WEIGHT DISCRIMINATION OF BOOLEAN FUNCTIONS WITH QUANTUM COMPUTATION

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

KIVANÇ UYANIK

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
PHYSICS

FEBRUARY 2014

Approval of the thesis:

WEIGHT DISCRIMINATION OF BOOLEAN FUNCTIONS WITH QUANTUM COMPUTATION

submitted by **KIVANÇ UYANIK** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Physics Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen
Dean, Graduate School of **Natural and Applied Sciences** ________________

Prof. Dr. Mehmet Zeyrek
Head of Department, **Physics** ________________

Assoc. Prof. Dr. Sadi Turgut
Supervisor, **Physics Department, METU** ________________

Examining Committee Members:

Prof. Dr. Yiğit Gündüç
Physics Engineering Department, Hacettepe University ________________

Assoc. Prof. Dr. Sadi Turgut
Physics Department, METU ________________

Prof. Dr. Namık Kemal Pak
Physics Department, METU ________________

Prof. Dr. Müge Boz Evinay
Physics Engineering Deptartment, Hacettepe University ________________

Assoc. Prof. Dr. Yusuf İpekoğlu
Physics Department, METU ________________

Date: ________________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name:	KIVANÇ UYANIK
Signature :	

ABSTRACT

WEIGHT DISCRIMINATION OF BOOLEAN FUNCTIONS WITH QUANTUM COMPUTATION

Uyanık, Kıvanç
Ph.D., Department of Physics
Supervisor : Assoc. Prof. Dr. Sadi Turgut

February 2014, 85 pages

In this thesis, we investigate solvability of the weight decision problem of two Boolean functions by quantum computation. In particular, we study this problem first from a general quantum operator discrimination perspective and second from a direct algorithmic viewpoint.

As quantum operator discrimination approach is concerned, we give two different formulations for two different cases. In one, the unitary transformations that correspond to the function evaluation are applied in a parallel fashion and in the other, they are applied only sequentially. Since the parallel case can always be simulated with a serial architecture, we put more emphasis on the serial approach and present a superior result in the serial setting. Specifically we show that any protocol with a serial application of p function evaluations interspersed with $p-1$ generic unitary operators in between can be uniquely mapped to a density matrix acting on some other Hilbert space.

In the direct approach, we generalize Grover's iteration in such a way that it can be run deterministically for the discrimination of Boolean functions. We show that sure-success weight distinguishability problem of two Boolean functions using a certain number of evaluations can be reduced to the problem of determining whether a point lies inside the convex hull of a curve. This convex analysis problem is further translated into a system of algebraic equations of a single variable. These equations are solved analytically for the case of single and two evaluations. For more evaluations numerical methods are utilized.

Keywords: Weight Decision Problem, Quantum Algorithm, Grover's Iteration, Quantum Information, Quantum Computation

ÖZ

KUANTUM HESAPLAMAYLA MANTIKSAL FONKSİYONLARIN
AĞIRLIKLARININ AYIRT EDİLMESİ

Uyanık, Kıvanç
Doktora, Fizik Bölümü
Tez Yöneticisi : Doç. Dr. Sadi Turgut

Şubat 2014 , 85 sayfa

Bu tezde kuantum hesaplamayla iki mantıksal fonksiyonun ağırlıklarının ayırt edilmesinin hesaplanabilirliği araştırılmıştır. Bu problem özel olarak önce genel bir kuantum operatör ayırt etme perspektifinden ve sonrasında da doğrudan bir algoritmik bakış açısıyla çalışılmıştır.

Kuantum operatör ayırt etme yaklaşımı söz konusu olduğunda iki farklı durum için iki farklı formülasyon verilmiştir. Bu durumlardan birinde fonksiyon hesaplamaya karşılık gelen üniter dönüşümler paralel olarak diğerinde ise seri olarak uygulanmaktadır. Paralel durum her koşulda seri bir mimariyle simüle edilebileceğinden seri yaklaşıma daha çok ağırlık verilmiş olup, çalışılan probleme daha ileri bir çözüm önerilmektedir. Özellikle, fonksiyonlara karşılık gelen p üniter operatörün aralarında $p-1$ genel üniter operatörle birlikte seri uygulamasından oluşan her protokolün p adet Hilbert uzayının tensör çarpımından oluşan uzayda etki eden farklı bir yoğunluk matrisine birebir eşlenebildiği gösterilmektedir.

Doğrudan yaklaşımdaysa Grover iterasyonu mantıksal fonksiyonların ayırt edilmesinde deterministik olarak çalışacak şekilde genellenmektedir. Mantıksal fonksiyonların deterministik olarak ve belli bir sayıda hesaplamayla ağırlık ayırt etme probleminin bir noktanın bir eğrinin konveks örtüsünün içinde olması problemine indirgenebileceği gösterilmektedir. Daha ötesi, bu konveks analiz problemi tek değişkenli cebirsel bir denklem sistemine indirgenmektedir. Bu denklemler bir ve iki hesaplama durumu için analitik olarak çözülmektedir. Daha fazla hesaplama durumu içinse numerik yöntemler kullanılmaktadır.

Anahtar Kelimeler: Ağırlık Ayırt Etme Problemi, Quantum Algoritma, Grover İterasyonu, Quantum Bilgi Sistemleri, Quantum Hesaplama

To my family

ACKNOWLEDGMENTS

It is a great honor to acknowledge my debt to my thesis advisor Sadi Turgut. Besides his patience, support and guidance throughout this work, I would always acknowledge him for being a great teacher and a scientist. I doubt any of his students would think any other way. My sincere appreciation extends to my jury committee also for their support and patience during this long project.

I would also like to thank all the academic and personnel stuff of METU. It may seem to be unnoticed but in my opinion we are all in debt for their hard labor.

I would especially like to express my thankfulness to my professors Ayşe and Atalay Karasu, Enver Bulur, Seçkin Kürkçüoğlu, Bayram Tekin and Hande Toffoli.

I also learned a lot from and had a lot of fun with my office roommates Sinan Değer and Cağan Aksak during my PhD. They always tolerated me and offered their best when I needed. I hope we will always be in close touch when we become colleagues.

I would like to thank my housemates Gökhan Kiper and Mehmet Ali Olpak for always cheering me up and being there for me in the times when I would most appreciate. I wish them a lifelong success.

The people I learned from never ends. I owe much to my dear friend and eternal Go rival Musa Burak and his family. I would like to thank him for always cheering me up and offering any kind of support without hesitation. I would also like to extend my gratitude to the members of METU Alpar Kılınç Go Club. I have always found a nice atmosphere and a friendly challenge in their companionship.

I would never pass without mentioning my thankfulness to Deniz for the most enjoyable times and drafts together. Also many thanks goes to Selma and "her big family" for their warm existence in my life.

Special thanks goes to Murat Öztürk, Zeynep Nilüfer Güven and Alireza Behtash for helping me reach my home after the long nights we studied in the empty corridors of physics department.

I can only mention my acknowledgement to my dear friends Umut Bostancı and Ertuğrul Karademir without whom my PhD years would have been very bleak.

There are many people I had wonderful time with, learned a lot from, appreciated for what he/she done with his/life, stood up for me, cheered me up and supported me numerous times during my PhD thesis. Here is a list of only some that I could mention in my limited time and space:

Ahmet Emre Onuk, Ayça and Volkan Hunerli, Ayşe Ilgın Sözen and Mehmet Demiray, Ayşen Pala, Berkin Yıldırım, Çagatay Menekay, Elif Uzcengiz and Bartu Şimşek, Kamil Çınar, Merve Demirtaş, Mustafa Devrim Kaba, Nader Ghazanfari, Nazım Dugan, Neli Gagua, Soner Albayrak, Sahin Kürekçi, Tahsin Çağrı Şişman.

All aside, this thesis would not be possible without my mother's unending love, support and blessings. Though far, I have always felt it channeling through long distances.

During my PhD study, I have been supported by TÜBİTAK with a 2211 PhD scholarship in the first five years.

TABLE OF CONTENTS

CHAPTERS

APPENDICES

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

Perhaps the most important progress in physics that took place during the early twentieth century was the transition from (what we call now)classical laws of physics to the modern theory of quantum mechanics. It is a theoretical framework for the construction of physical theories, rather than a collection of laws. Within this framework, it became possible to describe the interaction of fundamental particles and light up to an immense level of accuracy. Since then, it became an integral part of physical theories and its applications in science is widely acknowledged[1]. The implications of quantum mechanics were quite unfamiliar to the "classical" minds of physicists, however, as the experimental verifications accumulated, the idea that nature behaves in this bizarre way became more permanent. Until 1970's, on the other hand, scientists were unable to conduct these experiments by means of a complete control on the atomic scale. Even though they were able to explain numerous phenomena ranging from the interactions that take place inside the atoms to the working mechanism of neutron stars, scientists were deprived of such a low level manipulation of atoms or molecules.

In 1970's and 1980's this picture have begun to change. The capability of atomic force microscope to control single atoms on a surface, creation, detection, and manipulation of single photons and other similar techniques have become practical tools of physicists. Also in these years, the observation that the inefficiency of classical computers to simulate quantum mechanical processes led Feynman to speculate about a computer that its algorithmic operation is completely quantum mechanical[2, 3]. This idea was also suggested by Benioff[4, 5, 6] with a motivation from a self consistency perspective of quantum mechanics[7]. If such a general purpose quantum computer could be built, then it has been shown that, any local quantum system could be efficiently simulated[8]. However the striking thought of Deutsch was that whether any computational task, classical or quantum, could be simulated efficiently with a quantum computer[9]. Yet, this may be too strong a proposition and it hasn't been proven until now, however Deutsch and later Deutsch with Jozsa showed that there exist problems that quantum computers outperform classical computers on an exponential scale[9, 10].

Deutsch's influential papers have opened new avenues for both computer science and physics. His works are often attributed to mark the beginning of quantum computation. However before continuing on the short history of quantum computation, let us briefly review major breakthroughs in computer science. A good starting point may be the seminal paper by Alan Turing, which he wrote in 1936. In this work, he built a theoretical model for defining a class of functions known as "computable functions"[11]. With this formulation it was possible to show that every function that is "algorithmically computable" can be run on a hypothetical machine which later named as Turing machine[1]. There were two other models capturing the same idea by Church[12, 13] and Gödel and Herbrand. It was also shown that these three approaches are fundamentally equivalent[11], and the hypothesis of Turing is named as Church-Turing thesis honoring these two great mathematicians. In essence, as very well stated in [1]: "Turing claimed that the Universal Turing Machine completely captures what it means to perform a task by algorithmic means". These results are particularly useful for the quantification of the computational resources independent of the underlying computational model[11, 7]. Thanks to the basic architecture that Von Neumann developed a few years later, a computer that can work fully as capable as a Universal Turing Machine was ready to be built[1]. After the invention of transistors and later integrated circuits, electronics and computer engineers designed more efficient computer hardware ever since however the basic principle of computing remained the same.

Now let us return our attention to the Deutsch's arguments. He argued that whether a computational device that utilizes quantum mechanical laws of nature in order to *efficiently* simulate any physical system whether classical or quantum. An *efficient* simulation is defined as requiring only polynomial resources to simulate any computational model with a universal Turing machine therefore the Church-Turing thesis may take the strong form: “Any algorithmic process can be simulated efficiently using a Turing machine”[1]. Even when probabilistic algorithms are considered, the Church-Turing thesis is not fundamentally challenged. On the other hand when it is confronted with the idea of a quantum mechanical computer, a new field of research at the intersection of computer science and physics has been born. The classical versions of Turing machine[4, 5, 6, 9] and circuit model[9] are updated to the quantum versions and later their equivalence is proved[14]. The example of Deutsch and Jozsa distinguishes constant and balanced Boolean functions on a quantum computer exponentially faster than classical deterministic algorithms running on a classical computer. Nevertheless its performance could also be compared with the classical probabilistic algorithms and quantum speedup can not be achieved in this case when vanishingly small probability of error is allowed. Shortly after however Bernstein and Vazirani[15] gave instances of superpolynomially faster quantum algorithms. The next significant progress was due to Simon[16] in which he stated a problem where exponential speedup was achieved as compared to any classical algorithm.

Thus far, the problems that were posed to serve as an example of efficient quantum computation had little practical value. It was the remarkable paper by Shor that changed this picture. His algorithm could find prime factors of very large integers and discrete logarithm using polynomial time whereas the best classical algorithm that is known can solve this problem taking superpolynomial time[17]. This result has serious consequences in our daily lives. Most of the common cryptographic schemes are based on the difficulty of prime factoring. If someone can ever build a scalable quantum computer, the reliability of cryptographic systems would be compromised. Nonetheless a quantum computer that can operate Shor’s algorithm for large numbers is not expected to be built for quite a while. The other landmark in the field of quantum computation was the Grover's algorithm which can search through an unstructured database quadratically faster than classical search algorithms[18]. The other algorithms that we have mentioned provided much better speedups but they all depended on a premise in the problem. Indeed, later it was shown that all of those problems were instances of the more general hidden subgroup problem. In contrast, Grover's algorithm can be efficiently utilized regardless of the structure of the search space, thus making it much more applicable.

In this thesis, we study one of the applications of the Grover's algorithm. Specifically, we investigate solvability of the weight decision problem of two Boolean functions by quantum computation. We study this problem first from a general quantum operator discrimination perspective and second from a direct algorithmic viewpoint.

Quantum algorithms are fundamentally different from the classical ones. This is partly because it is highly nontrivial to utilize the nonclassical resources such as parallelism and entanglement that comes with the availability of quantum control at the atomic scale. An improvement in this area of science has consequences not only in the computer science and quantum physics but also in seemingly unrelated subjects such as astrophysics and biophysics. Therefore we aim to improve and expand our knowledge about this promising field of research.

Organization

Since the weight decision problem can be reformulated as an operator discrimination problem like many other quantum algorithms we start with a brief review of quantum operator discrimination in Chapter 2. All quantum operator discrimination protocols necessarily include a quantum state discrimination step in the end. For this reason, the first part of Chapter 2 is devoted to quantum state discrimination problem. The discrimination of two orthogonal states is rather trivial in quantum mechanics. However quantum states have a non-zero overlap in general and more sophisticated mathematical tools are necessary to distinguish such states. In particular we mainly review the two most common strategies for quantum state discrimination problem: minimum error and unambiguous discrimination. We also mention mixed strategies and possible extensions. After giving necessary background information about quantum state discrimination, we continue with the quantum operator discrimination problem. Even though they are reduced to state discrimination in the end, there are fundamental differences between quantum operator and quantum state discrimination tasks. Unlike states, operators can be utilized more than once in parallel, sequential or any combination of these and there is the choice freedom in the initial state we fed to the operator. The rest of Chapter 2 continues with a short survey of literature on operator discrimination and a powerful mathematical relation between quantum states and quantum operators, namely Choi-Jamiolkowski isomorphism[19, 20].

In Chapter 3, we present our first contribution. We interpret the weight decision problem of two functions as a set discrimination problem and first try to solve it only for the case, in which, quantum operators are applied only in parallel. Then we continue with a more extensive approach where we discuss the problem of set discrimination with operators applied sequentially. This approach is more generic because it can be shown that any scenario involving combinations of both serial and parallel applications of operators can be equivalently described with a protocol that consist of only sequential applications[21, 22].

In Chapter 4, we start with a brief review of Grover's algorithm as our second contribution and the other related works in the literature on weight decision problem are based on Grover iteration. We continue with quantum counting[23], a quantum algorithm that can give the number of solutions to an unstructured database search problem faster than classical algorithms. Then, we advance to our specific problem: weight decision of Boolean functions. It is first studied by Braunstein, Choi and others in [24, 25, 26], thus we review their algorithm and results thoroughly before giving our treatment to the problem. The other approach to the problem is our work given in [27] where we introduce an exact version of Grover iteration and utilize it to discriminate weights. We present our findings with this method, give analytical and numerical results towards the end of Chapter 4. We finalize this chapter by comparing the complexities of classical algorithms, quantum counting, Braunstein and Choi's method and our analysis.

Our conclusions and a short future outlook is the content of Chapter 5.

CHAPTER 2

QUANTUM OPERATOR DISCRIMINATION

In quantum computation, we usually model an unknown classical function as an oracle. This way we do not worry about the internal workings of the function. In other words, it is like having a black box; we can only change its input and make observations at its output. This is a quantum oracle, usually realized by a unitary evolution, which means that we can also feed superposition of different qubits as input and obtain corresponding superposition of evaluations at the output. This is a direct consequence of quantum parallelism. On the other hand, calling such a function, whether classical or quantum, costs a precious resource: time. So we try to make use of counterintuitive properties of quantum mechanics such as parallelism and entanglement in order to solve a problem significantly faster than we would have with a classical implementation. The examples where quantum implementation of such a problem is much faster than the classical ones can usually be reduced to the identification of a feature of such a function or complete identification of the function itself. Therefore, oracle identification or oracle discrimination is a critical part of quantum computation.

In quantum operator discrimination we try to optimize both the circuit architecture and the input state in order to minimize the number of operator calls. In a typical quantum operator discrimination scenario, we start with a fixed input state consisting of register qubits and ancilla qubits, evolve the joint system unitarily with the oracle operator given to us and with some other unitary operators we introduced for optimization and make a measurement to distinguish output states in the end. If we are allowed we can use the operator to be distinguished more than once, however this would also increase the cost, so we should be careful while designing such protocols. If the final states are orthogonal to each other, then there won't be any problem. In that case we can always find a measurement that would give different results corresponding to each different state. However most of the time the states are not orthogonal to each other and non-orthogonal state discrimination is a nontrivial problem. Therefore, before delving into quantum operator discrimination we should at least briefly review quantum state discrimination.

This chapter will continue with a short presentation of quantum state discrimination problem in section 2.1. Two main strategies and other possibilities will be briefly mentioned. Based on the discussions on state discrimination, an overview of quantum operator discrimination will be reviewed in section 2.2.

2.1 Quantum State Discrimination

State discrimination is itself a main topic in quantum information. In most general form, we deal with ensembles of quantum states with known (or sometimes unknown) probabilities. Basically, we are confronted with the following problem. N mixed states ρ_i are prepared with probabilities η_i, $\sum_i \eta_i = 1$. One of them is secretly selected and given to us from this ensemble. The only accessible information we have beforehand is the states ρ_i and corresponding probabilities η_i. Our task is to determine which state out of these is prepared and given using quantum measurements. Simplest case would be distinguishing mutually orthogonal pure states, $|\psi_i\rangle$:

$$\langle\psi_i|\psi_j\rangle = \delta_{ij}. \tag{2.1}$$

To distinguish these states, it is enough to make a simple Von Neumann measurement. Applying the projection operators

$$M_i \equiv |\psi_i\rangle \langle\psi_i| \tag{2.2}$$

is sufficient to successfully determine the given unknown state. Notice that since the states $|\psi_i\rangle$ are orthogonal to each other, the operators M_i add up to identity, $\sum_i M_i = \mathbb{1}$, therefore no inconclusive or wrong result is possible. On the other hand, nonorthogonal states cannot be perfectly discriminated. This had been known long before the quantum information tasks came into existence. The problem of discrimination of nonorthogonal states is an active area of research in quantum information. There are two main strategies for discrimination of non-orthogonal quantum states

- Minimum error strategy (ME)
- Unambiguous discrimination (UD).

There other also other strategies like maximal confidence discrimination or minimax discrimination; but quantum state discrimination literature is mostly gathered around these two fundamental approaches.

2.1.1 Minimum Error Strategy

The first strategy is due to independent works of Helstrom[28] and Holevo[29] and have been introduced as early as late 70's. In this method, we are given a state ρ that is unknown to us and has been chosen from an ensemble of N mixed states ρ_i, which are not necessarily mutually orthogonal. In this ensemble, each state ρ_i has a corresponding probability η_i to be prepared, such that they add up to one: $\sum_{i=1}^{N} \eta_i = 1$. We aim to specify which state has been selected from the ensemble while allowing minimal rate of erroneous results. Finding the operators, which give rise to minimum total probability of error would be sufficient. These operators could be formally described with positive operator valued measures (POVMs), $\{\Pi_j\}_{j=1}^{N}$, $\sum_j \Pi_j = \mathbb{1}$ such that each Π_j represents an operator that its measurement outcome, j, indicates that the prepared state was ρ_j. If we denote the unknown state given to us by ρ, then the probability of concluding that it was prepared as ρ_j is given by $\mathrm{Tr}\,(\Pi_j \rho)$. In [28, 29], the minimal probability of error for distinguishing two states is found to be

$$p_E = \frac{1}{2} - \frac{1}{2} \left\| \Lambda \right\| , \tag{2.3}$$

for an ensemble of two mixed states ρ_1 and ρ_2, where $\| \ \|$ is the trace norm defined as $\|\Lambda\| \equiv \mathrm{Tr}\sqrt{\Lambda^\dagger \Lambda}$ and Λ is the Hermitian operator $\Lambda \equiv \eta_2\rho_2 - \eta_1\rho_1$. This result is known as Helstrom bound in the literature. Finding optimal solutions analytically for more than two states is nontrivial in general and may be a difficult problem, however the conditions that must be satisfied by the POVMs are known [28, 30, 31]. Various alternative approaches for the known solutions and solutions for special cases have been studied in several papers and the interested reader may check [32] for a comprehensive review of the subject.

