m$  be i.i.d random variables taking 0 or 1, and  $\Pr[X_i = 1] = p$ . If  $p \leq \frac{1}{4}$ , then for any  $t \geq 0$

$$\Pr \left[ \left( \sum_{i=1}^m X_i - \mu \right) > t \right] \geq \frac{1}{4} \exp\left(\frac{-2t^2}{\mu}\right). \tag{65}$$

If  $p \leq \frac{1}{2}$ , then for any  $0 \leq t \leq m(1 - 2p)$

$$\Pr \left[ \left( \sum_{i=1}^m X_i - \mu \right) > t \right] \geq \frac{1}{4} \exp\left(\frac{-2t^2}{\mu}\right), \quad (66)$$

where  $\mu = E[\sum_{i=1}^m X_i] = mp$ .

**Theorem 5.8.** Any classical randomized algorithm for Balancedness Testing Problem 3 requires  $\Omega(\frac{1}{\varepsilon^2})$  queries.

**Proof:**

Let  $Y = (y_1, y_2, \dots, y_N)$ . Consider a computation of the partial function  $H_1(Y)$  from the proof of Theorem 5.4. Select  $m$  strings  $y'_1, y'_2, \dots, y'_m$  in  $Y$  independently and randomly. Let  $p = \frac{1-\varepsilon}{2}$  and  $t = \frac{m\varepsilon}{2}$ . By Lemma 5.7, we obtain

$$\begin{aligned} & \Pr \left[ \left( \sum_{i=1}^m y'_i - \mu \right) > t \right] \\ &= \Pr \left[ \left( \sum_{i=1}^m y'_i - m \frac{1-\varepsilon}{2} \right) > \frac{m\varepsilon}{2} \right] \\ &= \Pr \left[ \sum_{i=1}^m y'_i > \frac{m}{2} \right] \\ &\geq \frac{1}{4} \exp\left(\frac{-2t^2}{\mu}\right) \\ &= \frac{1}{4} \exp\left(\frac{-m\varepsilon^2}{1-\varepsilon}\right). \end{aligned} \quad (67)$$

If  $m = o(\frac{1}{\varepsilon^2})$ , then  $\Pr \left[ \sum_{i=1}^m y'_i > \frac{m}{2} \right] \geq \frac{1}{4}$ . At this point, we cannot tell which case it is. Repeating a constant number of experiments, the error probability can reach  $\frac{1}{3}$ . Therefore,  $m = \Omega(\frac{1}{\varepsilon^2})$ .  $\square$

**Remark 5.9.** According to Theorem 5.6 and Theorem 5.8, the randomized query complexity of our Problem 3 is  $\Theta(\frac{1}{\varepsilon^2})$ .

## 6. Conclusions

In this paper, we have given classical and quantum algorithms for the problem of testing the identity of Boolean functions, which may be thought of as a special case of testing the isomorphism and the affine equivalence of Boolean functions. At the same time, we have also proved optimality of presented algorithms. By a combination of those results, we have obtained an optimal separation in the query complexities of the following identity testing problems:  $\Theta(\frac{1}{\varepsilon})$  versus  $\Theta(\frac{1}{\sqrt{\varepsilon}})$ . In addition, we have extend the idea of the D-J algorithm to the balancedness testing and correlation testing problems.

Our results naturally raise a variety of related open problems for future explorations. For example, what are properties of isomorphic Boolean functions and affine equivalent Boolean functions? In particular, what is the (optimal) quantum query complexity for affine equivalent Boolean functions? In more general cases, given two unknown Boolean functions  $f, g$ , the testing problem is to determine whether  $|C(f, g)| \leq a$  or  $|C(f, g)| \geq b$ , where  $0 < \varepsilon \leq b - a \leq 1$ , under the promise that one of these cases holds. Obviously, we have already solved some special cases ( $\varepsilon$  should be a relatively small number, see the following Table 1 for more details, where CQC means Classical query complexity and QQC means Quantum query complexity). We found that the query complexity of quantum or classic algorithms with an error depends not only on the distance between  $b$  and  $a$ , but also on the location of  $b$  and  $a$  and is independent of  $n$ . The question is, for example, when  $b - a = \varepsilon$ , the quantum query complexities of the Identity testing problem and the Balancedness testing problem are  $\Theta(\frac{1}{\sqrt{\varepsilon}})$  and  $\Theta(\frac{1}{\varepsilon})$  respectively. At the same time, when  $b - a = \varepsilon$ , the classical query complexities of the above two questions are  $\Theta(\frac{1}{\varepsilon})$  and  $\Theta(\frac{1}{\varepsilon^2})$  respectively. Besides, although the expression of quantum or classical query complexity function is different in above two problems, the separation between quantum and classical algorithms is same in these two special cases. Whether the same separation properties between quantum and classical algorithms can be generalized to general situations (i.e., for different  $a, b$  values) remains to be further investigated.

## Acknowledgements

The authors would like to thank the referees and Professor Calude for important comments that help us improve the quality of the manuscript. This work is supported in part by the National Natural Science Foundation of China (Nos. 61572532, 61876195), and the Natural Science Foundation of Guangdong Province of China (No. 2017B030311011).

Table 1. Query complexity results

Problem	Algorithm	Parameter values	QQC	CQC
Identity testing	Bounded Error	$a = 1 - 2\varepsilon, b = 1$	$\Theta(1/\sqrt{\varepsilon})$	$\Theta(1/\varepsilon)$
Deutsch-Jozsa Problem	Exact	$a = 0, b = 1$	$O(1)$	$\Theta(N)$
Correlation testing	Exact	$a = \varepsilon, b = 1$	$O(1)$	$\Theta(N)$
Balancedness testing	Bounded Error	$a = 0, b = \varepsilon$	$\Theta(1/\varepsilon)$	$\Theta(1/\varepsilon^2)$

## References

- [1] Fischer E, Kindler G, Ron D, Safra S, Samorodnitsky A. Testing juntas. *Journal of Computer and System Sciences*, 2004. **68**(4):753–787. doi:10.1016/j.jcss.2003.11.004.
- [2] Blais E, Weinstein A, Yoshida Y. Partially symmetric functions are efficiently isomorphism testable. *SIAM Journal on Computing*, 2015. **44**(2):411–432. doi:10.1137/140971877.