2.1.2 Unambiguous Discrimination

In unambiguous discrimination strategy, we are not allowed to make an error. But it is possible to have an **inconclusive** result. It is first proposed by Ivanovic[33] and solved for $N = 2$ by Dieks[34] and Peres[35]. To give an idea about how the method works, let us assume that we are trying to distinguish two states ρ_1 and ρ_2. In this case, we can formulate the strategy with the use of two POVM operators Π_1 and Π_2, in order to decide whether the unknown state was ρ_1 or ρ_2 respectively. In this process there is zero probability that the first state ρ_1 would be inferred by the measurement Π_2 or vice versa, however as a drawback of this accuracy, we have to introduce another measurement. To identify this inconclusive decision, we need another POVM operator, Π_0, such that all the operators add up to identity: $\Pi_0 + \Pi_1 + \Pi_2 = \mathbb{1}$. If the measurement gives the result i, $i = (1, 2)$, corresponding to the operator Π_i, we can be sure that the supplied state was ρ_i. On the other hand if the measurement gives 0, we do not obtain any information about the identity of the state. Like the minimal error strategy, we are interested in the optimum results, so the task can be reformulated as finding the operators Π_i^{OPT}, which give minimum total probability of inconclusive result. For two pure states $|\psi_1\rangle$ and $|\psi_2\rangle$, one can choose $\Pi_1 = c_1 \left|\psi_2^\perp\right\rangle \left\langle\psi_2^\perp\right|$ and $\Pi_2 = c_2 \left|\psi_1^\perp\right\rangle \left\langle\psi_1^\perp\right|$ where $\left|\psi_i^\perp\right\rangle$'s denote the vectors orthogonal to $|\psi_i\rangle$ and c_1 and c_2 are the coefficients that are to be found from the optimization conditions. One can easily see these operators will not lead to incorrect results. Π_1 and Π_2 are positive definite by construction, but the last operator Π_0 should also be positive definite since it corresponds to a physical operation too [32]. In [32], the optimum failure probability for this case is found as

$$Q^{OPT} = \begin{cases} 2\left(\eta_1\eta_2\right)^{\frac{1}{2}}\cos\left(\Theta\right) & \text{if } \frac{\cos^2(\Theta)}{1+\cos^2(\Theta)} \leq \eta_1 \leq \frac{1}{1+\cos^2(\Theta)}, \\ \eta_1 + \eta_2\left|\langle\psi_1|\psi_2\rangle\right|^2 & \text{if } \eta_1 < \frac{\cos^2(\Theta)}{1+\cos^2(\Theta)}, \\ \eta_1\left|\langle\psi_1|\psi_2\rangle\right|^2 + \eta_2 & \text{if } \eta_1 > \frac{\cos^2(\Theta)}{1+\cos^2(\Theta)}, \end{cases} \tag{2.4}$$

where $\cos\Theta = |\langle\psi_1|\psi_2\rangle|$ and η_i are the preparation probabilities of ρ_i. Derivation of this result is straightforward but there is a bit of some laborious algebraic manipulation. The solution is partially covered in [32]. Since the main focus of this work is not the state discrimination, we do not include the intermediate steps. However, the interpretation of Eq. 2.4 is important. This result reveals that the POVM $\{\Pi_i\}$'s that we have defined lead to the minimal error, however only when it is possible to define them. In the extreme cases where the probability of preparing one of the states is too high, the POVM's doesn't exist and simply guessing with an operator using the orthogonal direction to that state give the lowest probability of error. In figure 2.1, the results of minimum error discrimination and unambiguous discrimination for two pure states is shown. To be able to make a comparison with the minimum error strategy, error probabilities corresponding to ME strategy are also included.

When the number of states to be distinguished is larger than two, the problem becomes complicated. Discrimination of n pure states has been worked out in [36] for example, however an analytical formula has not been found at the time this thesis was written. Pure states to be discriminated should be linearly independent [37], but if there are more than one copies available, linearly dependent states can also be unambiguously distinguished [38].

Unlike minimum error discrimination, unambiguous discrimination is progressed very differently for pure and mixed states [32]. Unfortunately, unambiguous discrimination even for two mixed states is difficult in general and there is ongoing research about the subject. For mixed states to be unambiguously distinguishable, they should satisfy the following condition: the states $\{\rho_i\}$ are unambiguously distinguishable if and only if $\text{supp}\,(\rho_k) \neq \text{supp}\,(\rho_1, ..., \rho_{k-1}, \rho_{k+1}, ..., \rho_n) \quad \forall k : 1 \leq k \leq n$ [39], where the support of a density matrix ρ, $\text{supp}\,(\rho)$, is the subspace spanned by the eigenvectors corresponding to positive eigenvalues of ρ and the support of a set of density matrices is defined to be the sum of each one's support [40, 41].

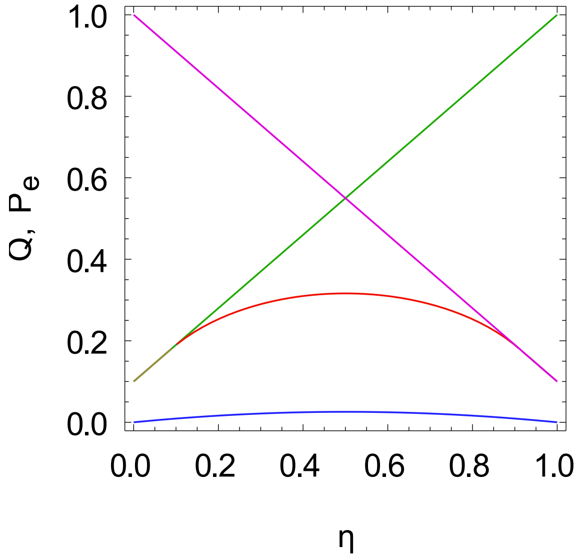

Figure 2.1: $\{Q, P_e\}$(optimum error probabilities) vs. η_1(probability of preparing $|\psi_1\rangle$) for minimum error (P_e - curve below) and unambiguous discrimination (Q - curve tangent to the linear lines) strategies evaluated for a representative value of $|\langle\psi_1|\psi_2\rangle|^2 = 0.1$. For too small η_1 below some threshold or too high above some threshold, error corresponding to UD is linear with η_1 as given in Eq. (2.4)[32].

2.1.3 Other Strategies

Unambiguous discrimination approach can be generalized to the cases where the states are not linearly independent. One way to do this is through the method of maximal confidence discrimination[42]. In this approach, one optimizes the quantity C_i, which is called confidence. It is defined as the conditional probability, $P(\rho_i|i)$, of the initial state being ρ_i, given the outcome i is detected. For linearly independent states this strategy reduces to unambiguous discrimination.

One can also combine different aspects of these strategies together to search for further optimized solutions in between. As indicated in [32], for linearly independent pure states, a strategy that interpolates between ME and UD strategies [43, 44] and for mixed states a unification of ME and maximum confidence strategy [45] has been studied. Another interesting approach is constructing POVMs directly if possible and test if they give optimized results. A successful example of this can be found in [46].

In a different approach, one relaxes the condition that the probabilities η_i are known beforehand. Having this restriction in an approach is called Bayesian. In contrast with the methods we have discussed up to now, there may be scenarios where actual probabilities are not relevant or not known as in a noncooperative cryptographic scenario[47]. In a non-Bayesian setting one applies a so-called minimax approach to find the optimal minimum error or unambiguous discrimination rate. It was first introduced in [47]. In [47], the maximum of the smallest probabilities of correct detection for both ME and UD constraints were considered.

In general, these optimization problems are very difficult to be solved analytically, however some of them are straightforwardly applicable to the known numerical optimization methods. References to the examples of these applications can be found in the comprehensive review by Bergou[32].

2.2 Quantum Operator Discrimination

One of the immediate applications of state discrimination is quantum operator/channel discrimination. The quantum channels or operations are the mathematical models that represent the total effect that is acted on a quantum state within a physical medium or a black box without dealing with the inner workings of the process. We ignore such information because we do not have any access to the information about the details or knowing such details would merely complicate our calculations. On the other hand, we can control the state that we feed as the input. An example that is particularly important for this thesis will be the unitary quantum operators which are used to implement classical functions. These operators are also called “oracle”, “quantum oracle” or “oracle operator”. The importance comes from the observation that all well-known quantum algorithms can be interpreted as an oracle discrimination problem [48].

A common formulation of the simplest version of quantum operator discrimination can be formulated as follows: We act the operators $\varepsilon_i(\cdot)$ we wish to distinguish on an input state ρ. Then discriminate the output states $\varepsilon_i(\rho)$ with minimum error discrimination[49, 50, 51, 52], unambiguous discrimination [41, 48, 53] or any other state discrimination scheme. With this primitive approach, the operator discrimination problem inherits many features of the quantum state discrimination, thus most of the known results can be immediately applied to the operator discrimination problem as well. Nevertheless this is the simplest scheme to be devised. Unlike quantum states, operators can be applied more than once, in a parallel, sequential or a mixed arrangement. This additional feature makes it possible to distinguish unitary quantum channels perfectly with a finite number of uses, in contrast with quantum states, for which, infinite number of copies are required for perfect discrimination in general[54].

As an introductory example, we can follow the steps in Sacchi’s work[51], where ME discrimination of two states is studied. Using Equation (2.3), for error probability of distinguishing output states, we obtain

$$P_E = \min_{\{\rho\in\mathcal{H}\}} \left\{ \frac{1}{2} - \frac{1}{2} \left\| \eta_2\varepsilon_2(\rho) - \eta_1\varepsilon_1(\rho) \right\| \right\} = \frac{1}{2} - \frac{1}{2} \max_{\{\rho\in\mathcal{H}\}} \left\| \eta_2\varepsilon_2(\rho) - \eta_1\varepsilon_1(\rho) \right\| , \tag{2.5}$$

where the minimization is over the input states. It is sufficient to deal with pure states, since quantum operations are linear and the trace norm satisfy the following property[55]

$$\left\| cA + (1-c)\, B \right\| \leq c \left\| A \right\| + (1-c) \left\| B \right\| \tag{2.6}$$

with $0 \leq c \leq 1$. If we add an ancilla with a Hilbert space $\mathcal{H}'$ of dimensionality $\dim \mathcal{H}'$ and use entangled states from $\mathcal{H}\otimes\mathcal{H}'$, the expression for error probability becomes

$$P_E^{ent} = \frac{1}{2} - \frac{1}{2} \max_{\{\rho\in\mathcal{H}\otimes\mathcal{H}'\}} \left\| \eta_2(\varepsilon_2 \otimes \mathbb{1})\rho - \eta_1(\varepsilon_1 \otimes \mathbb{1})\rho \right\| . \tag{2.7}$$

As a simplification considering the evaluation of the norm in Eq. 2.7, for a finite dimensional Hilbert space, it is enough to choose $\dim \mathcal{H}'$ to be at most $\dim \mathcal{H}$ [56, 57].

As it was done with the states, discrimination of operators can be also performed unambiguously. In this case we apply an unambiguous discrimination procedure to a set of output states $\varepsilon_i(\rho)$ for an input ρ which now minimize the rate of inconclusive result. Unambiguous discrimination of states was first applied to operators in Bergou's work[53]. UD of oracle operators were thoroughly discussed in Chefles' paper [48]. Necessary and sufficient conditions for unitary operators to be unambiguously distinguishable are presented by Wang and Ying [41].

One of the most vital parameters that is needed to be discussed in this context is entanglement. Even though "All entangled states are useful for channel discrimination"[58], there are cases where local entanglement[59] in multipartite operators, or entanglement with an auxiliary system [48, 21, 60] is not essential for perfect discrimination between unitary operations. Nevertheless, entanglement "can enhance the distinguishability of entanglement-breaking channels"[51], "can be used to improve the precision of quantum measurements for either precision or stability"[50] and "can allow us to better distinguish operations" [61]. In addition to these, entanglement help increasing success probability of discrimination[62] and is found to be useful in many discrimination cases[22, 41, 63, 64, 65] as compared to protocols without entanglement.

We would like to finalize this chapter with disclosing an effective theoretical tool for associating quantum channels/operators with quantum states and vice versa. It was proved that any trace-preserving completely positive operator $\varepsilon(\cdot)$ acting on a D dimensional quantum system can be associated to a density matrix $\omega_\varepsilon = (\mathbb{1} \otimes \varepsilon)\,\Omega$ which is an element of a $D \times D$ dimensional Hilbert space $\mathcal{H}^D \otimes \mathcal{H}^D$, where $\Omega \equiv \left|\Psi_D^+\right\rangle \left\langle\Psi_D^+\right|$ and $\left|\Psi_D^+\right\rangle \equiv \sum_{j=1}^{D} |j\rangle \otimes |j\rangle$ is the (unnormalized) maximally entangled state on the Hilbert space $\mathcal{H}^D \otimes \mathcal{H}^D$. This result is due to the work of Jamiolkowski[19] and the improvements of Choi[20], hence it is usually called Choi-Jamiolkowski isomorphism in the literature. With this correspondence, one can apply almost all the main theorems concerning the states to their operator analogue and vice versa[66]. One example may be given from the discrimination problem of unitary operations. Suppose we would like to distinguish two distinct unitary operators U and V acting on D dimensional Hilbert space $\mathcal{H}^D$. From the Choi-Jamiolkowski isomorphism, we know that there exist two density matrices ω_U and ω_V which are elements of $\mathcal{H}^D \otimes \mathcal{H}^D$ corresponding to these operators. Since the supports of ω_U and ω_V are different, these states are always unambiguously distinguishable, thus we can immediately say that the operators U and V are always distinguishable in an unambiguous fashion.

CHAPTER 3

QUANTUM SET DISCRIMINATION

In this chapter, we give our solutions to the specific problem of set discrimination. These sets will correspond to the Boolean functions we want to distinguish through their weights. These functions of consideration will be implemented by unitary operators that mutually commute with each other. As the operators can be applied multiple times before a measurement is performed, how these steps will be composed make quite a difference in the analysis. There can be many architectures possible however there are two main schemes of combinations. First it is possible to apply these operators in a parallel fashion so that they act on their input states independent from each other. They may act on an entangled input state however their individual actions would not have any effect on each other's outputs. We call this type of architecture, *parallel* scheme. The second one is the serial scheme where the unitaries are concatenated one after another so that the input of one operator is either the initial state or an output of another operator. Additionally it is possible to alter the state through numerous processes between two operators but it is possible to take into account of all these effects with a single unitary. We call such a scenario of acting operators a *serial* or a *sequential* scheme.

3.1 A Parallel Scheme

Most quantum algorithms are implemented in a "state preparation - a quantum process involving finite number of calls to an oracle - final measurement" form. We are able to choose an input state and design a circuit architecture but finally decide with a finite number of measurements in the end. Therefore the algorithm has been eventually recast as an oracle identification task. In Deutsch-Jozsa algorithm[10] this is the most transparent. However, some other famous quantum algorithms, Grover's algorithm[18] and even Shor's[17] algorithm can be reformulated as discrimination of quantum operators or sets of quantum operators[48].

3.1.1 Construction of the problem

Let $\mathscr{F}^n$ be a set of vectors $f = (f_1, \ldots, f_n)$ such that $f_i = 0$ or 1:

$$\mathscr{F}^n = \{f = (f_1, \ldots, f_n) : f_i = 0 \text{ or } 1\} . \tag{3.1}$$

Let us define a subset $\mathscr{F}^n_r \subset \mathscr{F}^n$ such that

$$\mathscr{F}^n_r = \{f : \text{exactly } r \text{ components are } 1, \text{other } n - r \text{ components are } 0\} \tag{3.2}$$

An example with $n = 5$ and $r = 2$ is given in Table 3.1.

Table 3.1: An example set :F^5_2

	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}
x_1	1	1	1	1	0	0	0	0	0	0
x_2	1	0	0	0	1	1	1	0	0	0
x_3	0	1	0	0	1	0	0	1	1	0
x_4	0	0	1	0	0	1	0	1	0	1
x_5	0	0	0	1	0	0	1	0	1	1

We can interpret the vectors as functions by $f(i) = f_i$. Quantum implementation of these functions, namely unitaries U_f, corresponding to the functions f act on $\log N$ bit input states such that

$$U_f \left|x\right\rangle_I \left|y\right\rangle_R = \left|x\right\rangle_I \left|f(x) \oplus y\right\rangle_R , \tag{3.3}$$

where x can take values from 1 to N. Here the first qubit is the input qubit which doesn't change through the operation and the second qubit is the result qubit where the evaluated function is added to the previous value of the register. This choice of implementation allows us to make use of either $|-\rangle \equiv \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right)$ or $|+\rangle \equiv \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right)$ for initial result qubit states depending on whether change of phase is needed or not. This can be seen when the actions of unitaries on these states are evaluated

$$U_f\,|x\rangle\,|\pm\rangle = (\pm 1)^{\,f_i(x)}\,|x\rangle\,|\pm\rangle\,. \tag{3.4}$$

Our problem is to find out whether it is possible to discriminate between $\mathscr{F}_r^n$ and $\mathscr{F}_{r'}^n$ by processing p identical unitaries U_f in parallel with zero error probability. In this thesis, we will consider only sets $\mathscr{F}_r^n$ defined in Eq. (3.1) where of course many other subsets can also be studied. However, before that, we will deal with the problem with two sets $\mathscr{F}_r^n$ and $\mathscr{F}_0^n$ where the second set have only one function with a constant output zero. Here one can quickly recognize Deutsch - Jozsa problem with the choice $r = n/2$ with an addition of the other constant function $f(x) = 1$ to the second set, i.e. distinguishing $\mathscr{F}_{n/2}^n$ and $\mathscr{F}_0^n \cup \mathscr{F}_n^n$. In the parallel scheme, there are p input registers and p result qubits which are prepared in an initial entangled state. The initial state can be chosen as

$$|\psi_p\rangle = \sum_{l=0}^{p} \sqrt{\alpha_l} \left(\sum_{1\le i_1 < i_2 < \cdots < i_l \le n} \left| i_1 i_1 \cdots i_l \underbrace{111\cdots 1}_{p-l} \right\rangle \left| \underbrace{--\cdots}_{l} \underbrace{++\cdots+}_{p-l} \right\rangle \right) \tag{3.5}$$

where α_l will be determined later. For example for $p = 3$, the initial state $|\psi_3\rangle$ of 3 input registers and 3 result qubits would be

$$\begin{aligned} |\psi_3\rangle \;\; = \;\; & \sqrt{\alpha_0}\,|111\rangle\,|+++\rangle + \sqrt{\alpha_1}\sum_{i=1}^{n}|11i\rangle\,|++-\rangle \\ & + \sqrt{\alpha_2}\sum_{1\le i<j\le n}|1ij\rangle\,|+--\rangle + \sqrt{\alpha_3}\sum_{1\le i<j<k\le n}|ijk\rangle\,|---\rangle\,. \end{aligned} \tag{3.6}$$

We chose the state in Eq. (3.5) because this state is optimal for a permutationally symmetric problem like the one we are discussing now. α_l's are parameters we want to find which add up to one with the normalization condition:

$$\langle \psi_p | \psi_p \rangle = \sum_{l=0}^{p} \binom{n}{l} \alpha_l = 1. \tag{3.7}$$

To make a decision with zero error probability, all possible combinations of unitaries U_f, U_g chosen from different sets $\mathscr{F}_1$ and $\mathscr{F}_2$ (i.e., $U_f \in \mathscr{F}_1$, $U_g \in \mathscr{F}_2$ and $\mathscr{F}_1 = \mathscr{F}_0^n$, $\mathscr{F}_2 = \mathscr{F}_r^n$ in our case) applied to the state $|\psi_p\rangle$ should give vanishing inner products

$$\left\langle \psi_p | \left(U_f^\dagger\right)^{\otimes p} U_g^{\otimes p} | \psi_p \right\rangle = 0. \tag{3.8}$$

Choosing one of the sets to be the constant set $\mathscr{F}_0^n$, which is the set such that it includes only one function with constant output zero, may seem to be too much of a simplification, however it is a proper first step towards the $\mathscr{F}_r^n - \mathscr{F}_{r'}^n$ discrimination problem. To show that, let us first consider a single call ($p = 1$) for the discrimination of any sets $\mathscr{F}_1$ and $\mathscr{F}_2$. The initial state is

$$|\psi_1\rangle = \sqrt{\alpha_0}\, |1\rangle\, |+\rangle + \sqrt{\alpha_1} \sum_i |i\rangle\, |-\rangle\,. \tag{3.9}$$

States after unitaries corresponding to $f_1 \in \mathscr{F}_1$ and $f_2 \in \mathscr{F}_2$ applied are

$$\begin{aligned} U_{f_1}\, |\psi\rangle &= \sqrt{\alpha_0}\, |1\rangle\, |+\rangle + \sqrt{\alpha_1} \sum_i (-1)^{f_1(i)}\, |i\rangle\, |-\rangle \text{ and} & (3.10)\\ U_{f_2}\, |\psi\rangle &= \sqrt{\alpha_0}\, |1\rangle\, |+\rangle + \sqrt{\alpha_1} \sum_i (-1)^{f_2(j)}\, |j\rangle\, |-\rangle\,. & (3.11) \end{aligned}$$

The inner product becomes

$$\begin{aligned} \left\langle \psi_1 | U_1^\dagger U_2 | \psi_1 \right\rangle &= \alpha_0 + \alpha_1 \sum_i \sum_j (-1)^{f_1(i) \oplus f_2(j)} \langle i | j \rangle & (3.12)\\ &= \alpha_0 + \alpha_1 \sum_i \sum_j (-1)^{f_1(i) \oplus f_2(j)} \delta_i^j & (3.13)\\ &= \alpha_0 + \alpha_1 \sum_i (-1)^{(f_1 \oplus f_2)(i)}. & (3.14) \end{aligned}$$

It can be seen that if we had been started with $\mathscr{F}_0$ and $\mathscr{H} \equiv \mathscr{F}_1 \oplus \mathscr{F}_2$ instead, where $\mathscr{H}$ contains only the functions $h_{12} = f_1 \oplus f_2$ the result wouldn't have been changed. It is also straightforward to prove it for more than one oracle calls. Even though this is not equal to the $r - r'$ discrimination problem yet, it is possible to extend it to our problem, at least theoretically. Now, the equations to be satisfied for p-shot discrimination becomes

$$\begin{aligned}\left\langle \psi_p | U_0^\dagger U_i | \psi_p \right\rangle &= \langle \psi_p | \mathbb{1} U_i | \psi_p \rangle = \langle \psi_p | U_i | \psi_p \rangle \\ &= \sum_{l=0}^{p} A(n,r,l)\alpha_l = 0 \qquad (3.15)\end{aligned}$$

where $A(n, r, l)$ are defined as

$$A(n,r,l) \equiv \sum_{m=0}^{min(r,l)} (-1)^m \binom{r}{m}\binom{n-r}{l-m}. \qquad (3.16)$$

Using the identity $\sum_m \begin{pmatrix} n \\ r \end{pmatrix} \begin{pmatrix} n-r \\ l-m \end{pmatrix} = \begin{pmatrix} n \\ l \end{pmatrix}$, we can express the Eq. 3.15 as

$$\left\langle \psi_p | U_0^\dagger U_i | \psi_p \right\rangle = \sum_{l=0}^{p} B(n,r,l)\alpha_l = \frac{1}{2} \qquad (3.17)$$

where $B(n, r, l)$ are defined as

$$B(n,r,l) \equiv \sum_{m \;\; odd}^{min(r,l)} \binom{r}{m}\binom{n-r}{l-m}. \qquad (3.18)$$

Eq. (3.15) (or Eq. (3.17)) together with the normalization condition in Eq. (3.7) for α_l's completely describes the problem.