- [3] Alon N, Blais E, Chakraborty S, Garciasoriano D, Matsliah A. Nearly tight bounds for testing function isomorphism. *SIAM Journal on Computing*, 2013. **42**(2):459–493. doi:10.1137/110832677.
- [4] Chen X, Servedio RA, Tan L, Waingarten E, Xie J. Settling the query complexity of non-adaptive junta testing. *Journal of the ACM*, 2018. **65**(6):40. doi:10.1145/3213772.
- [5] Blais E, Canonne CL, Eden T, Levi A, Ron D. Tolerant junta testing and the connection to submodular optimization and function isomorphism. In: Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018. SIAM, 2018 pp. 2113–2132. doi:10.1137/1.9781611975031.138. URL <https://doi.org/10.1137/1.9781611975031.138>.
- [6] Blais E, Odonnell R. Lower bounds for testing function isomorphism. In: 2010 IEEE Conferences on Computational Complexity, Cambridge, Massachusetts, USA, 9-12 June, 2010. IEEE, 2010 pp. 235–246. doi:10.1109/CCC.2010.30.
- [7] Bellare M, Coppersmith D, Hastad J, Kiwi M, Sudan M. Linearity testing in characteristic two. *IEEE Transactions on Information Theory*, 1996. **42**(6):1781–1795. doi:10.1109/18.556674.
- [8] Chakrabarty D, Seshadhri C. An  $o(n)$  monotonicity tester for boolean functions over the hypercube. *SIAM Journal on Computing*, 2016. **45**(2):461–472. doi:10.1137/13092770X.
- [9] Chen X, Servedio RA, Tan L. New algorithms and lower bounds for monotonicity testing. In: 55th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2014, Philadelphia, PA, USA, October 18-21, 2014. IEEE Computer Society, 2014 pp. 286–295. doi:10.1109/FOCS.2014.38. URL <https://doi.org/10.1109/FOCS.2014.38>.
- [10] Friedl K, Santha M, Magniez F, Sen P. Quantum testers for hidden group properties. *Fundamenta Informaticae*, 2009. **91**(2):325–340. doi:10.3233/FI-2009-0046.
- [11] Gruska J. Quantum computing. McGraw Hill, 1999. ISBN 978-0077095031.
- [12] Qiu D, Zheng S. Revisiting Deutsch-Jozsa algorithm. *Information and Computation*, Received 4 March 2020, Accepted 13 July 2020, Available online 6 August 2020. doi:10.1016/j.ic.2020.104605. URL <https://doi.org/10.1016/j.ic.2020.104605>.
- [13] Grover LK. A Fast quantum mechanical algorithm for database search. In: Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing, Philadelphia, Pennsylvania, USA, May 22-24, 1996. ACM, 1996 pp. 212–219. doi:10.1145/237814.237866. URL <https://doi.org/10.1145/237814.237866>.
- [14] Shor PW. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 1997. **26**(5):1484–1509. doi:10.1137/S0097539795293172.
- [15] Buhrman H, Fortnow L, Newman I, Rohrig H. Quantum property testing. *SIAM Journal on Computing*, 2008. **37**(5):1387–1400. doi:10.1137/S0097539704442416.
- [16] Hillery M, Andersson E. Quantum tests for the linearity and permutation invariance of Boolean functions. *Physical Review A*, 2011. **84**(6):062329. doi:10.1103/PhysRevA.84.062329.
- [17] Bravyi S, Harrow AW, Hassidim A. Quantum algorithms for testing properties of distributions. *IEEE Transactions on Information Theory*, 2011. **57**(6):3971–3981. doi:10.1109/TIT.2011.2134250.
- [18] Montanaro A, De Wolf R. A Survey of quantum property testing. *Theory of Computing*, 2016. **7**(7):1–81. doi:10.4086/toc.gs.2016.007.

- [19] Hou X. Classification of cosets of the Reed-Muller code  $R(m-3, m)$ . *Discrete Mathematics*, 1994. **128**(1):203–224. doi:10.1016/0012-365X(94)90113-9.
- [20] Gao G, Guo Y, Zhao Y. Recent results on balanced symmetric Boolean functions. *IEEE Transactions on Information Theory*, 2016. **62**(9):5199–5203. doi:10.1109/TIT.2015.2455052.
- [21] Cusick T, Cheon Y. Affine equivalence for rotation symmetric Boolean functions in  $2^k$  variables. *Designs, Codes and Cryptography*, 2012. **63**(2):273–294. doi:10.1007/s10623-011-9553-6.
- [22] Canright DR, Chung JH, Stănică P. Circulant matrices and affine equivalence of monomial rotation symmetric Boolean functions. *Discrete Mathematics*, 2015. **338**(12):2197–2211. doi:10.1016/j.disc.2015.05.017.
- [23] Fuller JE. Analysis of affine equivalent boolean functions for cryptography. Ph.D. thesis, Queensland University of Technology, 2003. URL <https://eprints.qut.edu.au/15828/>.
- [24] Batu T, Fortnow L, Fischer E, Kumar R, Rubinfeld R, White P. Testing Random Variables for Independence and Identity. In: 42nd Annual Symposium on Foundations of Computer Science, FOCS 2001, Las Vegas, Nevada, USA, 14-17 October, 2001. IEEE Computer Society, 2001 pp. 442–451. doi:10.1109/SFCS.2001.959920. URL <https://doi.org/10.1109/SFCS.2001.959920>.
- [25] Diakonikolas I, Kane DM, Nikishkin V. Testing identity of structured distributions. In: Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015. SIAM, 2015 pp. 1841–1854. doi:10.1137/1.9781611973730.123. URL <https://doi.org/10.1137/1.9781611973730.123>.
- [26] Cai J, Green F, Thierauf T. On the correlation of symmetric functions. *Mathematical Systems Theory*, 1996. **29**(3):245–258. doi:10.1007/BF01201278.
- [27] Castro FN, Medina LA. Linear recurrences and asymptotic behavior of exponential sums of symmetric Boolean functions. *The electronic journal of combinatorics*, 2011. **18**(2):8. doi:10.37236/2004.
- [28] Castro FN, Medina LA. Asymptotic behavior of perturbations of symmetric functions. *Annals of Combinatorics*, 2014. **18**(3):397–417. doi:10.1007/s00026-014-0230-0.
- [29] Castro FN, Medina LA, Stănică P. Generalized Walsh transforms of symmetric and rotation symmetric Boolean functions are linear recurrent. *Applicable Algebra in Engineering, Communication and Computing*, 2018. **29**(5):433–453. doi:10.1007/s00200-018-0351-5.
- [30] Kavut S, Maitra S, Tang D. Construction and search of balanced Boolean functions on even number of variables towards excellent autocorrelation profile. *Designs, Codes and Cryptography*, 2019. **87**(2):261–276. doi:10.1007/s10623-018-0522-1.
- [31] Zhang Y, Yang G, Hung WNN, Zhang J. Computing affine equivalence classes of Boolean functions by group isomorphism. *IEEE Transactions on Computers*, 2016. **65**(12):3606–3616. doi:10.1109/TC.2016.2557329.
- [32] Qiu D, Zheng S. Generalized Deutsch-Jozsa problem and the optimal quantum algorithm. *Physical Review A*, 2018. **97**(6):062331. doi:10.1103/physreva.97.062331.
- [33] Ehlich H, Zeller K. Schwankung von polynomen zwischen gitterpunkten. *Mathematische Zeitschrift*, 1964. **86**(1):41–44. doi:10.1007/BF01111276.
- [34] Buhrman H, De Wolf R. Complexity measures and decision tree complexity: a survey. *Theoretical Computer Science*, 2002. **288**(1):21–43. doi:10.1016/S0304-3975(01)00144-X.

- [35] Deutsch D, Jozsa R. Rapid solution of problems by quantum computation. *Proceedings of The Royal Society A: Mathematical, Physical and Engineering Sciences*, 1992. **439**(1907):553–558. doi:10.1098/rspa.1992.0167.
- [36] Brassard G, Hoyer P, Mosca M, Tapp A. Quantum amplitude amplification and estimation. *Contemporary Mathematics*, 2002. **305**:53–74. doi:10.1090/conm/305/05215.
- [37] Gruska J, Qiu D, Zheng S. Generalisation of the Distributed Deutsch-Jozsa promise problem. *Mathematical Structures in Computer Science*, 2017. **27**(3):311–331. doi:10.1017/S0960129515000158.
- [38] Nayak A, Wu F. The quantum query complexity of approximating the median and related statistics. In: *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing*, Atlanta, Georgia, USA, May 1-4, 1999. ACM, 1999 pp. 384–393. doi:10.1145/301250.301349. URL <https://doi.org/10.1145/301250.301349>.
- [39] Johnson RA, Miller I, Freund JE. *Probability and Statistics for Engineers*. Pearson Education, 2017. ISBN 978-0321986245.
- [40] Mousavi N. How tight is Chernoff bound. 2010 URL <https://ece.uwaterloo.ca/~nmousavi/Papers/Chernoff-Tightness.pdf>.