3.1.2 Solution of the problem

Let's consider k^{th} extreme solution. Namely, the solution that all α_i are zero except α_0 and α_k

$$\alpha_l^{(k)} = \begin{cases} 0 & if \;\; l \geq 1, \quad l \neq k, \\ \frac{1}{B(n,r,k)} & if \;\; l = k, \\ -\frac{A(n,r,k)}{2B(n,r,k)} & if \;\; l = 0. \end{cases} \tag{3.19}$$

Since it is necessary that $\alpha_l^{(k)} \geq 0$ for an acceptable solution, the coefficient $A(n,r,k)$ should be less than zero. Any acceptable solution is a convex combination of extreme solutions, namely $\alpha^{(k)}$'s.

$$\overrightarrow{\alpha} = \lambda^{(1)}\alpha^{(1)} + \lambda^{(2)}\alpha^{(2)} + \cdots + \lambda^{(p)}\alpha^{(p)} \tag{3.20}$$

where $\lambda^{(j)} \geq 0 \;\forall j, \;\; 1 \leq j \leq p$ and $\sum \lambda^{(j)} = 1$. Therefore finding minimum k satisfying $A(n,r,k) < 0$ is enough. First few terms of $A(n,r,k)$ and $B(n,r,k)$ are given below

$$\begin{aligned} A(n,r,0) &= 1, & B(n,r,0) &= 0, & (3.21)\\ A(n,r,1) &= n - 2r, & B(n,r,1) &= r, & (3.22)\\ A(n,r,2) &= \frac{1}{2}\left((n-2r)^2 - n\right), & B(n,r,2) &= r(n-r), & (3.23)\\ A(n,r,3) &= \bigg(\frac{1}{6}(n-2r)^3 & B(n,r,3) &= \frac{1}{2}r(n-r)(n-r-1) & \\ &\quad + (3n-2)(n-2r)\bigg), & &\quad + \frac{1}{6}r(r-1)(r-2). & (3.24)\\ &\vdots & &\vdots & \end{aligned}$$

At this point, we can advance to the next step, that is distinguishing the union of two sets $\mathscr{F}_r^n \cup \mathscr{F}_{r'}^n$ from $\mathscr{F}_0^n$. This time we have an extra equation for r' that must be satisfied,

$$\sum_{l=0}^{p} B(n,r',l)\alpha_l = \frac{1}{2} \tag{3.25}$$

Solving either Eq. (3.15) or Eq. (3.17) seems to be difficult for a general p-evaluation discrimination. We start again with the case $p=1$ first

$$\alpha_0 B(n,r,0) + \alpha_1 B(n,r,1) = 1/2, \tag{3.26}$$

$$\alpha_0 B(n,r',0) + \alpha_1 B(n,r',1) = 1/2, \tag{3.27}$$

$$\alpha_0 + \binom{n}{1}\alpha_1 = 1. \tag{3.28}$$

There are three equations and two variables, therefore most probably there is no solution at all. Indeed this is the case since the solution is

$$\frac{1}{2r} = \alpha_1 = \frac{1}{2r'} \tag{3.29}$$

$$1 - \frac{n}{2r} = \alpha_0 = 1 - \frac{n}{2r'} \tag{3.30}$$

and it is meaningless to have $r = r'$.

For the case $p=2$, the equations are

$$\alpha_0 B(n,r,0) + \alpha_1 B(n,r,1) + \alpha_2 B(n,r,2) = 1/2, \tag{3.31}$$

$$\alpha_0 B(n,r',0) + \alpha_1 B(n,r',1) + \alpha_2 B(n,r',2) = 1/2, \tag{3.32}$$

$$\alpha_0 + \binom{n}{1}\alpha_1 + \binom{n}{2}\alpha_2 = 1. \tag{3.33}$$

Employing the convex analysis that has been introduced in the beginning of the subsection with Eq. (3.19) we reach the solution

$$\alpha_0 = 1 - \frac{n(r - r' - \frac{n+1}{2})}{2rr'}, \tag{3.34}$$

$$\alpha_1 = \frac{r + r' - n}{2rr'}, \tag{3.35}$$

$$\alpha_2 = \frac{1}{2rr'}. \tag{3.36}$$

Positivity of the coefficients α_i gives possible r and r' values for a given size of the problem n for which the discrimination problems can be solved with only $p = 2$ evaluations. The conditions in terms of n, r, and r' can be summarized as

$$(r - r')^2 \leq r + r' \quad \text{and} \quad n \leq r + r' \tag{3.37}$$

$$\begin{aligned} (r - r')^2 > r + r' \quad \text{and} \quad n &\leq r + r' - \frac{1}{2} \\ &- \sqrt{(r - r')^2 - (r + r') + \frac{1}{4}} \end{aligned} \tag{3.38}$$

In Figure 3.1, the solution set for $n = 100$ is plotted. The solution is symmetric in r and r'. This figure gives an idea for which r values we can discriminate sets $\mathscr{F}_r^n$ from $\mathscr{F}_1^n$. r must be at least approximately half of n. That is because, distinguishing $\mathscr{F}_r^n$ from $\mathscr{F}_1^n$ and distinguishing $\mathscr{F}_{r-1}^n \cup \mathscr{F}_{r+1}^n$ from $\mathscr{F}_0^n$ is the same problem. This follows from the fact that any function $h_{12} = f_1 \oplus f_2$ is (where $f_1 \in \mathscr{F}_1^n$ and $f_2 \in \mathscr{F}_r^n$) has a number of roots $r - 1$ or $r + 1$.

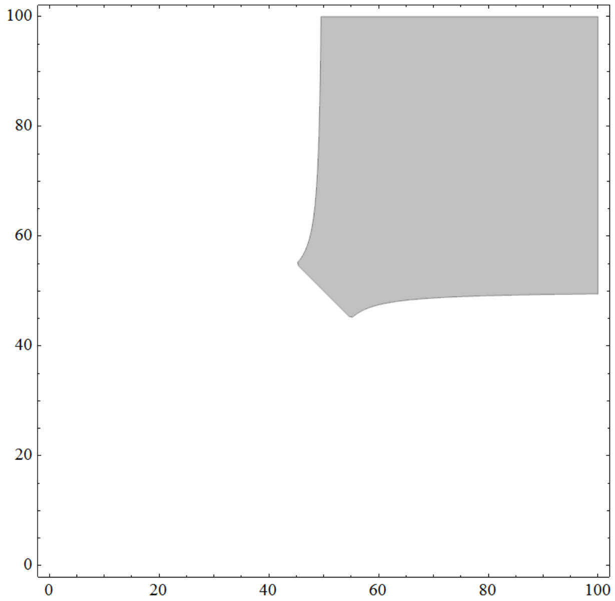

Figure 3.1: The solution set for 2-evaluations with $n = 100$. Horizontal and vertical axes represent r and r' respectively. It can be seen that for 2-evaluations $\mathscr{F}_r^n \cup \mathscr{F}_{r'}^n$ can be distinguished from $\mathscr{F}_0^n$ for values located in the upper right quarter. From this, at least one can deduce that $\mathscr{F}_r^n$ can be distinguished from $\mathscr{F}_1^n$ for sufficiently large r $\left(r >\approx 50 = \frac{n}{2}\right)$.

Case $p = 3$ is more difficult. There are three extremal solutions for given r

$$\begin{aligned} \alpha(\mathbf{r})^{(1)} &= \left(1 - \frac{n}{2r}, \frac{1}{2r}, 0, 0\right), && (3.39) \\ \alpha(\mathbf{r})^{(2)} &= \left(1 - \frac{n(n-1)}{4r(n-r)}, 0, \frac{1}{2r(n-r)}, 0\right), && (3.40) \\ \alpha(\mathbf{r})^{(3)} &= \left(1 - \frac{n(n-1)(n-2)}{2r\beta}, 0, 0, \frac{3}{r\beta}\right), && (3.41) \end{aligned}$$

where $\beta = 3(n-r)(n-r-1) + (r-1)(r-2)$. The two solutions for different values of r and r' will be convex combinations of these vectors

$$\alpha = \sum \lambda^{(i)} \alpha(\mathbf{r})^{(i)} = \sum \mu^{(j)} \alpha(\mathbf{r}')^{(j)} \tag{3.42}$$

where $0 \leq \lambda^{(i)}, \mu^{(j)} \leq 1$ and they add up to 1. Numerically we can evaluate $\lambda^{(i)}$ and $\mu^{(i)}$ for given n, r and r', thus solve the equation for the coefficients α_l, however we couldn't find an analytical expression for these coefficients. From these examples we can see that it would probably be a difficult problem to solve for p-shot in general. Numerical results from $p = 2$ to $p = 5$ is shown in Fig. 3.2. In this figure, we can see that increasing p increases the distinguishability, but with a decreasing rate.

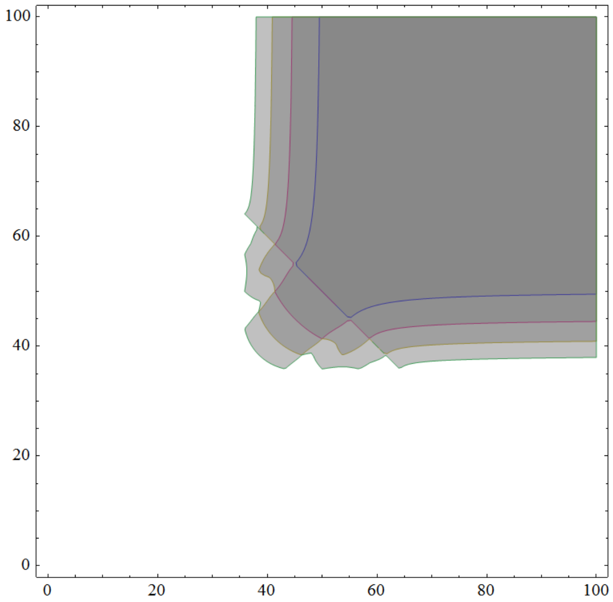

Figure 3.2: The solution set for 2 to 5 evaluations with $n = 100$. Horizontal and vertical axes represent r and r' respectively.

3.2 Density Matrix Correspondence

In the previous section, we considered only a number of parallel calls to the oracle. Nevertheless we are not restricted to parallel calls in general. In fact, any protocol consisting of only parallel calls or only serial calls or any mixture of parallel and serial calls require the same number of oracle invocations for the optimal scenario [21, 22]. In this section, we show that any protocol with any combination of parallel and serial calls can be uniquely represented by a density matrix operating on a Hilbert space with a much higher dimension. This will enable us to decide if the set discrimination problem is solvable without dealing with the design of the protocol.

3.2.1 General form of the problem

Let X be an n-level system with the standard orthonormal basis $\{|1\rangle_X, |2\rangle_X, \ldots, |n\rangle_X\}$ which spans the Hilbert space $\mathcal{H}_X$. We model the black box realizing the function f with a unitary U_f on X such that it acts on the standard basis vectors as

$$U_f \left|i\right\rangle = f_i \left|i\right\rangle , \tag{3.43}$$

where f_i are complex numbers with unit length: $|f_i| = 1$, thus, f_i are eigenvalues of U_f. We now define f_i directly as the eigenvalues of U_f[1]. We can define vector $\left|f\right\rangle$ in $\mathcal{H}_X$ by

$$\left|f\right\rangle \equiv \sum_{i=1}^{n} f_i \left|i\right\rangle . \tag{3.44}$$

Hence, if we call $\left|u\right\rangle \equiv \sum_i^n \left|i\right\rangle = \left|1\right\rangle + \left|2\right\rangle + \cdots + \left|n\right\rangle$, then $\left|f\right\rangle = U_f \left|u\right\rangle$.

Let $\mathscr{F}$ be a set of mutually commuting unitaries. Let $\mathscr{G} \subset \mathscr{F}$ and $\mathscr{G}' \subset \mathscr{F}$ be two disjoint subsets of $\mathscr{F}$. Let a black box carry out a unitary U_f which is either in $\mathscr{G}$ or $\mathscr{G}'$. By using the processor at most p-times can we decide with certainty whether $U_f \in \mathscr{G}$ or $U_f \in \mathscr{G}'$?

3.2.2 Protocol

Let A be an ancilla without any constraint on the dimension of the Hilbert space $\mathcal{H}_A$ associated with it. A p-evaluation protocol consist of an initial state

$$\left|\varphi\right\rangle_{XA} = \sum_i \left|i\right\rangle_X \otimes \left|\varphi_i\right\rangle_A \tag{3.45}$$

where

$$\sum_i^n \| \left|\varphi_i\right\rangle \|^2 = 1 \tag{3.46}$$

by normalization. Let

$$\mathbb{V}^{(k)}, \quad (k = 1, \ldots, p-1) \tag{3.47}$$

be $p-1$ unitaries on $\mathcal{H}_{XA}$. The sequence of the protocol is given in Algorithm 1.

[1] Remember that, we were choosing binary $f_i \in \{0, 1\}$ and defined the unitary action as in Eq. 3.3 which gave rise to an effective action $U_f \left|x\right\rangle \left|\pm\right\rangle = (\pm 1)^{f_i(x)} \left|x\right\rangle \left|\pm\right\rangle$ in the previous section.

Algorithm 1 A p-evaluation protocol.

```
0          The system $XA$ is prepared in the state $|\varphi\rangle_{XA}$
1          Processor evaluates $U_f$ on $X$
1'         The unitary $\mathbb{V}^{(1)}$ is applied on $XA$
2          Processor evaluates $U_f$ on $X$
2'.        The unitary $\mathbb{V}^{(2)}$ is applied on $XA$
           ⋮
(p − 1)    Processor evaluates $U_f$ on $X$
(p − 1)'   The unitary $\mathbb{V}^{(p-1)}$ is applied on $XA$
p          Processor evaluates $U_f$ on $X$
           Final state is called $|\Phi_f\rangle$.
```

Thus for any given protocol with the unitary U_f the final state $|\Phi_f\rangle$ becomes

$$|\Phi_f\rangle \equiv \mathbb{U}_f \mathbb{V}^{(p-1)} \mathbb{U}_f \mathbb{V}^{(p-2)} \cdots \mathbb{V}^1 \mathbb{U}_f \, |\varphi\rangle \,, \tag{3.48}$$

where $\mathbb{U}_f \equiv (U_f)_X \otimes \mathbb{1}_A$. We can define an overlap matrix for the protocol by

$$S_{fg} \equiv \langle \Phi_f | \Phi_g \rangle \,. \tag{3.49}$$

The protocol discriminates between $\mathscr{G}$ and $\mathscr{G}'$, if $S_{ff'} = 0$ for every $U_f \in \mathscr{G}$ and $U_{f'} \in \mathscr{G}'$.

3.2.3 Correspondence

For any isometry $\mathbb{V}$ on XA, define these n^2 block operators V_{ij} on A by

$$\mathbb{V} \, |i\rangle_X \otimes |\psi\rangle_A = \sum_{j=1}^{n} |j\rangle_X \otimes \left(V_{ji} \, |\psi\rangle\right)_A . \tag{3.50}$$

Here $\mathbb{V} : H_{XA} \to \mathcal{H}_{XA}$ but $V_{ji} : \mathcal{H}_A \to \mathcal{H}_A$. One can also define block operators with

$$V_{ij} = \langle j | \mathbb{V} | i \rangle \,. \tag{3.51}$$

Thus we have

$$\mathbb{V} = \sum_{i,j=1}^{n} \left(|j\rangle \langle i|\right)_X \otimes \left(V_{ji}\right)_A . \tag{3.52}$$

Consider the Hilbert space of p-copies of X: $(\mathcal{H}_X)^{\otimes p}$. The standard basis is

$$|i_p i_{p-1} \dots i_1\rangle = |i_p\rangle_{X_p} \otimes |i_{p-1}\rangle_{X_{p-1}} \otimes \cdots \otimes |i_1\rangle_{X_1} \tag{3.53}$$

where $i_1, i_2, \dots, i_p = 1, \dots, n$. Let $\rho^{(p)}$ be a mixed state on $(\mathcal{H}_X)^{\otimes p}$. We define $\rho^{(p-1)}, \rho^{(p-2)}, \dots, \rho^{(1)}$ as reduced density matrices as follows:

$$\rho^{(k)} = \rho^{(k)}_{X_k X_{k-1} \dots X_1} \equiv \mathrm{Tr}_{X_p X_{p-1} \dots X_{k+1}} \rho^{(p)}. \tag{3.54}$$

We say "$\rho^{(k)}$ is diagonal in X_k" if $\left\langle i_k i_{k-1} \dots i_1 | \rho^{(k)} | i'_k i'_{k-1} \dots i'_1 \right\rangle = 0$ when $i_k \neq i'_k$. Thus

$$\rho^{(k)} = \sum_{i=1}^{n} \left(|i\rangle\langle i|\right)_{X_k} \otimes (F_i)_{X_{k-1} X_{k-2} \dots X_1} \tag{3.55}$$

in such a case.

Theorem 1 *There is a one to one correspondence between p-evaluation protocols and density matrices $\rho^{(p)}$ on $(\mathcal{H}_X)^{\otimes p}$ which satisfy the condition that $\rho^{(k)}$ are diagonal in X_k for all $k = 1, 2, \cdots, p$; so that*

$$S_{fg} = \langle \Phi_f | \Phi_g \rangle = \left\langle f^{\otimes p} | \rho^{(p)} | g^{\otimes p} \right\rangle. \tag{3.56}$$

Note that $\langle \Phi_f | \Phi_f \rangle = 1 \quad \forall f$. The condition that $\rho^{(k)}$ is diagonal in X_k is necessary for this relation to hold true. The idea is

$$\left\langle f \otimes \psi | \rho^{(k)} | f \otimes \varphi \right\rangle = \left\langle \psi | \rho^{(k-1)} | \varphi \right\rangle \tag{3.57}$$

for any $|\psi\rangle, |\varphi\rangle \in (\mathcal{H}_X)^{\otimes (k-1)}$ and any f with $|f_i| = 1$ for all i. As a result of this

$$\begin{aligned} \left\langle f^{\otimes k} | \rho^{(k)} | f^{\otimes k} \right\rangle &= \left\langle f^{\otimes (k-1)} | \rho^{(k-1)} | f^{\otimes (k-1)} \right\rangle & (3.58) \\ &= \quad \vdots & (3.59) \\ &= \left\langle f | \rho^{(1)} | f \right\rangle & (3.60) \\ &= \mathrm{Tr} \rho^{(1)} & (3.61) \\ &= 1. & (3.62) \end{aligned}$$

Proof. Proof of Theorem 1 is given in Appendix A. ■

A useful aspect of Thm 1 is the following. Suppose that there are two possible protocols that solve the problem. By Thm 1, there are two density matrices ρ and ρ'satisfying the diagonality requirement and Eq. 3.56. Since Eq. 3.56 is linear in ρ, we consequently reach the conclusion that any mixture $\tilde{\rho} = \lambda\rho + (1-\lambda)\,\rho'$ $(0 \leq \lambda \leq 1)$ of these two density matrices is also a solution to the problem. Involving Thm 1, once again, we can infer the existence of new protocols (a protocol for each value of λ) corresponding to $\tilde{\rho}$. Showing the existence of the new protocols is very trivial if one uses the corresponding density matrices but accomplishing the same task would be very complicated if one

This approach is useful in the cases where the problem has symmetry. In this case, we can choose a symmetric density matrix thus work with protocols that utilize symmetry maximally. Some examples are shown below.

3.2.4 Parallel schemes corresponds to diagonal density matrices

Let the initial state be $|\varphi\rangle = \sum_{i_1\ldots i_p} c_{i_1\ldots i_p} |i_p\rangle \otimes |i_{p-1}\rangle \otimes \cdots \otimes |i_1\rangle \in (\mathcal{H}_X)^{\otimes p}$. Then the final state is evaluated as

$$|\Phi_f\rangle = (U_f)^{\otimes p}|\varphi\rangle = \sum_{i_1\ldots i_p} c_{i_1\ldots i_p} f_{i_1}\ldots f_{i_p}|i_p\ldots i_1\rangle \tag{3.63}$$

which gives rise to an inner product

$$\begin{aligned}\langle\Phi_g|\Phi_f\rangle &= \sum_{i_1\ldots i_p}\left|c_{i_1\ldots i_p}\right|^2\left(g_{i_1}\ldots g_{i_p}\right)^*\left(f_{i_1}\ldots f_{i_p}\right) && (3.64)\\ &= \left\langle g^{\otimes p}|\rho|f^{\otimes p}\right\rangle && (3.65)\end{aligned}$$

where the density matrix ρ is defined as

$$\rho = \sum_{i_1\ldots i_p}\left|c_{i_1\ldots i_p}\right|^2|i_p\ldots i_1\rangle\langle i_p\ldots i_1|\,. \tag{3.66}$$

Consequently, the protocol can be carried out with a parallel scheme if and only if the corresponding density matrix $\rho^{(p)}$ is diagonal in the standard basis.

3.2.5 Our Problem

Let $\mathscr{F}$ be a set of vectors $f = (f_1, \ldots, f_n)$ such that $f_i = \pm 1$:

$$\mathscr{F} = \{f = (f_1, \ldots, f_n) : f_i \pm 1\}. \tag{3.67}$$

Let us define a subset $\mathscr{F}_r \in \mathscr{F}$ such that

$$\begin{aligned} \mathscr{F}_r &= \left\{f : \sum f_i = n - 2r\right\} & (3.68) \\ &= \{f : \text{exactly } r \text{ components } f_i \text{ are } -1\}, & (3.69) \end{aligned}$$

and the set spanned by the vectors $\left|f^{\otimes k}\right\rangle$ as

$$\mathcal{V}_{r,k} = \operatorname{span}\left\{\left|f^{\otimes k}\right\rangle : f \in \mathscr{F}_r\right\}. \tag{3.70}$$

Now we can state our problem: Can we discriminate between $\mathscr{F}_r$ and $\mathscr{F}_{r'}$ by a p-evaluation protocol? Using Theorem 1 we can immediately restate this as follows. Can we find a density matrix ρ in $(\mathcal{H}_X)^{\otimes p}$ satisfying the diagonality requirement stated in Thm 1, in such a way that $\langle\psi|\rho|\psi'\rangle = 0$ for any $\psi \in V_{r,p}$ and $\psi \in V_{r',p}$, or equivalently is $\sqrt{\rho} V_{r,p}$ orthogonal to $\sqrt{\rho} V_{r',p}$?

3.2.6 Two-evaluation protocols

Since we restated the problem in terms of density matrices, let us try to solve it for the simplest non-trivial case. Consider that we are restricted to two evaluations. Then the density matrix corresponding to this protocol will be of the form:

$$\rho_{X_2X_1} = \sum_{ijk} R_{ijk} \left(|i\rangle\langle i|\right)_{X_2} \left(|j\rangle\langle k|\right)_{X_1} = \sum_{ijk} R_{ijk} |ij\rangle\langle ik|. \tag{3.71}$$

Since it is a valid protocol, by Theorem 1, ρ should be diagonal in X_1. By our construction,

$$\mathrm{Tr}_{X_2} \rho_{X_2X_1} = \rho_{X_1} \text{is diagonal.} \tag{3.72}$$

ρ is a density matrix, so it is positive semidefinite, also second trace should give one:

$$\rho = \rho_{X_2X_1} \geq 0, \tag{3.73}$$

$$\mathrm{Tr}\rho_{X_1} = 1. \tag{3.74}$$

Let $\mathscr{P}$ be a permutation of $\{1, 2, \ldots, n\}$ and let $\mathscr{P}\,|i\rangle \equiv |\mathscr{P}_i\rangle$. We can see that, if $\rho_{X_2X_1}$ solves the discrimination problem, then

$$\rho'_{X_1X_2} = \mathscr{P}^\dagger \otimes \mathscr{P}^\dagger \rho_{X_1X_2} \mathscr{P} \otimes \mathscr{P} \tag{3.75}$$

also solves the same problem and has the required diagonality property. As a result,

$$\tilde{\rho}_{X_1X_2} = \sum_{\mathscr{P}} \mathscr{P}^\dagger \otimes \mathscr{P}^\dagger \rho_{X_1X_2} \mathscr{P} \otimes \mathscr{P} \tag{3.76}$$

also solves the same problem but it is also symmetric under permutations. Hence, we can assume that $\rho_{X_2X_1}$ is permutation symmetric

$$\rho_{X_1X_2} = \sum_{\mathscr{P}} \mathscr{P}^\dagger \otimes \mathscr{P}^\dagger \rho_{X_1X_2} \mathscr{P} \otimes \mathscr{P}. \tag{3.77}$$

Similarly, the complex conjugated density matrix $\rho^*_{X_2X_1}$ also solves the same problem and hence

$$\hat{\rho}_{X_2X_1} = \frac{1}{2}\left(\rho_{X_2X_1} + \rho^*_{X_2X_1}\right) \tag{3.78}$$

does too, i.e., there is a permutationally symmetric density matrix with real entries that solves the problem.

First we observe that some of the R_{ijk} values are not different from each other because of the permutational symmetry given in Eq. (3.75). For example $R_{123} = R_{135} = R_{236} = \ldots$, $R_{112} = R_{338} = \ldots$, ... So the general form of R_{ijk} is

$$R_{ijk} = \alpha + \beta_{12}\delta_{ij} + \beta_{13}\delta_{ik} + \beta_{23}\delta_{jk} + \chi_{123}\delta_{ij}\delta_{ik}, \tag{3.79}$$

where by Eq. (3.78), $\alpha, \beta_{12}, \beta_{12}, \beta_{12}, \chi_{123}$ can be chosen real. Next, Hermiticity of ρ gives

$$\beta_{13} = \beta_{12}. \tag{3.80}$$

We have the density matrix of system X_1 as

$$\rho_{X_1} = \mathrm{Tr}\rho_{X_2X_1} = \sum_{jk} R'_{jk}\,|j\rangle\,\langle k|\,, \tag{3.81}$$

where

$$R'_{jk} = \sum_i R_{ijk} = (n\alpha + 2\beta_{12}) + (n\beta_{23} + \chi_{123})\,\delta_{jk}. \tag{3.82}$$

For ρ_{X_1} to be diagonal we should have

$$n\alpha + 2\beta_{12} = 0. \tag{3.83}$$

Also trace of the partial density matrix must be equal to 1

$$\mathrm{Tr}\rho_{X_1} = \sum_j R'_{jj} = n\left(n\beta_{23} + \chi_{123}\right) = 1. \tag{3.84}$$

Next we use the positive definiteness property and by Sylvester criterion we find

$$\begin{aligned} \beta_{23} &\geq 0 && (3.85)\\ \beta_{23} + (n-1)\alpha &\geq 0 && (3.86)\\ \beta_{23}^2 + \beta_{23}\chi_{123} + (n-1)\left(\chi_{123}\alpha - \beta_{12}^2\right) &\geq 0. && (3.87) \end{aligned}$$

Up to this point, we preferred a general treatment and didn't make use of the condition specific to our problem, which was stated in Subsection 3.2.5. Hereafter we continue with the restrictions that comes with our choice of the problem, also what we will call the *weight decision problem.* We can formulate it as distinguishing two sets $\mathscr{F}_r^n$ and $\mathscr{F}_{r'}^n$ with different number of roots r and r' for the time. In the next section we will introduce weights of the functions as the ratio of roots to total number of inputs weights, for instance, the weight of a function from the set $\mathscr{F}_r^n$ will be $\rho = \frac{r}{n}$. However this form of the problem is difficult to make use of in the current formalism, so, before that, let us continue with an easier but related problem: the discrimination of no roots ($\mathscr{F}_0^n$) and r roots ($\mathscr{F}_r^n$). We restricted the protocol to two evaluations at most so we have

$$\langle g \otimes g|\rho|f \otimes f\rangle = 0 \text{ for all } f \in \mathscr{F}_r \text{ and } g \in \mathscr{F}_0, \tag{3.88}$$

where $\mathscr{F}_0$ is the set of function without any non-zero outputs. Then it follows that

$$\sum_{ijk} g_i g_j R_{ijk} f_i f_k = \sum_{ijk} R_{ijk} f_i f_k = 0. \tag{3.89}$$

The sum over j can be evaluated separately in Eq. (3.89). Let us define a new matrix T_{ik} as follows:

$$\begin{aligned} T_{ik} \equiv \sum_j R_{ijk} &= n\alpha + \beta_{12} + \beta_{23} + n\beta_{13}\delta_{ik} + \chi_{123}\delta_{ik} && (3.90)\\ &= \left(n\beta_{13} + \chi_{123}\right)\delta_{ik} + \left(n\alpha + \beta_{12} + \beta_{23}\right). && (3.91) \end{aligned}$$

Then, we can write Eq. (3.89) as

$$\begin{aligned}\sum_{ik} T_{ik} f_i f_k &= (n\beta_{13} + \chi_{123}) \sum f_i^2 + (n\alpha + \beta_{12} + \beta_{23}) \left(\sum f_i\right)^2 && (3.92)\\ &= n\left(n\beta_{13} + \chi_{123}\right) + (n-2r)^2 \left(n\alpha + \beta_{12} + \beta_{23}\right) && (3.93)\\ &= 0. && (3.94)\end{aligned}$$

Making use of Eq. (3.83) and Eq. (3.84) we obtain

$$\begin{aligned} n\chi_{123} + n^2\beta_{13} + (n-2r)^2 \left(n\alpha + \beta_{12} + \beta_{23}\right) &= 0 && (3.95)\\ \left(1 - n^2\beta_{23}\right) + n^2\beta_{13} + (n-2r)^2 \left(\beta_{23} - \beta_{12}\right) &= 0 && (3.96)\\ 1 + \left((n-2r)^2 - n^2\right)\left(\beta_{23} - \beta_{12}\right) &= 0. && (3.97)\end{aligned}$$

Let us denote

$$\kappa = \left((n-2r)^2 - n^2\right) = 4r(n-r),$$

so that we get

$$\beta_{23} - \beta_{12} = \frac{1}{\kappa}. \tag{3.98}$$

So that the inequality given by Eq. (3.86) and Eq. (3.83) leads to

$$\beta_{23} \leq \frac{(n-1)}{(n-2)}\frac{2}{\kappa}. \tag{3.99}$$

Also third inequality given by Eq. (3.87) along with the Equations (3.83), (3.84) and (3.98) leads to

$$\beta_{23} \leq \frac{(n-1)}{(n-2)}\frac{2\kappa - n^2}{\kappa^2}. \tag{3.100}$$

So

$$0 \leq \beta_{23} \leq \frac{(n-1)}{(n-2)}\frac{2\kappa - n^2}{\kappa^2} \tag{3.101}$$

and there is a solution only if $2\kappa - n^2 \geq 0$. Any solution, even $\beta_{23} = 0$ is acceptable, therefore

$$\frac{n^2}{2} \leq \kappa = n^2 - (n-2r)^{\,2} \tag{3.102}$$

which leads to

$$(0.15n \simeq)\, n\frac{\sqrt{2}-1}{2\sqrt{2}} \leq r \leq n\frac{\sqrt{2}+1}{2\sqrt{2}}\,(\simeq 0.85n)\,. \tag{3.103}$$

Before continuing to the case with three evaluation, let us compare this result with the one we obtained from the parallel scheme. In parallel scheme we found that nearly half of the functions could be discriminated from the zero function. On the other hand with the serial scheme we obtain approximately a 70% discrimination rate.

3.2.7 Three-evaluation

The protocol we have given in the Subsection 3.2.6 can be extended to larger number of evaluations however it becomes complicated very rapidly with increasing number of evaluations. In this part we will illustrate this behavior without doing any evaluations since doing so seems to be intractable. First, the density matrix corresponding to a three evaluation protocol will be of the form given in Eq. (3.104) instead of the one in Eq. (3.71)

$$\rho_{X_3X_2X_1} = \sum_{ijklm} Q_{ijklm} \left|ijk\right\rangle \left\langle ilm\right| . \tag{3.104}$$

Like we did in previous analysis, permutational symmetry will dictate the form of Q_{ijklm} such that

$$\begin{aligned} Q_{t_1t_2t_3t_4t_5} \quad = \quad & \sum_{t_1t_2t_3t_4t_5=1}^{n} \Bigg(\alpha + \sum_{i<j} \beta_{ij}\delta_{t_it_j} + \sum_{i<j<k} \chi_{ijk}\delta_{t_it_j}\delta_{t_it_k} \\ & + \sum_{i\neq j\neq k\neq l} \epsilon^{ij}_{kl}\delta_{t_it_j}\delta_{t_kt_l} + \sum_{i<j<k<l} \varphi_{ijkl}\delta_{t_it_j}\delta_{t_it_k}\delta_{t_it_l} \\ & + \sum_{i<j\neq k\neq l\neq m} \gamma_{ij}\delta_{t_it_j}\delta_{t_kt_l}\delta_{t_kt_m} + \nu\delta_{t_1t_2}\ldots\delta_{t_1t_5} \Bigg) . \end{aligned} \tag{3.105}$$

There are 52 coefficients in the most general form of Q_{ijklm} when the permutation symmetries are taken care of. From the Hermiticity of the density matrix, we have 20 equalities. $\rho^{(3)}$ is diagonal in X_2 and $\rho^{(2)}$ is diagonal in X_1 thus giving a total of 8 equalities. Finally the trace should be equal to 1, hence $52 - 29 = 23$ independent coefficients remain.

We now run the protocol with three evaluations so we have

$$\langle g \otimes g \otimes g|\rho|f \otimes f \otimes f\rangle = 0 \text{ for all } g \in \mathscr{F}_0 \text{ and } f \in \mathscr{F}_r. \tag{3.106}$$

which is followed by

$$\sum_{ijklm} g_i g_j g_j Q_{ijklm} f_i f_l f_m = \sum_{ijklm} Q_{ijklm} f_i f_l f_m = 0. \tag{3.107}$$

We again take the sum that does not depend on f_i's first and obtain the matrix elements T_{ilm}:

$$T_{ilm} \equiv \sum_{jk} Q_{ijklm}. \tag{3.108}$$

So that the sum gives

$$\begin{aligned} \sum_{ijklm} T_{ilm} f_i f_l f_m &= (n-2r)\, T_{111} + (n-2r)^3\, T_{123} \\ &\quad + n\,(n-2r)\,(T_{112} + T_{121} + T_{122}) = 0. \end{aligned} \tag{3.109}$$

Theoretically this will decrease the number of equations one more, however, this correction is not as simple as the other constraints. Even if we can make use of this last constraint, a method of exploiting positive definiteness of the density matrix is no longer clear as it was in the 2-evaluation protocol. Unfortunately we were not able to find a way to take it into account. We could continue with numerical methods however even if we could succeed, four and more evaluations would surely be unmanageable. Instead we try another method which we will cover in the next chapter.

CHAPTER 4

ALGORITHMS USING GROVER ITERATION

4.1 Grover's Algorithm

Suppose we have a set of N items identified by the numbers from 0 to $N-1$. Let us have a selector such that it accepts only one of the items and rejects the others. However we are sure that by no means we can inspect this selector and request which item will be marked in the end directly. This can formally defined by a class of Boolean functions $f : \{0, 1, \ldots, N-1\} \rightarrow \{0, 1\}$ such that preimage of $\{1\}$ is a set with only one element, $\{s\}$; $0 \leq s \leq N$, for example. Obviously the best one can do classically is to search the element exhaustively and it requires $\frac{N+1}{2}$ calls on average or the query complexity of this process is of the order $O(N)$. Grover constructed a quantum algorithm that solves this problem and it requires only $O\left(\sqrt{N}\right)$ function calls [18]. Even though it does not led to a logarithmic speedup like Shor's algorithm, its wide range of applicability makes Grover's algorithm promising on a large scale.

Grover's algorithm is based on amplification of the amplitudes of the sought after states by applying the same iteration until the amplitude of the solution state is close to one. Naturally this observation leads us to a more general class of algorithms that use amplitude amplification. In this chapter we will be covering some remarkable examples of these, indeed our original contribution is also an example based on Grover iteration and can be categorized into the class of algorithms that is based on amplitude amplification. In Algorithm 2 the flow of Grover's algorithm is given.

Algorithm 2 Grover's algorithm

1. Prepare $\log N$ bit quantum registers in an equal superposition of states of each possible input. This can be done with the application of an n bit Hadamard transform on the zero state. N can be chosen as a power of 2 or the problem can be adjusted to this choice without any complications. Let $N = 2^n$.

$$|\psi_1\rangle = H^{\otimes n} |0\rangle^{\otimes n} = \sum_x \frac{1}{\sqrt{N}} |x\rangle \equiv |\Sigma\rangle \tag{4.1}$$

2. Repeat the Grover iteration $\frac{\pi}{4}\sqrt{N}$ times. Grover iteration consists of two steps.

    (a) Inversion about the solution state. Its operation can be written as the unitary operator $U_f = 2|s\rangle\langle s| - \mathbb{1}$.

    (b) $(-)$ Inversion about the mean state $|\Sigma\rangle = \sum_{i=1}^{N} |i\rangle$. It can be given by the unitary operator $U_\Sigma = \mathbb{1} - 2|\Sigma\rangle\langle\Sigma|$.

$$|\psi_t\rangle = (U_\Sigma U_f)^t |\psi_1\rangle, \tag{4.2}$$

    where $t \approx \frac{\pi}{4}\sqrt{N}$

3. Perform a measurement on the final state. The measurement result will give the solution s with a small probability of error.

Note that in the second step of Algorithm 2, the operations U_f and U_Σ are given in a mathematical form rather than an operational layout that reflect their physical realizations. In fact, the inversion about the mean operator is realized with two Hadamard transformations and an inversion about the standard state operator so that their successive applications give

$$U_\Sigma = H^{\otimes n} \left(\mathbb{1} - 2 \left| 0 \right\rangle \left\langle 0 \right| \right) H^{\otimes n}. \tag{4.3}$$

The inversion about the solution state U_f is also realized with the help of an ancilla qubit such that its action on the full state and its effect on the work qubit are given in Eq. 4.4 and Eq. 4.5. Note that these definitions are the same as the ones in the previous chapter, specifically Eq. 3.3 and Eq. 3.4

$$U_f \left| x \right\rangle \left| y \right\rangle = \left| x \right\rangle \left| f(x) \oplus y \right\rangle, \tag{4.4}$$

$$U_f \left| x \right\rangle \left| \pm \right\rangle = (\pm 1)^{f(x)} \left| x \right\rangle \left| \pm \right\rangle. \tag{4.5}$$

In Nielsen and Chuang's book a beautiful geometrical visualization is given[1]. We redrew it in Fig. 4.1 to facilitate better understanding of how the algorithm works. We start with depicting the initial state $\left| \Sigma \right\rangle$ as a vector that is a linear combination of the normalized solution and non-solution basis states, which are denoted as $\left| s \right\rangle$ and $\left| ns \right\rangle$ respectively. In each iteration, the state is first reflected about the solution state by the oracle operator U_f which selectively inverts the phases of the solution states and then the resulting state is reflected about the superposition state Σ by the other part of the iterator, namely, U_Σ. The total effect of these two reflections is equal to a rotation with an angle $\theta = 2 \sin^{-1} \frac{1}{\sqrt{N}}$ in the $\left| s \right\rangle$ - $\left| ns \right\rangle$ basis. Thus, each iteration rotates the state by the same amount θ. Eventually when the state is rotated enough, that is after $\frac{\pi}{4}\sqrt{N}$ iterations, it becomes hardly distinguishable from the solution state. At this point a measurement in the standard basis give the solution state with maximum success probability.

One can generalize the two iterations such that a single Grover iteration consisting of a Hadamard transform and overall or selective phase inversions are given

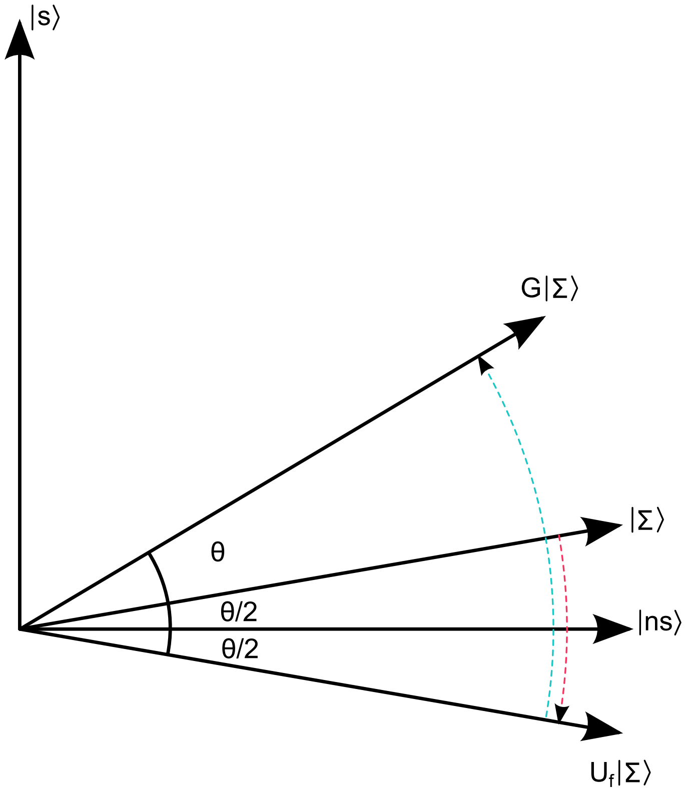

Figure 4.1: Grover's algorithm visualized. The initial state can always be written as a linear combination of the normalized superposition of solution states $|s\rangle$, and non-solution states $|ns\rangle$. In each iteration the state of the computer is rotated by an angle $\theta = 2\sin^{-1}\frac{1}{\sqrt{N}}$ towards the solution state. After $O(\sqrt{N})$ rotations, the state of the computer is the least distinguishable from the superposition of solution states.

in the form

$$G = -\Upsilon I_{|0\rangle}\left(\phi\right)\Upsilon^{-1} I_{|s\rangle}\left(\varphi\right) \tag{4.6}$$

where Υ is a unitary operator and $I_{|\psi\rangle}\left(\alpha\right)$ denotes the inversion operator by an angle α so that

$$I_{|\psi\rangle}\left(\alpha\right) \equiv \mathbb{1} - \left(1 - e^{i\alpha}\right)|\psi\rangle\langle\psi| \tag{4.7}$$

With the choice of $\Upsilon = H^{\otimes n}$ and $\phi = \varphi = \pi$ we recover the original algorithm by Grover.

Grover's algorithm can be applied without any modification to the same problem with more than one solutions. It is straightforward to show that the algorithm runs r times faster if there are r solutions instead of one but it gives only one of the solutions randomly at the output thus recovering the same complexity[67]. This is one example of a number of generalizations of Grover's algorithm. Another generalization is using an arbitrary unitary instead of the Hadamard transform given in Eq. 4.3. In a later work, Grover showed that almost any unitary can be used instead of the Hadamard transfom without changing the $O\left(\sqrt{N}\right)$ complexity as long as it is used consistently [68]. Grover's algorithm is unique among the major algorithms in the sense that it gives a quadratic speedup even though there is no structure imposed on the search space. However if there is an inner structure to be exploited, quadratic speedup for this new search space can be recovered by using appropriate heuristics for the problem[23]. The initial amplitude distribution and phase inversion angle can also be chosen different than the original. These versions of the algorithm were studied in papers by Biham *et al.*[69, 70]. Even though the original algorithm is probabilistic except for the case $N = 4$[71], with several methods it can be converted to a sure success algorithm. Brassard *et al.* accomplished this by applying the standard iteration up until the last step but changing the final step with a smaller step size[72]. Høyer designed an algorithm that adjusts the inversion phase angle in such a way that the rotations add up to a transformation exactly up to the solution state[73]. Long fine tuned the amplitude amplification operator with only one adjustable phase [74], and his algorithm is superior to the others in such a way that a single change in the phase which can be obtained in a simple closed form is enough to obtain 100% success probability[75]. Approaches by Høyer and Long can be combined in a more general form using an $SU(2)$ representation involving additional phases which is shown by Hsieh and Li in [76].

A natural question concerning performance is whether Grover's algorithm provide the best speedup among other possible quantum algorithms or is there a room for further improvement for quantum algorithms that solve unstructured database search problem. Unfortunately Grover's results are shown to be the upper limit for unstructured database search. It has been proven asmptotically by Bennett *et al.* and Boyer *et al.* [77, 67] and the exact proof is given by Zalka[78]. Another feature of the Grover's algorithm is that it is not a fixed point algorithm, that is, it does not monotonically converge to a solution. If we apply more iterations than we should, the state of the computer starts to become less distinguishable than the solution state again. This can be easily seen in the geometric picture that we drew in Fig. 4.1. The state of the computer rotates indefinitely if we continue to apply the iterations. As we will see in the next chapter, this property will become useful when the problem is not finding a solution but finding the number of solutions. Concerning the fixed point property, it has been shown that Grover's algorithm can be altered to become a fixed point algorithm[79].

4.2 Quantum Counting

In the generalizations we mentioned in the Section 4.1 the number of solutions to the reverse problem $f^{-1}(1) =?$ was known beforehand and we were trying to obtain those solutions. Suppose that we are not interested in the solutions themselves but only in the number of the solutions, that is now an unknown. This is known as quantum counting and it is first studied by Brassard *et al.*[23]. In this section we will repeat most of their results without giving proofs. Let us give a more formal definition of the problem: given a boolean function $f : \{0, \ldots, N-1\} \rightarrow \{0, 1\}$; find the cardinality of the inverse of the preimage $r \equiv |f^{-1}(1)|$ where $|\,|$ denotes the cardinality of a set. The insight to solve this problem comes from the observation that we made earlier. The application of Grover iteration cause an arbitrary amount of rotation in the state of the quantum computer (see Fig. 4.1). This amount depends on the set size as well as the number of solutions to the problem. If the state is rotated enough, we expect it to end up very close to the same amplitudes that was achieved before in a periodical fashion. Since we know the set size, learning the period or frequency will inform us about the number of solutions. In order to obtain such information we can employ Fourier analysis. The quantum algorithm **Count**(f, P) given in Algorithm 3 shows how this can be achieved with bounded error. The algorithm accepts two parameters: the Boolean function f whose number of roots we try to determine and an integer P we introduce to control the number of runs as well as precision. P can be assumed to be a power of 2 like N. The implementation of Grover iterations along with the quantum Fourier transform is carried out using two unitaries:

$$C_f \left|m\right\rangle \otimes \left|\psi\right\rangle = \left|m\right\rangle \otimes (G_f)^m \left|\psi\right\rangle, \tag{4.8}$$

$$F_P \left|k\right\rangle = \frac{1}{\sqrt{P}} \sum_{l=0}^{P-1} e^{2\pi ikl/P} \left|l\right\rangle. \tag{4.9}$$

Here G_f is the original Grover iteration with Hadamard transforms and phase inversions equal to π. Notice that C_f takes the parameter m that is recorded in the first register and apply that many Grover iterations to the state kept in the second register. The second operator F_P is the usual quantum Fourier transform operator[1]. The parameter P becomes an upper limit for m thus running C_f along with F_P becomes possible. The flow of **Count**(f, P) is given in Algorithm 3.

Brassard *et al.* proved that the function **Count**(f, P) finds an approximate number of roots $\tilde{r}$ that the maximum it can differ from the actual number of roots r by

$$|r - \tilde{r}| < \frac{2\pi}{P}\sqrt{rN} + \frac{\pi^2}{P^2}N \tag{4.13}$$

with probability at least $\frac{8}{\pi^2}$. Here $P > 4$. We would like to call **Count**(f, P) with $P = c\sqrt{\frac{N}{r}}$ to obtain an estimation on r with constant relative error. Ironi-

Algorithm 3 Function **Count**(f, P). It accepts two parameters, function f of interest which is implemented as a black box and an integer P to determine the precision of the algorithm. It returns an approximate number of roots $\tilde{r}$. P can be taken as a power of 2.

1. Initialize two quantum registers in the equal amplitude superposition of all states. This can be done the same way as we did in Algorithm 2.

$$|\psi_1\rangle = H \otimes H \left(|0\rangle |0\rangle\right) = \sum_{i,j} \frac{1}{N} |i\rangle |j\rangle \tag{4.10}$$

2. Apply the multiple Grover iteration operator C_F

$$|\psi_2\rangle = C_F |\psi_1\rangle . \tag{4.11}$$

3. Apply the quantum Fourier transform on the first register of the resulting state

$$|\psi_3\rangle = (F_P \otimes \mathbb{1}) |\psi\rangle . \tag{4.12}$$

4. Perform a measurement on the final state $|\psi_3\rangle$. Suppose that you obtain $\tilde{q}$. If $\tilde{q} > P/2$ then $\tilde{q} \leftarrow (P - \tilde{q})$.

5. Output $\tilde{r} = N \sin^2 \left(\tilde{t}\pi/P\right)$ and $\tilde{q}$ if necessary.

cally, we are trying to determine the number of roots r itself with this process. To overcome this dilemma we introduce an adaptive procedure given in Algorithm 4 that starts with the roughest estimation and increases the precision P until it is acceptable.

Algorithm 4 Function **CountRel**(f, c). It has two parameters: function f and constant c that limits error probability.

```
1. $P \leftarrow 2$
2. While $\left(\tilde{f} < 1\right)$
   {
   $P \leftarrow 2P$
   $\tilde{f} \leftarrow$ Maj($\Omega\left(\log\log N\right)$),Count($f, P$))
   }
3. Output Count($f, cP$)
```

Maj$(k,$**Count**$)$, in Algorithm **CountRel()**, denotes the majority vote of k runs of algorithm **Count**. For a problem of size N and function f with number of roots r, an estimate $\tilde{r}$ is generated by **CountRel**(f, c) such that

$$|r - \tilde{r}| < \frac{r}{c}. \tag{4.14}$$

with a probability larger than $\frac{3}{4}$, using $\Theta\left((c + \log\log N)\sqrt{N/r}\right)$ evaluations of f [23]. Best we could do classically is testing the function for random inputs. In that case we would need approximately $O(N/r)$ runs. In summary, by making use of Grover iteration as well as quantum Fourier transform, a quantum algorithm counting the number of roots of a Boolean function can be constructed. Please note again that the algorithms and results in this section are due to the work of Brassard et al.[23] and no original work of the author of this thesis is included for this part.

4.3 Weight Decision Problem

As we have seen in the previous section, periodical nature of repeated applications of Grover's iterations allowed us construct a quantum algorithm to extract the weight of a Boolean function faster than a classical algorithm. By weight of a function f we mean the ratio of number of inputs r for which outputs of f are 1 over the number of all possible inputs N, in other words if the weight of a function f is ρ then $\rho = \frac{r}{N}$ where r is the number of roots of the equation $f(x) = 1$. What if we are sure that the weight of given function is one of the possible known values let us say ρ_1 and ρ_2 and we try to determine the correct weight with minimum number of calls to the function. This problem is known as the weight decision problem and has many applications. Some examples can be given from cryptanalysis[80], coding theory[81], built-in self-testing circuits[82] and fault-tolerant circuit design[83].

The first application of Grover iteration to weight decision problem is realized by Braunstein *et al.* and Choi and Braunstein in a series of papers. They first consider a "symmetric" case of weight discrimination where there is a complementarity condition on weights such that they add up to one[24]. Then they remove the restriction and give a solution to the general problem of "asymmetric" weights and multiple weights[25] and also gave a formal proof of the asymptotic optimality of their results[26]. Their algorithms works perfectly in the sense that there is zero probability of incorrectly distinguishing functions with different weights. The other solution to the same problem is by Uyanik and Turgut [27]. In this work we used a new approach to alter Grover iterations so that the process becomes a sure-success one. Specifically we transformed the decidability problem into a system of algebraic equations. Therefore it is now easier to decide whether two weights are distinguishable or not especially in the small number of iterations regime. Since quantum weight decision problem is a special case of quantum counting, the algorithm by Brassard *et al.* can also be used to discriminate weights, however as one would intuitively expect the two algorithms special to the premise problem are faster than the generic counting problem. Even in that case their speedup are eventually limited to the order of square root [26, 27]. In the remainder of this chapter we will thoroughly review the weight decision problem by first starting with the works by Braunstein, Choi and others given in Ref. [24, 25, 26] and continue with the work given in Ref. [27].

4.3.1 Braunstein and Choi's method

We have seen that successive applications of Grover iterations rotate the state of the system towards the uniform superpositions of solution basis states. This is in fact a neat example of quantum operator discrimination where the operator is applied multiple times sequentially. Successive applications of the same operator on the same initial state lead to final states that are almost distinguishable. Can we adapt this technique to distinguish functions with different weights by making corresponding operators rotate the initial state to mutually orthogonal states? Moreover can we do this with 100% success rate? The answer to both of the questions is "yes we can". However the path from original Grover algorithm to the exact discrimination of weights in general is non-trivial.

4.3.1.1 Symmetric weight decision problem

First let us observe that the initial state in the original Grover's algorithm can be written as

$$\left|\psi_0^{(\rho)}\right\rangle = \sin\frac{\beta_\rho}{2}\left|s\right\rangle + \cos\frac{\beta_\rho}{2}\left|ns\right\rangle \tag{4.15}$$

where $\rho = \sin^2\frac{\beta_\rho}{2}$, with $0 < \beta_\rho \leq \pi$ and the states $|s\rangle$ and $|ns\rangle$ denote the normalized superpositions of equal amplitude solution and nonsolution states, respectively. The standard Grover iteration

$$G = -I_{|\Sigma\rangle}(\pi)\, I_{|s\rangle}(\pi) \tag{4.16}$$

is recovered when the phase angles are chosen as π and the unitary operation as H in Eq. 4.6. The state that is obtained by applying standard Grover iteration k times can be given as

$$\left|\psi_k^{(\rho)}\right\rangle = \sin\left(2k+1\right)\frac{\beta_\rho}{2}\left|s\right\rangle + \cos\left(2k+1\right)\frac{\beta_\rho}{2}\left|ns\right\rangle. \tag{4.17}$$

Here, the original Grover algorithm is only one step of measurement away to obtain one of the solutions with a success probability $\sin\left(2k+1\right)\frac{\beta_\rho}{2}$ however our aim is to rotate the same initial state $\left|\psi_0^{(\rho)}\right\rangle$ to two mutually orthogonal states $\left|\psi_k^{(\rho_1)}\right\rangle$ and $\left|\psi_k^{(\rho_2)}\right\rangle$ corresponding to different weights ρ_1 and ρ_2. Unfortunately, a straightforward application of the standard Grover iteration achieves this only for weights that add up to a number very close to 1 and with bounded error. At this point, we take the exact discrimination condition into consideration. As we have reviewed in the paragraph after Eq. 4.7 there are several options to make Grover iteration exact. In [24], Braunstein *et al.* chooses a method that uses the original Grover iteration until the last step but changes one of the inversion phase angles only for the last step to obtain an exact rotation. Furthermore as the problem requires that the states corresponding to functions with different weights should be correctly rotated to the mutually orthogonal solution and non-solution states, one has to alter the last two steps instead of one. In general these modified phases are different than π and in [24] they are given as

$$\cos\theta_1 = \frac{\left(-1\right)^k\cos\beta_{\rho_1} - \cos 2\beta_{\rho_1}\cos\left(2k-2\right)\beta_{\rho_1}}{\sin 2\beta_{\rho_1}\sin\left(2k-2\right)\beta_{\rho_1}}, \tag{4.18}$$

$$\cos\theta_2 = \frac{\left(-1\right)^k\sin 2\beta_{\rho_1}\left(y\sin\theta_2 - \left(-1\right)^k\sin\beta_{\rho_1}\right)}{\cos\beta_{\rho_1}\cos 2\beta_{\rho_1} - \left(-1\right)^k\cos\left(2k-2\right)\beta_{\rho_1}}. \tag{4.19}$$

Therefore an algorithm that exactly discriminates functions with weights that add up to 1 is present. Braunstein *et al.* calls this symmetric weight decision problem because there is an additional symmetry condition that the sum of weights must be 1. We need this symmetry argument because if the weights are symmetric then on the Bloch sphere, corresponding initial states occupy symmetric locations in the $X-Y$ plane. Hence the Grover iteration rotates them symmetrically and an evolution resulting in orthogonal states is guaranteed to achieve. Detailed flow of the process for an exact discrimination of symmetric weights is given in Algorithm 5.

4.3.1.2 Asymmetric weight decision problem

The symmetry condition in Algorithm 5 is a strict restriction and we would like to avoid it if possible. Choi and Braunstein found an effective method

Algorithm 5 Symmetric weight decision algorithm. The standard Grover iteration $G = -I_{|\Sigma\rangle}(\pi) I_{|s\rangle}(\pi)$ is applied until the last two steps. By a modification in the phase angles in the steps $k-1$ and k, rotation to the orthogonal solution and non-solution states is achieved.

```
1. $\left|\psi_0^{(\rho)}\right\rangle = H^{\otimes n} |0\rangle^{\otimes n} |1\rangle = |\Sigma\rangle |1\rangle$, $i = 0$.
2. If $\rho_1 \leq \sin^2 \frac{\pi}{5}$, $k \leftarrow 2$, else $k$ satisfies $\sin^2\left(\frac{k-1}{2k-1}\frac{\pi}{2}\right) < w_1 \leq \sin^2\left(\frac{k}{2k+1}\frac{\pi}{2}\right)$.
   $\frac{k-1}{2k-1}\pi < \beta_{\rho_1} \leq \frac{k}{2k+1}\pi$. $\beta_{\rho_1} + \beta_{\rho_2} = \pi$.
3. While $(i < k * 2)$
   {
   $\left|\psi_{i+1}^{(\rho)}\right\rangle = -I_{|\Sigma\rangle}(\pi) I_{|s\rangle}(\pi) \left|\psi_0^{(\rho)}\right\rangle$,
   $i \leftarrow i + 1$
   }
4. $\left|\psi_{k-1}^{(\rho)}\right\rangle = -I_{|\Sigma\rangle}(-\theta_1) I_{|s\rangle}(\pi) \left|\psi_{k-2}^{(\rho)}\right\rangle$
5. $\left|\psi_k^{(\rho)}\right\rangle = -I_{|\Sigma\rangle}(-\theta_2) I_{|s\rangle}(\pi) \left|\psi_{k-1}^{(\rho)}\right\rangle$
6. Measure $\left|\psi_k^{(\rho)}\right\rangle$ in the computational basis. Let the result be $\hat{x}$.
7. If $k$ is odd and if $f(\hat{x}) = 1$ then $\rho \leftarrow \rho_1$ else $\rho \leftarrow \rho_2$.
   If $k$ is even and if $f(\hat{x}) = 1$ then $\rho \leftarrow \rho_2$ else $\rho \leftarrow \rho_1$.
```

to generalize the algorithm from symmetric weights to generic weights without changing the run-time but using only 2 qubits of extra space. They achieve such reduction by using extra inputs so that symmetry is recovered for the modified function. Let us briefly describe this adaptation. For a discrimination problem of functions with N inputs we add $3N$ extra inputs. Let ρ_1 and ρ_2 be the weights of the functions f_1 and f_2, which must be discriminated. It means that for exactly r_i out of N different inputs of f_i give 1 and $N - r_i$ inputs of f_i give 0 so that $\rho_i = \frac{r_i}{N}$, where $i = 1, 2$. For the first N inputs the original function remains as it is. Next, we duplicate the function for the next N inputs and for the remaining part we add l inputs that give 1 and $2N - l$ inputs that give 0. The modified function f' is given in Eq. 4.20

$$f'(x) \equiv \begin{cases} f(x) & 0 \leq x < N, \\ f(x-N) & N \leq x < 2N, \\ 1 & 2N \leq x < 2N + l, \\ 0 & 2N + l \leq x < 4N. \end{cases} \tag{4.20}$$

The altered weights for f_1' and f_2' are $\rho_+ = \frac{2\rho_1 + l}{4N}$ and $\rho_- = \frac{2\rho_2 + l}{4N}$. The symmetry should be recovered for the new function thus l is equal to $2N - (\rho_1 + \rho_2)$. The modified weights are therefore $\rho_\pm = \frac{1}{2} \pm (\rho_1 - \rho_2)$. Since l is an integer, Algorithm 5 can be used to distinguish the modified oracles corresponding to the modified functions f_1 and f_2. Therefore asymmetric weight decision problem can be solved on a quantum computer taking the same amount of time but using 2 extra qubits that is used for storing quadruple size of input. Note that if the actual weights are very close to each other, symmetrized weights are close to 0 and 1

4.3.1.3 Multiple weight decision problem

Asymmetric weight decision algorithm can be extended to the multiple weights. In this case the task is to decide which one of the m weights $0 < \rho_1 < \rho_2 < \ldots < \rho_m < 1$ is selected. The adaptation for multiple weight decision is given in Algorithm 6. Here we start with the set S containing all functions with the given weights. Since we can characterize any function with its weight without confusion we labeled them with their weights. We make $m-1$ rounds. In each round we update our information about the largest and smallest weights of the set S. Then we run the asymmetric weight decision algorithm as if we were distinguishing the maximum and minimum weights and discard the false result from the set S. This elimination process will continue until there is only one weight in S. Finally we output the set with only one element as the correct weight. This algorithm operates as intended because once the correct weight is encountered inside the loop, there is no possibility of rejecting it due to the sure-success property of the asymmetric weight decision subroutine. Since we need $m-1$ calls to the oracle, the complexity of the quantum algorithm is $O\left(m\sqrt{N}\right)$. Comparing it with the classical complexity $O(N)$, it becomes advantageous when $m \leq \sqrt{N}$.

Algorithm 6 Multiple weight decision algorithm.

```
1. $S \leftarrow \{\rho_1, \rho_2, \ldots, \rho_m\}$

2. While ($|S| \geq 1$)
   {
   $\rho_{min} \leftarrow$ SmallestWeightOf ($S$)
   $\rho_{max} \leftarrow$ LargestWeightOf ($S$)
   $\{\rho_{selected}, \rho_{notSelected}\}$ = AsymmetricWeightDecision ($\rho_{min}, \rho_{max}$)
   $S \leftarrow S - \{\rho_{notSelected}\}$
   }

3. Return $S$.
```

This concludes the findings of Braunstein *et al.* and Choi and Braunstein's further improvements[24, 25]. We will now continue with our contribution to the subject of weight discrimination.

4.3.2 Our Method

In this subsection we present our method and findings in weight decision problem of Boolean functions. This subsection is an adaptation from our work given in

Ref. [27]. The problem definition is the same as Braunstein and Choi's. We are given two Boolean functions with different weights $\rho = \frac{r}{N}$ and $\rho' = \frac{r'}{N}$. We can only access the function through calling it with different inputs thus it is a black box from the computational perspective. We try to decide which of the functions is the actual one with minimum number of calls to that function. In our work we do not have a distinction of symmetric and asymmetric weight decision hence we use the name *weight decision problem* regardless of such symmetry.

4.3.2.1 Preliminaries

We implement the function with a unitary operator with its action

$$U_f \left|x\right\rangle_I \left|y\right\rangle_R = \left|x\right\rangle_I \left|y \oplus f(x)\right\rangle_R \tag{4.21}$$

as usual. It is designed to make use of its effective action

$$U_f \left|x\pm\right\rangle_{IR} = (\pm 1)^{f(x)} \left|x\pm\right\rangle_{IR}, \tag{4.22}$$

so that the phase is rotated by π radians only when $f(x) = 1$ and the R register contains the state $\left|-\right\rangle$. Here I and R indicate input register and result qubit, respectively. Let us add an ancilla system A, so that $\left|\beta_1\right\rangle, \ldots, \left|\beta_n\right\rangle$ be a set of orthonormal vectors in the state space of the composite system AIR. Note that the total Hilbert space formed by AIR is not necessarily spanned completely by these vectors.

Consider the unitary operator

$$S \equiv \mathbb{1} - 2\sum_{i=1}^{n} \left|\beta_i\right\rangle \left\langle\beta_i\right| \tag{4.23}$$

which makes an inversion in n-dimensional subspace, namely the subspace spanned by $\{\left|\beta_i\right\rangle_i\}$. Equipped with an oracle and an "inversion about mean" like operator we can introduce the generalized Grover iteration

$$Q_f \equiv -S\left(\mathbb{1}_A \otimes U_f\right). \tag{4.24}$$

We now define an $n \times n$ "cosine matrix"

$$\mathbb{C}_{ij} \equiv \langle\beta_i|\mathbb{1}_A \otimes U_f|\beta_j\rangle \tag{4.25}$$

to facilitate the expression of multiple iterations of Q_f on an arbitrary initial state. $\mathbb{C}_{ij}$ is Hermitian and its eigenvalues lie between -1 and $+1$ since the eigenvalues of Hermitian U_f matrices are ± 1. Thus we can consider $\mathbb{C}$ as the cosine of angle matrix Θ so that

$$\mathbb{C} = \cos\Theta. \tag{4.26}$$

We continue with the definition of $n \times n$ matrices $\mathbb{R}^{(m)}$. There are exactly m of them. For any positive integer k, $\mathbb{R}^{(k)}$ is given as

$$\mathbb{R}^{(k)} = \frac{\sin(k\Theta)}{\sin\Theta} \tag{4.27}$$

which is a polynomial function of $\mathbb{C}$. The matrices Θ, $\mathbb{C}$, and $\mathbb{R}$ depend on the function f but the dependence of these matrices on function f such as $\mathbb{R}(f)$ etc. will not be shown when it does not cause any confusion. Now we can present a fundamental result in terms of the matrices $\mathbb{R}^{(m)}$ as:

Theorem 2 *Provided that the initial state is one of the basis states $|\beta_i\rangle$, the final state after m iterations of $Q_f = -S(\mathbb{1} \otimes U_f)$ is*

$$(Q_f)^m |\beta_i\rangle = \sum_{j=1}^{n} \left(|\beta_j\rangle \mathbb{R}_{ji}^{(m+1)} - (\mathbb{1}_A \otimes U_f) |\beta_j\rangle \mathbb{R}_{ji}^{(m)} \right). \tag{4.28}$$

Proof. Theorem 2 can be proved by induction. Verification of Eq. 4.28 for $m = 0$ and $m = 1$ and showing that following recurrence relation is satisfied by the matrices $\mathbb{R}^{(m)}$ is rather straightforward

$$\mathbb{R}^{(m+2)} - 2\mathbb{C}\mathbb{R}^{(m)} + \mathbb{R}^{(m)} = 0. \tag{4.29}$$

■

For the weight decision problem, we prepare the initial state of the composite system as

$$|\beta\rangle \equiv \sum_{i=1}^{n} c_i |\beta_i\rangle$$

where c_i are the amplitudes that we will determine later. We then consider m successive applications of the generalized Grover iteration Q_f so that the final state of the system can be written as

$$|\Phi_f\rangle \equiv (Q_f)^m |\beta\rangle. \tag{4.30}$$

Finally, we carry out a measurement on an appropriate part of the system in order to determine which one of the weights is given. Like Braunstein *et al.* we consider only sure-success discrimination algorithms which means all final states corresponding to the functions with different weights are orthogonal to each other. That is to say, using a bit of notation of Chapter 3, if two functions f and g are chosen from the sets $\mathscr{F}_r^N$ and $\mathscr{F}_{r'}^N$ then m successive applications of Q_f should always lead to orthogonal final states $|\Phi_f\rangle$ and $|\Phi_g\rangle$ such that $\langle\Phi_f|\Phi_g\rangle = 0$. From a computational point of view, in order to obtain the most efficient algorithm we have to minimize the number of oracle calls thus the number of iterations, m.

We can now concentrate on the β_i 's. We make the following choice

$$|\beta_i\rangle = |\beta_i\rangle_{AIR} \equiv |\alpha_i\rangle_A \otimes \left(\sqrt{\mu_i}\frac{1}{\sqrt{N}}\sum_{x=0}^{N-1}|x-\rangle_{IR} + \sqrt{1-\mu_i}\,|0+\rangle_{IR}\right) \tag{4.31}$$

where μ_i are real parameters between 0 and 1. The ancilla states $|\alpha_i\rangle_A$ are mutually orthogonal to each other (i.e.,$\langle\alpha_i|\alpha_j\rangle = \delta_{ij}$). This way we guarantee that $|\beta_i\rangle_{AIR}$ are also normalized and mutually orthogonal to each other.

Suppose that the weight of the function f is ρ. In this case the i^{th} diagonal entry of $\mathbb{C}$ becomes

$$\mathbb{C}_{ii} = \cos\theta_i\left(f\right) = \langle\beta_i|\mathbb{1}_A \otimes U_f|\beta_i\rangle = 1 - 2\rho\mu_i, \tag{4.32}$$

thus we find the upper bound of the eigenvalues

$$0 \leq \theta_i\left(f\right) \leq \cos^{-1}\left(1-2\rho\right). \tag{4.33}$$

From Eq. 4.33, it also follows that $\mathbb{R}^{(m)}$ are also diagonal and their diagonal entry in the i^{th} position is

$$\mathbb{R}_{ii}^{(m)} = \frac{\sin m\theta_i\left(f\right)}{\sin\theta\left(f\right)}. \tag{4.34}$$

4.3.2.2 Discrimination of zero function

We start with a simpler problem by setting one of the weights to zero like we did in Chapter 3. Given that the unknown function is either f with an arbitrary weight ρ or the zero function z which is identically zero for all of its inputs we try to determine which one is the actual weight with using minimum number of function evaluations. We immediately observe that $U_z = \mathbb{1}$, thus application of any number of generalized grover iterations corresponding to zero function does not change the basis states

$$(Q_z)^m |\beta_i\rangle = |\beta_i\rangle . \tag{4.35}$$

For all the functions with weight ρ to be distinguished from the zero function after m iterations, the condition $\langle\Phi_z|\Phi_f\rangle = 0$ should be satisfied. Therefore it follows that

$$\begin{aligned} 0 &= \langle\Phi_z|\Phi_f\rangle = \langle\beta| (Q_f)^m |\beta\rangle && (4.36)\\ &= \sum_{ij} a_i^* \langle\beta_i| (Q_f)^m |\beta_j\rangle a_j && (4.37)\\ &= \sum_{ij} a_i^* \left(\mathbb{R}^{(m+1)}(f) - \mathbb{C}(f)\,\mathbb{R}^{(m)}(f)\right)_{ij} a_j && (4.38)\\ &= \sum_{ij} a_i^* \left(\cos m\Theta(f)\right)_{ij} a_j. && (4.39)\end{aligned}$$

where we made use of Theorem 2 to advance from Eq. 4.37 to Eq. 4.38. Regardless of the number of iterations, Eq. 4.39 is satisfied with $n = 1$. It means that, the subspace where S is an inversion is spanned only by $|\beta_1\rangle$. Consequently, $\Theta(f)$ has only one entry which we may denote by $\theta_1(f)$. Then, we have

$$\cos m\theta_1(f) = 0 \tag{4.40}$$

for all the functions f with weight ρ. This leads to a value of $\theta_1(f) = \pi/2m$ for the fastest algorithm (i.e. minimum m). Note that, the definition of β_i in Eq. 4.31 is given so that, $\theta_1(f)$ depends only on the weight of the function f. Therefore the minimum number of iterations, m, is the smallest integer that satisfies

$$\cos\frac{\pi}{2m} = 1 - 2\rho\mu_1 \tag{4.41}$$

with the condition $0 \leq \mu_1 \leq 1$. Hence the expression for minimum number of iterations can be found as

$$m_{min}(\rho) = \left\lceil \frac{\pi}{2\arccos(1-2\rho)} \right\rceil , \tag{4.42}$$

where $\lceil y \rceil$ denotes the smallest integer not less than y. We can also find a lower bound on zero-distinguishable ρ in terms of iteration number m

$$\rho \geq \rho_{min}(m) = \frac{1}{2}\left(1 - \cos\frac{\pi}{2m}\right). \tag{4.43}$$

This inequality is the answer of the reverse problem: which weights can be distinguished from the zero function for a given number of iterations.

Setting $m = 1$ in Eq. 4.43, we observe that any weight larger than 0.5 can be discriminated from the zero function. This is in fact a special case that corresponds to a variation of the Deutsch-Jozsa problem[10], where in the Deutsch-Jozsa problem the functions with weight 0.5 are discriminated from functions with weights either zero or one. For a weight smaller than $\frac{1}{2}$, we would need more than one iterations. More generally, as we will discuss further in this section, the more the weights are close to each other, the more evaluations are necessary to distinguish them without error. This is also the case for the special case of zero function discrimination. When ρ is much less than 1, we require $m \sim \frac{\pi}{4\sqrt{\rho}}$ function evaluations to distinguish a non-zero function from a zero function.

Table 4.1: Minimum weights of functions, which can be distinguished from the zero function z by only m function evaluations

m	$\rho_{min}(m)$
1	0.5^1
2	0.15
3	0.067
4	0.038
5	0.024
10	0.0062

4.3.2.3 Discrimination of two non-zero weights

Now, let us continue with the general case of two non-zero weights, $\rho \neq 0 \neq \rho'$. The choice of $|\beta_i\rangle$ in Eq. 4.31 make sure that all the matrices Θ, $\mathbb{C}$ and $\mathbb{R}$ are diagonal. Let us have two functions f and g with respective weights ρ and ρ'. We are interested in the inner product of the final states

$$\begin{aligned}\langle\Phi_f|\Phi_g\rangle &= \sum_{i=1}^{n}|c_i|^2\Bigg(\mathbb{R}_{ii}^{(m+1)}(f)\,\mathbb{R}_{ii}^{(m+1)}(g)-\mathbb{R}_{ii}^{(m+1)}(f)\,\mathbb{R}_{ii}^{(m)}(g)\cos\theta_i(g)\\ &-\mathbb{R}_{ii}^{(m)}(f)\,\mathbb{R}_{ii}^{(m+1)}(g)\cos\theta_i(f)\\ &+\mathbb{R}_{ii}^{(m)}(f)\,\mathbb{R}_{ii}^{(m)}(f)\cos\theta_i(f\oplus g)\Bigg)=0. \qquad (4.44)\end{aligned}$$

where we utilized Theorem 2 again to obtain this result. The last term inside the sum of Eq. 4.44 comes from the inner product

$$\begin{aligned}\langle\beta_i|\left(\mathbb{1}\otimes U_f\right)\left(\mathbb{1}\otimes U_g\right)|\beta_i\rangle &= \langle\beta_i|\left(\mathbb{1}\otimes U_{f\oplus g}\right)|\beta_i\rangle \qquad (4.45)\\ &= \cos\theta_i\left(f\oplus g\right) \qquad (4.46)\end{aligned}$$

and this cosine term can be evaluated as

$$\cos\theta_i\left(f\oplus g\right)=\langle\beta_i|\mathbb{1}_A\otimes U_{f\oplus g}|\beta_i\rangle=1-\mu_i\frac{2t}{N}, \qquad (4.47)$$

where t is the number of solutions to the equation $(f\oplus g)(x)=1$. Equivalently, the weight of the function $f\oplus g$ is t/N. The possible values t can have are $t=|r-r'|,|r-r'|+2,\ldots,r+r'-2,r+r'$, considering any f with weight ρ and any g with weight ρ'. Observe that Eq. 4.44 is linear in t. Hence, it can be reduced to two independent equations

$$\begin{aligned}\sum_{i=1}^{n}|c_i|^2 A_i &= 0, \qquad (4.48)\\ \sum_{i=1}^{n}|c_i|^2 B_i &= 0 \qquad (4.49)\end{aligned}$$

subject to the condition

$$\sum_{i=1}^{n}|c_i|=1, \qquad (4.50)$$

where A_i and B_i are given as

$$\begin{aligned}A_i &\equiv \cos\left(m\theta_{if}\right)\cos\left(m\theta_{ig}\right)\\ &\quad+\frac{\sin\left(m\theta_{if}\right)\sin\left(m\theta_{ig}\right)}{\sin\left(\theta_{if}\right)\sin\left(\theta_{ig}\right)}\left(1-\cos\left(\theta_{if}\right)\cos\left(\theta_{ig}\right)\right), \qquad (4.51)\\ B_i &\equiv \frac{\sin\left(m\theta_{if}\right)\sin\left(m\theta_{ig}\right)}{\sin\left(\theta_{if}\right)\sin\left(\theta_{ig}\right)}\left(2-\cos\left(\theta_{if}\right)\cos\left(\theta_{ig}\right)\right)\\ &= 2\left(\rho+\rho'\right)\mu_i\frac{\sin\left(m\theta_{if}\right)\sin\left(m\theta_{ig}\right)}{\sin\left(\theta\right)\sin\left(\theta_{ig}\right)}. \qquad (4.52)\end{aligned}$$

Therefore our task reduces to finding n-tuplets (A_i, B_i) on a two dimensional $A-B$ plane such that $\sum_i |c_i|^2 (A_i, B_i) = (0,0)$ which means that the convex hull of the set of points $\{(A_i, B_i)\}$ contains the origin, the point $(0,0)$. Let us call the condition that the convex hull of a set of points should include a fixed point as "convex hull property".

Note that, once the problem is fixed with some arbitrary weights ρ and ρ' and a number of iterations m, A_i and B_i depend only on the adjustable parameter μ_i. Let us show only this dependence explicitly (i.e., $A_i = A(\mu_i)$ and $B_i = B(\mu_i)$), so that the set of points (A_i, B_i) is a finite subset of the continuous curve $(A(\mu), B(\mu))$ where $0 \leq \mu \leq 1$. This allows us to make the observation that the origin should also be in the convex hull of the whole curve $(A(\mu), B(\mu))$, but this means that the problem can be solved with $n = 2$. This is because finding only two points on the curve such that the line connecting them includes the origin is enough to show the convex hull property.

In the last paragraph, we have transformed the problem of distinguishability of two weights ρ and ρ' with m function evaluations into a problem in convex analysis: whether a curve and the point origin satisfies convex hull property. We can express the coordinates of the curve $(A(\mu), B(\mu))$ as

$$A(\mu) = T_m(y)\,T_m(y') + U_{m-1}(y)\,U_{m-1}(y')\,(1-yy'), \tag{4.53}$$

$$B(\mu) = U_{m-1}(y)\,U_{m-1}(y')\,(2-y-y'), \tag{4.54}$$

where $y = 1 - 2\rho\mu$, $y' = 1 - 2\rho'\mu$ and T_m and U_m denote the Chebyshev polynomials. Their definition are given as follows:

$$T_m(x) \equiv \cos(m \arccos(x)), \tag{4.55}$$

$$U_m(x) \equiv \frac{1}{m}\frac{d}{dx}T_m(x). \tag{4.56}$$

Even though the parameters $(A(\mu), B(\mu))$ are only polynomials in terms of μ, the problem of determining whether the origin is within the convex hull of the curve can become intractable with increasing number of iterations. In that case we may resort to numerical techniques.

When $m = 1$, $A(\mu) = 1$ for any the value μ, therefore the convex hull of the curve can not contain the origin. Consequently, a single evaluation of oracle is never enough for sure-success discrimination of two weights with both different from 0 or 1. Distinguishability starts with $m = 2$ iterations. An example with $m = 2$ is shown in Fig. 4.2 where the weights are taken to be $\rho = 0.95$ and $\rho' = 0.45$. Observe that the parametric curve $(A(\mu), B(\mu))$ starts from the point $(A(0), B(0)) = (1, 0)$. It turns out that this is the case for all the cases even for $m = 1$. However the curve ends at a nontrivial point. Since it starts from a point on the positive part of the horizontal axis, if we can find a second intersection on the negative part of the axis, we can easily construct a convex combination that gives $(0, 0)$. Even though this simplification covers most of the solutions, for a complete analysis we have to find compact inequalities in terms of the weights ρ and ρ'.

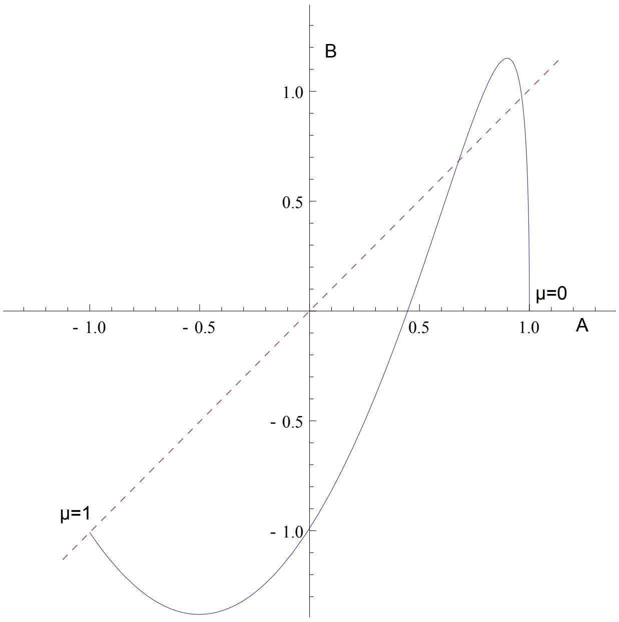

Figure 4.2: $B(\mu)$ vs. $A(\mu)$ for $m = 2$, $\rho = 0.95$ and ρ=0.45

m=2 case

Before studying the general case, we start with the simplest cases where we can find analytical solutions. For $m = 2$ we will see that compact formulas for distinguishability is possible. When $m = 2$, μ dependence of A and B is found as

$$\begin{aligned} A^{(2)}(\mu) &= 1-8\mu^2(\rho-\rho')^2, & (4.57)\\ B^{(2)}(\mu) &= 8\mu(\rho+\rho')(1-2\mu\rho)(1-2\mu\rho'), & (4.58)\end{aligned}$$

where we introduced the superscripts in order to emphasize the number of iterations. Our requirement is that, there exists two distinct points $\left(A^{(2)}(\mu_1), B^{(2)}(\mu_1)\right)$ and $\left(A^{(2)}(\mu_2), B^{(2)}(\mu_2)\right)$ on the curve, such that the following is satisfied

$$\sum_{i=1,2} |c_i|^2 \begin{pmatrix} A^{(2)}(\mu_i) \\ B^{(2)}(\mu_i) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \qquad (4.59)$$

Hence there should exist two values, μ_1 and μ_2 such that

$$\frac{B^{(2)}(\mu_1)}{A^{(2)}(\mu_1)} = \frac{B^{(2)}(\mu_2)}{A^{(2)}(\mu_2)}, \qquad (4.60)$$

and both $\left(A^{(2)}(\mu_1), A^{(2)}(\mu_2)\right)$ and $\left(B^{(2)}(\mu_1), B^{(2)}(\mu_2)\right)$ pairs should have opposite signs, i.e., if $A^{(2)}(\mu_1)$ is positive then $A^{(2)}(\mu_2)$ is negative or vice versa and same thing for $B(\mu)$'s as well. A helpful observation is that if there is a solution to Eq. 4.60, then there is always a solution with $\mu_2 = 1$. It can be proved by considering all the possible orderings of the two roots $\mu_0 \equiv (2\rho)^{-1}$, $\mu_0' \equiv (2\rho')^{-1}$ and the pole $\mu_p \equiv \left(2\sqrt{2}\,|\rho-\rho'|\right)^{-1}$ of $B^{(2)}(\mu)/A^{(2)}(\mu)$. The details of the proof can be found in Ref. [27]. By setting $\mu_2 = 1$, we obtain the following

$$B^{(2)}(\mu_1)A^{(2)}(1) - A^{(2)}(\mu_1)B^{(2)}(1) = \left(K\mu_1^2 + L\mu_1 + M\right)(\mu_1 - 1) = 0 \qquad (4.61)$$

with

$$\mu_1 \leq \mu_p \qquad (4.62)$$

where

$$\begin{aligned} K &= 4\rho\rho'\left(1-8(\rho-\rho')^2\right) & (4.63)\\ L &= 8(\rho-\rho')^2 + 4\rho\rho' - 2(\rho+\rho') & (4.64)\\ M &= (2\rho-1)(2\rho'-1). & (4.65)\end{aligned}$$

As it can be seen from Eq. 4.61, for two iterations, the hardest step to find the conditions on the weights is solving a quadratic equation. For more than two iterations, the algebraic equation to be solved will be of higher degrees and it will be harder to solve analytically.

General case

Applying the same procedure for three iterations, we obtain the following $(A(\mu), B(\mu))$ coordinates

$$
\begin{aligned}
A^{(3)}(\mu) &= -128\mu^4\rho\rho'(\rho-\rho')^2 + 64\mu^3(\rho-\rho')^2(\rho+\rho') \\
&\quad -48\mu^2(\rho-\rho')^2 + 1, \qquad (4.66) \\
B^{(3)}(\mu) &= \mu\left(16\mu^2\rho^2 - 16\mu\rho + 3\right)\left(16\mu^2\rho'^2 - 16\mu\rho' + 3\right). \qquad (4.67)
\end{aligned}
$$

Observe that for the case where $m = 3$, the order of $A^{(3)}(\mu)$ and $B^{(3)}(\mu)$ are 4 and 3 respectively. The most difficult equation to be solved is now 4th order. Comparing it with the 2nd order Eq. 4.61, we see that it would be impractical to try to solve these equations for $m > 2$ analytically. Fortunately, we can employ numerical methods without so much effort. Note that it is also possible to solve this problem directly without applying Theorem 2. However without Theorem 2, one needs to optimize $|\beta_i\rangle$'s one by one. In contrast, in the current approach, optimization of only one parameter, μ is enough and it is clearly an advantage for both analytical and numerical approaches.

Numerical results obtained with our method for several iterations $(m = 2, 3, \ldots, 8)$ are plotted in Fig 4.3 along with a few results obtained by the algorithm given by Choi *et al.* Observe that, starting farthest from the diagonal, each iteration contributes enclosing more area representing distinguishable space of weights. With each run, amount of this enclosure decreases, as the problem becomes harder for weights that are closer to each other.

Symmetries

Notice that the patterns are symmetric with respect to the main diagonal (i.e., $\rho = \rho'$ line). This observation is anticipated because whether we swap the weights or not, starting from the same initial state, the inner product of the final states obtained from these weights doesn't change. However we would expect one more symmetry: distinguishing ρ from ρ' should not be different from distinguishing $1-\rho$ from $1-\rho'$ since this corresponds to inverting the function output for all the inputs. The complexity of the problem does not

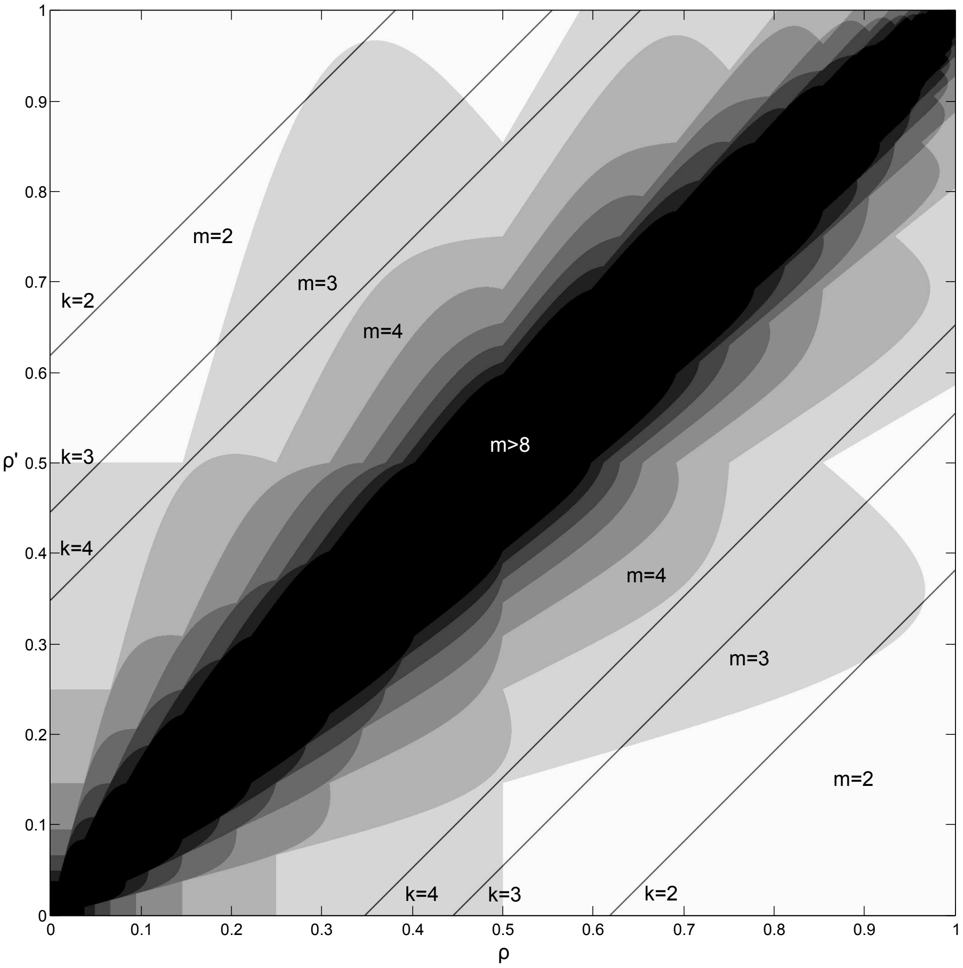

Figure 4.3: Number of iterations up to 8 as a function of ρ and ρ'. The areas with darker shade correspond to higher m values and the areas with lighter shade or no shade correspond to lower m values. The lightest region is for $m = 2$. The disc-like black region on the diagonal corresponds to weight pairs that have no solutions with $m \leq 8$ iterations. To be able to make a comparison, the whole region is also divided by 6 lines which are parallel to the diagonal. These divisions, from the outermost to the inner ones, correspond to the weight pairs distinguishable by $k = 2, 3, 4$ iterations with the algorithm given in [24, 25].

change with such modification. Unfortunately our formalism does not detect this kind of bit flip in any part of the computation. It is possible to remove this deficiency by adjusting the unitary S in the algorithm by changing it to $S' = S(\mathbb{1} \otimes U_u)$ where U_u corresponds to the unitary for the constant function $u(x) = 1$. It can easily be checked that S' has complex eigenvalues and also many of the eigenvectors that correspond to -1 are not in the form of Eq. 4.31. Thus, our conclusion is that, even though it is possible to design an algorithm with an inherent $(\rho, \rho') \leftrightarrow (1 - \rho, 1 - \rho')$ symmetry, it would be unnecessarily complicated. Alternatively we can run the algorithm one more time with the inverted outputs and make a decision after considering both results with the expense of doubling the run-time of the algorithm. The decidability of weights when such a symmetry is taken into consideration is drawn in Fig. 4.4.

4.3.3 A comparison of query complexities

We end this chapter by a brief comparison of query complexities of the algorithms given in Chapter 4.

Classical complexities

Let us start with the classical case first. A deterministic classical algorithm would use approximately

$$m_{cl,det} \sim N\left(1 - |\rho - \rho'|\right) \tag{4.68}$$

steps to make a decision on two weights in the worst case. However a probabilistic algorithm could achieve better. Let us assume that, with such an algorithm, minimum s random queries are needed to discriminate two weights ρ and ρ'. Then, the sum of standard deviations should be smaller than the difference of the weights

$$\sigma_\rho - \sigma_{\rho'} < |\rho - \rho'| \tag{4.69}$$

where σ_λ denotes the standard deviation of the random shooting process of a function with weight λ. This scenario is one of the typical examples of hypergeometric distribution, but calculating the standard deviation may be difficult for

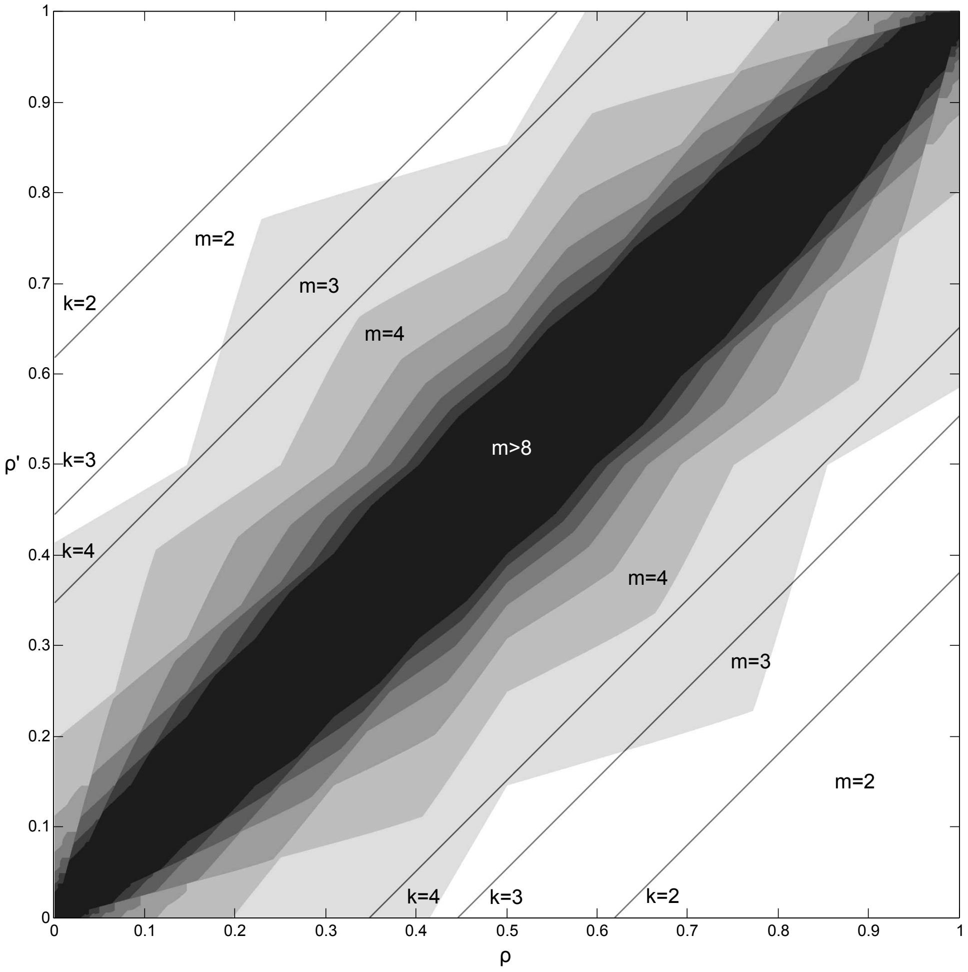

Figure 4.4: Symmetrized version of Fig. 4.3.

it. Luckily, for large N the hypergeometric distribution can be approximated to a binomial process, thus the standard deviation can be given as

$$\sigma_\rho \simeq \sqrt{\frac{\rho\,(1-\rho)}{m}} \tag{4.70}$$

and using Eq. 4.69 it is found that at least

$$m_{cl,prb} \sim \frac{4\rho\,(1-\rho)}{|\rho-\rho'|^2} \tag{4.71}$$

evaluations are necessary to discriminate two weights with negligible error.

Complexity of quantum counting and the method by Choi *et al.*

Since weight discrimination is a special case of quantum counting, we would like to present the query complexity of the quantum counting for the problem of discrimination of two weights first. In Refs. [24, 25] the query complexity for quantum counting is given with the following argumentation: Since any weights can be symmetrized, finding the complexity for symmetric weights is sufficient. For symmetric weights one can write the weights as

$$\begin{aligned} \rho_1 &= \sin^2\left(\frac{m\pi}{4m+2}\right) && (4.72)\\ \rho_2 &= \cos^2\left(\frac{m\pi}{4m+2}\right) && (4.73) \end{aligned}$$

and setting $P = 4m+2$ is enough to check that whether the weight is ρ_1 or ρ_2. On the other hand it has been shown in [24, 25] and also its optimality is verified in [26] that m queries are needed for the method by Choi *et al.* In [26], optimal m has been expressed in terms of the weights as

$$m_{BC} \sim \frac{\pi}{2}\frac{\sqrt{\rho\,(1-\rho')}}{|\rho-\rho'|} \tag{4.74}$$

in the regime where the weights are close to each other and to $\frac{1}{2}$. Comparing them with the classical case, a square-root speedup has been achieved by both quantum counting and the method of Choi et al. Braunstein and Choi's method is 4 times faster than the quantum counting and does not need a quantum Fourier transform however it is a specialized algorithm for weight discrimination whereas quantum counting is a more general algorithm in the sense that it can discriminate any number of weights.

Complexity of our method

As we discussed before, we always have the point $(A, B) = (1, 0)$ as the starting point in the geometrical picture we presented to understand the μ dependence of the parameters A and B. Finding a second point on the negative side of the horizontal axis would lead to a solution since the $(0, 0)$ point can be obtained by a convex combination of these two points. Even though better solutions can be found for $m > 2$, this approximation is good enough to estimate the order of quantum complexities. Therefore we search for the roots of B and check the sign of A at these points. Note that, either $m\theta_{if}$ or $m\theta_{ig}$ is an integer multiple of π at the roots of B. in Eq. 4.52 so, the second term in Eq. 4.51 vanish. From negativity of A, Eq. 4.75 follows:

$$|m\theta_{if} - m\theta_{ig}| > \frac{\pi}{2}. \tag{4.75}$$

Since the limits of the algorithm are tested when the weights are close to each other, we can assume so and linearize the expression as

$$\begin{aligned}
m\left(\theta_{if} - \theta_{ig}\right) &= m\left(\cos^{-1}\left(1 - 2\mu\rho\right) - \cos^{-1}\left(1 - 2\mu\rho'\right)\right) && (4.76)\\
&\simeq m\left(\rho - \rho'\right)\frac{\partial}{\partial\rho}\cos^{-1}\left(1 - 2\mu\rho\right) && (4.77)\\
&= m\frac{\rho - \rho'}{\sqrt{\rho\left(\frac{1}{\mu} - \rho\right)}} > \frac{\pi}{2} && (4.78)
\end{aligned}$$

and by making use of the observation that for an optimal solution $\mu \lesssim 1$ we obtain

$$m_{UT} \sim \frac{\pi}{2}\frac{\sqrt{\rho\left(1 - \rho\right)}}{|\rho - \rho'|}. \tag{4.79}$$

Let us compare this result with the previous ones. A quadratic speedup in comparison with classical algorithms is achieved. Similar to Braunstein and Choi's method, this method also works four times faster than quantum counting. There is almost no difference between the two specialized weight decision algorithms in the large number of iterations regime. However as shown in Fig. 4.3, for most of the weights that can be discriminated by small number of iterations, our method gives slightly improved results, whereas Braunstein and Choi's results are better than ours for some of the weights. The former observation is partly because of the fact that, Braunstein and Choi's method sets the number of iterations in the beginning of their algorithm only by looking at the difference of the weights regardless of the weights themselves. On the contrary, our method is more flexible so that we are able to adjust the parameters of the algorithm for optimal number of iterations for all the weight decision scenarios. However the latter observation also takes place since our method lacks a bit flip symmetry. As we have seen in Fig. 4.4, symmetrized version of our method is strictly superior to the one in [25] for the first few iterations.

CHAPTER 5

CONCLUSIONS AND OUTLOOK

In this thesis we studied the sure-success weight decision problem of Boolean functions using quantum algorithms. In particular, we progressed through two methods.

First, we explored the problem within the framework of quantum operator discrimination in Chapter 3 after giving necessary background information on quantum state discrimination and quantum operator discrimination in Chapter 2. In this approach we observed that only parallel calls to the function augments the distinguishability of functions but not so much. Alternatively, the unitaries corresponding to the function can be applied in a sequential manner. Thus we followed this approach next and obtained more promising results. The serial formulation is advantageous in the sense that any other formulation can be translated into a serial formalism[21, 22]. Consequently, the main focus in Chapter 3 was serial applications. We intoduced a powerful theorem that creates a connection between quantum protocols that consist of p sequential applications of the unitary evaluation that realizes the function and the density matrices in the product Hilbert space $\mathcal{H}^{\otimes p}$. Through this connection we were able to establish the conditions for $p = 2$ evaluations without much effort. In contrast, even for $p = 3$ evaluations, the equations for distinguishability became intractable.

In the second method, we applied Grover iteration directly to the problem. Grover's iteration in its original form is not exact. We modified the algorithm in a novel way so that an exact discrimination of functions with different weights is possible. As of this writing, sure success discrimination can be achieved with three algorithms, quantum counting[23], Braunstein and Choi's algorithm[24, 25] and our method[27]. We reviewed all these methods in Chapter 4. As its name suggests, quantum counting algorithm gives the number of inputs of a Boolean function that lead to the output 1. Since it can distinguish any number of functions with different weights, decision of two functions becomes a special case. The other two algorithms are specialized versions in the sense that they can only discriminate two weights and all three achieve a square-root speedup as compared to classical algorithms. This should not be a surprising fact because they all are based on the Grover iteration and the query complexity reduction of Grover algorithm is shown to be no more than square root[67, 77, 78]. There is however a slight difference of complexities between these algorithms. Specialized algorithms turns out to be approximately 4 times faster than the quantum counting algorithm. When the weights are closer to each other, more queries are necessary. We made a query comparison of two algorithms in this region of the weight space and found that our method gives approximately the same speed as the Braunstein and Choi's. In a closer look, our results indicate that for different weight combinations in the small number of iterations regime, better quantum discrimination scenarios are possible.

Future work

There can be several possible extensions of these two approaches. Here we list some of them. We have tested the density matrix correspondence only for the weight decision problem of two functions. From the quantum operator discrimination perspective, "the weight" in our problem, becomes the common property that defines which set a unitary operator belongs to. The task then can be reformulated as the quantum operator discrimination of these sets of unitaries. One can define a more general property such as a class of weights versus some other class or weights. An example could be discriminating zero function from any other non-zero function. Conversely one can define a more specific version or consider a completely different property of a Boolean function. Weight is only one of the components in a Walsh transform of a Boolean function. A possible direction may be taking other components of the Walsh transform into consideration. Special functions such as bent functions which always lead to a balanced Walsh transform have useful applications in cryptography[84]. The methods we introduced in this thesis may be extended to determine or make use of such properties.

Weight is in a sense the Boolean counterpart of the definite integral of a function over its domain. A possible long-term continuation of our work may be discrimination of real functions whose integral over a specific domain is guaranteed to give one of the two predetermined vaues.

We used a novel technique to make Grover's algorithm a sure-success one for the weight decision problem. Even though there exists several other methods in the literature for the same purpose, our method may be preferrable in other problems with fastest quantum algorithm solutions are available using Grover iteration. Searching for such problems can be counted as one of the many possible continuations of our works compiled in this thesis.

REFERENCES

[1] M.A. Nielsen and I.L. Chuang, Quantum Computation and Quantum Information, Cambridge University Press, Cambridge (2000).

[2] R. Feynman, Intl. J. Theor. Phys., **21**, 467–488 (1982).

[3] R. Feynman, Feynman lectures on computation. Addison-Wesley, Menlo Park, California (1996).

[4] P. Benioff, J. Stat. Phys., **22**, 563–591 (1980).

[5] P. Benioff, J. Stat. Phys., **29**, 515–546 (1982).

[6] P. Benioff, Phys. Rev. Lett., **48**, 1581–1585, (1982).

[7] D.M. Bacon, PhD thesis (2001); preprint:quant-ph/0305025v1.

[8] S. Lloyd, Science, **273**, 1073 (1996).

[9] D. Deutsch, Proc. Roy. Soc. London A, **400**, 97–117 (1985).

[10] D. Deutsch and R. Jozsa, Proc. Roy. Soc. London A, **439**, 553–558 (1992).

[11] A.M. Turing, Proc. Lond. Math. Soc. 2, **42**, 230–265, (1936).

[12] A. Church, American J. Math., **58**, 345–363 (1936).

[13] A. Church, Annals of Mathematics, second series, **33**, 346–366, (1936).

[14] A.C. Yao, In Proc. of the 34th Ann. IEEE Symp. on Foundations of Computer Science, 352–361 (1993).

[15] E. Bernstein and U. Vazirani, In Proceedings of the 25th Annual ACM Symposium on Theory of Computation, 11–20, (1993).

[16] D. Simon, Proc. 35th Ann. Symp. on Foundations Comput. Sci., 116–123 (1994). Journal version appeared in SIAM J. Comput. **26**, 1474–1483 (1997) .

[17] P.W. Shor, Proc. 35th Ann. Symp. on Foundations of Comp. Sci., 124–134, (1994). Journal version appeared in SIAM J. Comput. **26**, 1484–1509 (1997).

[18] L.K. Grover, Proc. 28th Ann. ACM Symp. on the Theory of Computing, ACM Press, New York, 212–219, (1996). Journal version appeared in Physical Review Letters, **79**, 325–328 (1997).

[19] A. Jamiolkowski, Rep. Math. Phys. **3**, 275 (1972).

[20] M. Choi, Linear Algebr. Appl. **10**, 285 (1975).

[21] R. Duan, Y. Feng, and M. Ying, Phys. Rev. Lett. **98**, 100503, (2007).

[22] X.D.Wu and R.Y. Duan, Phys. Rev. A **78**, 012303 (2008).

[23] G. Brassard, P. Høyer and A. Tapp, Proceedings of the 25th International Colloquium on Automata, Languages and Programming (Lecture Notes In Computer Science) **1443**, 820–31 (1998).

[24] S.L. Braunstein, B.S. Choi, S. Ghosh and S. Maitra, J. Phys. A: Math. Theor. **40**, 8441–8454 (2007).

[25] B.S. Choi and S.L. Braunstein, Quantum Information Processing, **10**, 177–188 (2011).

[26] B.S. Choi, Quantum Inf. Process. **11**, 123 (2012).

[27] K. Uyanik and S. Turgut, Quantum Information Processing, **12**, 3395–3409 (2013).

[28] C.W. Helstrom, Quantum Detection and Estimation Theory; Academic Press: New York (1976).

[29] A.S. Holevo, Probabilistic and Statistical Aspects of Quantum Theory; North-Holland: Amsterdam, (1976).

[30] A.S. Holevo, J. Multivariate Anal., **3**, 337–394 (1973).

[31] H.P. Yuen and R.S. Kennedy, Lax, M. IEEE Trans. Inform. Theory, **21**, 125–134 (1975).

[32] J.A. Bergou, J. Mod. Opt. **57**, 160–280 (2010).

[33] I.D. Ivanovic, Phys. Lett. A, **123**, 257–259 (1987).

[34] D. Dieks, Phys. Lett. A, **126**, 303–306 (1988).

[35] A. Peres, Phys. Lett. A, **128**, 19 (1988).

[36] S. Pang and A. Wu, Phys. Rev. A, **80**, 052320 (2009).

[37] A. Chefles, Phys. Lett. A, **239**, 339–347, (1998).

[38] A. Chefles, Phys. Rev. A, **64**, 062305, (2001).

[39] Y. Feng, R. Y. Duan and M. S. Ying, Phys. Rev. A, **70**, 012308 (2004).

[40] T. Rudolph, R.W. Spekkens, and P.S Turner, Phys. Rev. A, **68**, 010301R (2003).

[41] G. Wang and M. Ying, Phys. Rev. A, **73**, 042301, (2006).

[42] S. Croke, E. Andersson, S.M. Barnett, C.R. Gilson and J. Jeffers, Phys. Rev. Lett., **96**, 070401 (2006).

[43] A. Chefles and S.M. Barnett, J. Mod. Opt., **45**, 1295–1302 (1998).

[44] C.W. Zhang, C.F. Li and G.C. Guo, Phys. Lett. A, 1999, 25–29 (1999).

[45] J. Fiurasek and M. Jezek, Phys. Rev. A, **67**, 012321 (2003).

[46] S.M. Barnett and S. Croke, Adv. Opt. Photon., **1**, 238 (2009).

[47] G.M. D'Ariano, M.F. Sacchi and J. Kahn, Phys. Rev. A, **72**, 032310 (2005).

[48] A. Chefles, A. Kitagawa, M. Takeoka, M. Sasaki, and J. Twamley, J. Phys. A, **40**, 10183–10213 (2007).

[49] A. Acin, Phys. Rev. Lett. **87**, 177901, (2001).

[50] G.M. D'Ariano, P.L. Presti and M.G.A. Paris, Phys. Rev. Lett. **87**, 270404 (2001).

[51] M.F. Sacchi, Phys. Rev. A, **71**, 062340, (2005).

[52] M.F. Sacchi, J. Opt. B: Quantum Semiclass Opt., **7**, 333–336 (2005).

[53] J.A. Bergou and M. Hillery, Phys. Rev. A, **72,** 012302 (2005).

[54] L.M. Duan and G.C. Guo, Phys. Rev. Lett. **80**, 4999 (1998).

[55] R. Bhatia, Matrix Analysis, Springer Graduate Texts in Mathematics, **169** Springer, New York, (1996).

[56] V.I. Paulsen, Completely Bounded Maps and Dilations Longman Scientific and Technical, New York, (1986).

[57] D. Aharonov, A. Kitaev, and N. Nisan, in Proceedings of the 30th Annual ACM Symposium on Theory of Computation STOC, unpublished, 20. (1997).

[58] M. Piani and J.Watrous, Phys. Rev. Lett. **102**, 250501, (2009).

[59] L.Z. Li and D.W. Qiu Phys. Rev. A, **77**, 032337 (2008).

[60] R. Duan, Y. Feng and M. Ying, Phys. Rev. Lett. **103**, 210501 (2009).

[61] A. Childs, J. Preskill, and J. Renes., Journal of Modern Optics, **47**(2–3), 155–176, (2000).

[62] J. Fiurasek and M. Micuda, Phys. Rev. A, **80**, 042312 (2009).

[63] G.M. D'Ariano, M.F. Sacchi, and J. Kahn, Phys. Rev. A, **72**, 052302 (2005).

[64] J. Watrous, Quantum Inf. Comput. **8**(8–9), 0819–0833 (2008).

[65] M. Sedlak and M. Ziman, Phys. Rev. A, **79**, 012303, (2009).

[66] P. Arrighi and C. Patricot, Ann. Phys. N.Y. **311**, 26 (2004).

[67] M. Boyer *et al.*, in Proceedings of the 4th Workshop on Physics and Computation, Los Alamitos, CA, (1996) ~Institute of Electrical and Electronics Engineers Computer Society Press, Los Alamitos, 36–43 (1996); see also Fortsch. Phys. **46**, 493 (1998); preprint:quant-ph/9605034.

[68] L.K. Grover, Phys. Rev. Lett. **80**, 4329 (1998).

[69] E. Biham, O. Biham, D. Biron, M. Grassl and D. Lidar, Phys. Rev. A, **60**, 2742 (1999).

[70] E. Biham, O. Biham, D. Biron, M. Grassl, D. A. Lidar and D. Shapira, Phys. Rev. A, **63**, 012310 (2000).

[71] Z. Diao, Phys. Rev. A, **82**, 044301 (2010).

[72] G. Brassard, P. Høyer, M. Mosca, and A. Tapp, preprint:quant-ph/0005055.

[73] P. Høyer, Phys. Rev. A, **62**, 052304 (2000).

[74] G.L. Long, Phys. Rev. A, **64**, 022307 (2001).

[75] F.M. Toyama, W. van Dijk and Y. Nogami, Quantum Information Processing, **12**, 1897–1914, (2013).

[76] J.Y. Hsieh and C.M. Li, Phys. Rev. A, **65**, 052322, (2002).

[77] C. Bennett *et al.*, SIAM J. Comput. **26**, 1510 (1997).

[78] C. Zalka, Phys. Rev. A **60**, 2746 (1999).

[79] T. Tulsi, L.K. Grover and A.Patel, Quantum Inf. Comput., **6**, 483 (2006).

[80] E. Filiol and C. Fontaine, In Proceedings of Advances in Cryptology-EUROCRYPT'98, InternationalConference on the Theory and Application of Cryptographic Techniques. Lecture Notes in Computer Science, **1403**, 475, (1998).

[81] F.J. MacWilliams and N.J.A. Sloane, The Theory of Error-Correcting Codes. North Holland, Amsterdam (1996).

[82] K. Chakrabarty and J.P. Hayes, IEEE Trans. VLSI Syst., **3**(1), 72 (1995).

[83] K. Chakrabarty and J.P. Hayes, J. Electron Test. Theory Appl. **8**(1), 71 (1996).

[84] T.W. Cusick and P. Stanica. Cryptographic Boolean Functions and Applications. Elsevier Inc., San Diego, USA, (2009).

APPENDIX A

PROOF OF THE DENSITY MATRIX CORRESPONDENCE THEOREM

Here we give the proof of Theorem 1

Proof. For a given protocol $\{|\varphi\rangle, \mathbb{V}^{(k)}\}$ define

$$\left|\varphi^{(k)}(i_k, i_{k-1}, \ldots, i_1)\right\rangle \equiv V^{(k-1)}_{i_k,i_{k-1}} V^{(k-2)}_{i_{k-1},i_{k-2}} \cdots V^{(1)}_{i_2,i_1} |\varphi(i_1)\rangle \tag{A.1}$$

thus, the final state becomes

$$|\Phi_f\rangle = \sum_{i_1 \cdots i_p} f_{i_1} \ldots f_{i_p} |i_p\rangle_X \otimes \left|\varphi^{(p)}(i_p, \ldots, i_1)\right\rangle_A \tag{A.2}$$

and the overlap matrix is

$$S_{fg} = \langle \Phi_f | \Phi_g \rangle = \sum_{\substack{i_1 \ldots i_p \\ j_1 \ldots j_p}} f^*_{i_1} \ldots f^*_{i_p} g_{j_1} \ldots g_{j_p} \delta_{i_p j_p} \left\langle \varphi^{(p)}(i_p, \ldots, i_1) | \varphi^{(p)}(j_p, \ldots, j_1) \right\rangle. \tag{A.3}$$

For the forward implication, define $\rho = \rho^{(p)}$ by

$$\rho \equiv \sum_{\substack{i_1 \ldots i_p \\ j_1 \ldots j_p}} \delta_{i_p j_p} \left\langle \varphi^{(p)}(i_p, \ldots, i_1) | \varphi^{(p)}(j_p, \ldots, j_1) \right\rangle |i_1 \ldots i_p\rangle \langle j_1 \ldots j_p| \tag{A.4}$$

so that

$$\langle i_p \ldots i_1 | \rho | j_p \ldots j_1 \rangle = \delta_{i_p j_p} \left\langle \varphi^{(p)}(i_p, \ldots, i_1) | \varphi^{(p)}(j_p, \ldots, j_1) \right\rangle \tag{A.5}$$

and we recover the overlap matrix given in Equation (A.3)

$$S_{fg} = \sum_{\substack{i_1 \ldots i_p \\ j_1 \ldots j_p}} \left(f_{i_1} \ldots f_{i_p}\right)^* \left(g_{j_1} \ldots g_{j_p}\right) \langle i_p \ldots i_1 | \rho | j_p \ldots j_1 \rangle \tag{A.6}$$

$$\times \left\langle f^{\otimes p} | \rho | g^{\otimes p} \right\rangle \tag{A.7}$$

We will prove the diagonality property of ρ by induction. To do this, first observe that $\rho^{(p)}$ is diagonal in X_p:

$$\begin{aligned}\rho^{(p)} &= \sum_{\substack{i_1\ldots i_p\\ j_1\ldots j_p}} \left(|i_p\rangle\langle i_p|\right)_{X_p} \otimes \delta_{i_k j_k} \Big(\left\langle \varphi^{(p)}(i_{p-1},\ldots,i_1)|\varphi^{(p)}(j_{p-1},\ldots,j_1)\right\rangle_{X_{p-1}\ldots X_1} \\ &\qquad |i_{p-1}\ldots i_1\rangle\langle j_{p-1}\ldots j_1|\Big) \qquad \text{(A.8)}\end{aligned}$$

Now let us assume $\rho^{(k)}$ is diagonal in X_k, thus $\rho^{(k)}$ is of the form

$$\begin{aligned}\rho^{(k)} &= \sum_{\substack{i_1\ldots i_k\\ j_1\ldots j_k}} \left(|i_k\rangle\langle i_k|\right)_{X_k} \otimes \delta_{i_k j_k} \Big(\left\langle \varphi^{(k)}(i_k,\ldots,i_1)|\varphi^{(k)}(j_k,\ldots,j_1)\right\rangle \\ &\qquad |i_k\ldots i_1\rangle\langle j_k\ldots j_1|\Big)_{X_{p-1}\ldots X_1} \qquad \text{(A.9)}\\ &= \sum_{\substack{i_1\ldots i_k\\ j_1\ldots j_{k-1}}} \left(|i_k\rangle\langle i_k|\right)_{X_k} \\ &\qquad \otimes \Big(\left\langle \varphi^{(k-1)}(i_{k-1},\ldots,i_1)|V^\dagger_{i_k i_{k-1}} V_{i_k j_{k-1}}|\varphi^{(k-1)}(j_{k-1},\ldots,j_1)\right\rangle \\ &\qquad |i_k\ldots i_1\rangle\langle j_k\ldots j_1|\Big)_{X_{p-1}\ldots X_1}. \qquad \text{(A.10)}\end{aligned}$$

Then $\rho^{(k-1)}$ can be obtained as

$$\begin{aligned}\rho^{(k-1)} &= \mathrm{Tr}_{X_k}\rho^{(k)} \qquad \text{(A.11)}\\ &= \sum_{\substack{i_1\ldots i_k\\ j_1\ldots j_{k-1}}} \Big(\left\langle \varphi^{(k-1)}(i_{k-1},\ldots,i_1)|V^\dagger_{i_k i_{k-1}} V_{i_k j_{k-1}}|\varphi^{(k-1)}(j_{k-1},\ldots,j_1)\right\rangle \\ &\qquad |i_{k-1}\ldots i_1\rangle\langle j_{k-1}\ldots j_1|\Big)_{X_{p-1}\ldots X_1} \qquad \text{(A.12)}\\ &= \sum_{\substack{i_1\ldots i_{k-1}\\ j_1\ldots j_{k-1}}} \Big(\left\langle \varphi^{(k-1)}(i_{k-1},\ldots,i_1)|\delta_{i_{k-1}j_{k-1},}|\varphi^{(k-1)}(j_{k-1},\ldots,j_1)\right\rangle \\ &\qquad |i_{k-1}\ldots i_1\rangle\langle j_{k-1}\ldots j_1|\Big)_{X_{p-1}\ldots X_1} \qquad \text{(A.13)}\\ &= \sum_{\substack{i_1\ldots i_{k-1}\\ j_1\ldots j_{k-1}}} \Big(\delta_{i_{k-1}j_{k-1},}\left\langle \varphi^{(k-1)}(i_{k-1},\ldots,i_1)|\varphi^{(k-1)}(j_{k-1},\ldots,j_1)\right\rangle \\ &\qquad |i_{k-1}\ldots i_1\rangle\langle j_{k-1}\ldots j_1|\Big)_{X_{p-1}\ldots X_1}. \qquad \text{(A.14)}\end{aligned}$$

Thus $\rho^{(k-1)}$ is also diagonal in X_{k-1}. We were able to write Eq. (A.13) because of the special structure of the $\mathbb{V}$ matrices:

$$\begin{aligned}
\mathbb{V}\mathbb{V}^\dagger &= \mathbb{1} && \text{(A.15)}\\
\sum_{ijkl} \mathbb{V}\left(|i\rangle\,|\psi_l\rangle\right)\left(\langle i|\,\langle\psi_l|\right)\mathbb{V}^\dagger &= \mathbb{1} && \text{(A.16)}\\
\sum_{ijkl}\left(|i\rangle\, V_{ji}\,|\psi_l\rangle\right)\left(\langle i|\,\langle\psi_l|\, V_{ki}^\dagger\right) &= \mathbb{1} && \text{(A.17)}\\
\sum_{ijkl}\left(|i\rangle\,\langle i|\right)_A\left(V_{ji}\,|\psi_l\rangle\,\langle\psi_l|\, V_{ki}^\dagger\right)_X &= \mathbb{1}_{XA}. && \text{(A.18)}
\end{aligned}$$

This implies

$$\begin{aligned}
\sum_{ilk} V_{ji}\left(\sum_l |\psi_l\rangle\,\langle\psi_l|\right)V_{ki}^\dagger &= \mathbb{1}_A && \text{(A.19)}\\
\sum_{ijk} V_{ji}V_{ki}^\dagger &= \mathbb{1}_A = \delta_{jk}, && \text{(A.20)}
\end{aligned}$$

and therefore we have

$$\sum_{ijk} V_{ik}^\dagger V_{ij} = \delta_{kj}. \tag{A.21}$$

Since $\rho^{(1)}$ is a diagonal matrix, and the trace $\mathrm{Tr}\rho^{(1)}$ reduces to the normal trace and is equal to 1 because of Eq. (3.46). Lastly, to see that $\rho^{(p)}$ is a positive definite matrix, observe from the definition Eq. (A.4) that its matrix elements ρ_{ij} also constitute an overlap matrix and overlap matrices are always positive definite. Therefore these conditions we have shown up to now justify that the matrix we have defined in Eq. (A.4) is a density matrix and satisfy the property "diagonality in X_k".

Now, for the opposite implication, let a density matrix $\rho^{(p)}$ on $(\mathcal{H}_X)^{\otimes p}$be given to us such that $\rho^{(k)}$ is diagonal in X_k for all $k = 1, \ldots, p$. Define

$$\left|\omega^{(k)}(i_k, \ldots, i_1)\right\rangle = \sqrt{\rho^{(k)}}\,|i_k \ldots i_1\rangle\,. \tag{A.22}$$

This is actually a vector in $\mathcal{H}_{X_k} \otimes \cdots \otimes \mathcal{H}_{X_1}$. We know

$$\left\langle\omega^{(k)}(j_k, \ldots, j_1)|\omega^{(k)}(i_k, \ldots, i_1)\right\rangle = \langle j_k, \ldots, j_1|\rho|i_k, \ldots, i_1\rangle \tag{A.23}$$

and this is 0 if $j_k \neq i_k$. Define the $|\theta\rangle$ vector as

$$\left|\theta^{(k)}(j_{k-1}, \ldots, j_1)\right\rangle = \sum_{i=1}^{n} |i\rangle \otimes \left|\omega^{(k)}(i, j_{k-1}, \ldots, j_1)\right\rangle. \tag{A.24}$$

Then,

$$\begin{aligned}
\Theta(j_{sk-1},\ldots;i_{k-1},\ldots) &= \left\langle\theta^{(k)}(j_{k-1},\ldots,j_1)|\theta^{(k)}(i_{k-1},\ldots,i_1)\right\rangle && \text{(A.25)}\\
&= \sum_{i=1}^{n}\langle i|i'\rangle \\
&\quad \times\left\langle\omega^{(k)}(i,j_{k-1},\ldots,j_1)|\omega^{(k)}(i',i_{k-1},\ldots,i_1)\right\rangle && \text{(A.26)}\\
&= \sum_{i=1}^{n}\left\langle i,j_{k-1},\ldots,j_1|\rho^{(k)}|i',i_{k-1},\ldots,i_1\right\rangle && \text{(A.27)}\\
&= \left\langle j_{k-1},\ldots,j_1|\rho^{(k-1)}|i_{k-1},\ldots,i_1\right\rangle && \text{(A.28)}\\
&= \left\langle\omega^{(k-1)}(j_{k-1},\ldots,j_1)|\omega^{(k-1)}(i_{k-1},\ldots,i_1)\right\rangle && \text{(A.29)}\\
&= \langle j_{k-1}|i_{k-1}\rangle \\
&\quad \times\left\langle\omega^{(k-1)}(j_{k-1},\ldots,j_1)|\omega^{(k-1)}(i_{k-1},\ldots,i_1)\right\rangle && \text{(A.30)}
\end{aligned}$$

Hence there is an isometry $\Omega^{(k-1)}:(\mathcal{H}_X)^{\otimes k}\to(\mathcal{H}_X)^{\otimes k}$ such that

$$\left|\theta^{(k)}(j_{k-1},\ldots,j_1)\right\rangle=\Omega^{(k-1)}\left|j_{k-1}\right\rangle_X\otimes\left|\omega^{(k-1)}(j_{k-1},\ldots,j_1)\right\rangle_A. \tag{A.31}$$

At this point we can pass to an ancilla instead of using $(\mathcal{H}_X)^{\otimes k-1}$ as ancilla. Let A be a system with an infinite-dimensional Hilbert space (i.e., $\mathcal{H}_A$ can be required to be sufficiently large.) Form the association

$$\begin{aligned}
\left|\omega^{(k)}(j_k,\ldots,j_1)\right\rangle &\longleftrightarrow \left|\varphi^{(k)}(j_k,\ldots,j_1)\right\rangle_A && \text{(A.32)}\\
(\mathcal{H}_X)^{\otimes k} &\longleftrightarrow \mathcal{H}_A && \text{(A.33)}
\end{aligned}$$

Let $F_k:\mathcal{H}_X{}^{\otimes k}\to\mathcal{H}_A$ be an isometry $(k=1,\ldots,p)$. We can define this anyway we want. F_k is an isometry means

$$F_k^\dagger F_k=\mathbb{1}_{kX}. \tag{A.34}$$

Let $\left|\varphi^{(k)}(j_{k-1},\ldots,j_1)\right\rangle\equiv F_k\left|\omega^{(k)}(j_k,\ldots,j_1)\right\rangle$, thus

$$F_k^\dagger\left|\varphi^{(k)}(j_k,\ldots,j_1)\right\rangle=\left|\omega^{(k)}(j_k,\ldots,j_1)\right\rangle. \tag{A.35}$$

Now, rewrite Eq. (A.31) as

$$\begin{aligned}
\sum_{i=1}^{n}|i\rangle\otimes\left|\omega^{(k)}(i,j_{k-1},\ldots,j_1)\right\rangle &= \Omega^{(k-1)}\left|j_{k-1}\right\rangle \\
&\quad \otimes\left|\omega^{(k-1)}(j_{k-1},\ldots,j_1)\right\rangle && \text{(A.36)}\\
&= \Omega^{(k-1)}\left|j_{k-1}\right\rangle \\
&\quad \otimes F^\dagger\left|\varphi^{(k-1)}(j_{k-1},\ldots,j_1)\right\rangle && \text{(A.37)}
\end{aligned}$$

and apply $\mathbb{1}_X \otimes F_k$, to obtain

$$\sum_{i=1}^{n} |i\rangle \otimes \left|\omega^{(k)}(i, j_{k-1}, \ldots, j_1)\right\rangle = (\mathbb{1}_X \otimes F_k)\, \Omega^{(k-1)} \left(\mathbb{1}_X \otimes F_{k-1}^{\dagger}\right) |j_{k-1}\rangle \tag{A.38}$$

$$\otimes \left|\varphi^{(k-1)}(j_{k-1}, \ldots, j_1)\right\rangle . \tag{A.39}$$

Now define $\mathbb{V}$ as

$$\mathbb{V}^{(k-1)} \equiv (\mathbb{1}_X \otimes F_k)\, \Omega^{(k-1)} \left(\mathbb{1}_X \otimes F_{k-1}^{\dagger}\right), \ \ (k = 2, 3, \ldots, p). \tag{A.40}$$

So $\mathbb{V}^{(k-1)}$ is an isometry on XA and

$$\sum_{i=1}^{n} |i\rangle \otimes \left|\varphi^{(k)}(i, j_{k-1}, \ldots, j_1)\right\rangle = \mathbb{V}^{(k-1)} |j_{k-1}\rangle \otimes \left|\varphi^{(k-1)}(j_{k-1}, \ldots, j_1)\right\rangle , \tag{A.41}$$

hence

$$\left|\varphi^{(k)}(i, j_{k-1}, \ldots, j_1)\right\rangle = V_{ij_{k-1}}^{(k-1)} \left|\varphi^{(k-1)}(j_{k-1}, \ldots, j_1)\right\rangle . \tag{A.42}$$

If we take i as j_k

$$\left|\varphi^{(k)}(j_k, j_{k-1}, \ldots, j_1)\right\rangle = V_{j_k j_{k-1}}^{(k-1)} V_{j_{k-1} j_{k-2}}^{(k-2)} \cdots V_{j_2 j_1}^{(1)} \left|\varphi^{(1)}(j_1)\right\rangle . \tag{A.43}$$

So define $|\varphi\rangle_{XA} = \sum_{i=1}^{n} |i\rangle \otimes \left|\varphi_i^{(1)}\right\rangle$ and then

$$\mathbb{U}_f \mathbb{V}^{(p-1)} \cdots \mathbb{V}^{(1)} \mathbb{U}_f |\varphi\rangle = \sum f_{i_p} \ldots f_{i_1} |i_p\rangle_X \otimes \left|\varphi^{(p)}(i_p, \ldots, i_1)\right\rangle_A \tag{A.44}$$

$$= |\Phi_f\rangle . \tag{A.45}$$

Thus the overlap matrix becomes

$$\langle \Phi_f | \Phi_g \rangle = \left(f_{i_p} \ldots f_{i_1}\right)^* \left(g_{i_p} \ldots g_{i_1}\right) \delta_{i_p j_p} \times \left\langle \varphi^{(p)}(i_p, \ldots, j_1) | \varphi^{(p)}(j_p, \ldots, j_1) \right\rangle \tag{A.46}$$

$$= \left(f_{i_p} \ldots f_{i_1}\right)^* \left(g_{i_p} \ldots g_{i_1}\right) \times \left\langle \omega^{(p)}(i_p, \ldots, j_1) | \omega^{(p)}(j_p, \ldots, j_1) \right\rangle \tag{A.47}$$

$$= \left\langle f^{\otimes p} | \rho | g^{\otimes p} \right\rangle \tag{A.48}$$

This completes the second part of the proof. ■

Personal Information

Surname, Name: Uyanık, Kıvanç
Nationality: Turkish
Date and Place of Birth: 08 March 1982, İzmir
Phone: +90 312 210 4334
Email: kivancuyanik@gmail.com

Education

- **Ph. D. in Physics,** (February 2008 – February 2014),
  Middle East Technical University,
  Advisor: Assoc. Dr. Sadi Turgut,
  Thesis title: Weight Discrimination of Boolean Functions with Quantum Computation.

- **M. Sc. in Physics,** (February 2005 – February 2008),
  Middle East Technical University,
  Advisor: Assoc. Dr. Yusuf İpekoğlu,
  Thesis title: Entanglement Measures.

- **B. Sc. in Physics,** (September 2000 – January 2005),
  Middle East Technical University.

- **B. Sc. in Electrical and Electronics Engineering,** (September 1999 – January 2005),
  Middle East Technical University.

Employment and Experience

- Teaching assistant in the Graduate School of Sciences of Middle East Technical University, January 2006 – April 2007.

Honors and Awards

- Scholarship for PhD study from The Scientific and Technological Research Council of Turkey, October 2008 – February 2013.
- Second place in the Aegean Region National Olimpiad in Informatics 1st stage examinations for high school students organized by The Scientific and Technological Research Council of Turkey, 1998.

Publications

1. K. Uyanik and S. Turgut, "A new sure-success generalization of Grover iteration and its application to weight decision problem of Boolean functions" ' Quantum Inf. Process **12:3395**, (2013) [arXiv:1301.4461 [quant-ph]].
2. K. Uyanik and S. Turgut, "Geometric measures of entanglement" Phys. Rev. A **81**, 032306 (2010) [arXiv:0910.1365 [quant-ph]].

Conferences, Workshops, Schools

- *Summer School on Fundamental Notions of Quantum Information Theory*, ITAP Turunç / Muğla TURKEY, July, 2009.
- *A Short Course on Supergravity*, Istanbul Center for Mathematical Sciences, TURKEY, February 05–06, 2009.
- *International Summer School on High Energy Physics: Standard Model and Beyond*, Akyaka/ Muğla TURKEY September, 2006.