Grover's Quantum Search Algorithm and Mixed States

Dan Kenigsberg

Grover's Quantum Search Algorithm and Mixed States

Research Thesis

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Computer Science

Dan Kenigsberg

Submitted to the Senate of
the Technion — Israel Institute of Technology
Heshvan 5762 Haifa October 2001

The research thesis was done under the supervision of Prof. Eli Biham in the Computer Science Department.

I thank Eli Biham for his guidance throughout the course of this research.

I would like to thank Tal Mor for fruitful discussions and useful references, Gilad, Ziv, Yuval, Philip and Ari for their companionship and coffee breaks, Tzafrir and Nadav for online Unix/LaTeX/Hebrew support, Omer for his help with Matlab, and Julia for the first template of this thesis. Ido asked not to mention his intelligent remarks in print, and thus is excluded from this list.

A special thank is sent to Avital for her love and encouragement, to my parents who brought me up, to my grandfather who taught me to read and do arithmetic, and to my grandmother who did not see me for weeks because of this work.

The generous financial help of Monte H. and Bertha Tyson Memorial Fellowship and the Technion is gratefully acknowledged.

Contents

Abstract **1**

1 Introduction **3**
- 1.1 Quantum Computation Overview 3
- 1.2 Basic Ingredients of Quantum Computation 6
  - 1.2.1 States 6
  - 1.2.2 Operations 7
  - 1.2.3 Measurement 8
  - 1.2.4 Mixed States 8
  - 1.2.5 Classical and Quantum Oracles 9
- 1.3 Well-Known Quantum Algorithms 10
  - 1.3.1 Deutsch-Jozsa 10
  - 1.3.2 Simon 11
  - 1.3.3 Shor 12
- 1.4 The Structure of this Thesis 13

2 Grover's Algorithm **14**
- 2.1 Classical Lower Bound 14
- 2.2 The Quantum Algorithm 15
- 2.3 Analysis 16
  - 2.3.1 Rotation about the Average 17
  - 2.3.2 Punctuated Execution 18
- 2.4 Algorithm Optimality 18

3 Generalizations **24**
- 3.1 Many Marked States 24
- 3.2 Unknown Number of Marked States 25
- 3.3 Arbitrary Pure Initial State 26

3.4 Amplitude Amplification 28
3.5 General Rotations 29
3.5.1 δ-sensitivity of ΔP 32
3.5.2 Finding a Marked State with Certainty 34
3.6 The Ultimate Generalization 35

4 Initialization with a Mixed State 37
4.1 Arbitrary Mixed Initial State 37
4.2 Examples 39
4.2.1 Pure Initial State 39
4.2.2 Pseudo-Pure Initial State 40
4.2.3 Initial State Where m of the Qubits Are Mixed 40
4.3 Algorithm Usefulness and Entropy 41

5 Summary and Conclusions 43

A Oracle Equivalence 45

Abstract in Hebrew ה

List of Figures

1.1 A classical reversible oracle 9
1.2 A regular quantum oracle 9
1.3 A controlled phase quantum oracle 10

3.1 The relation between β, γ, $\omega_\pm$, and δ 33
3.2 ΔP as a function of β and δ, with constant W_k 34
3.3 The number of eigenvalues which are not 1 as a function of n_a and n_b 36

4.1 The differences $P_i - \langle P_i \rangle$ as projections of a rotating $\Delta\vec{P}_i$ . . 38

5.1 Braunstein et al.'s separability bounds on the Bloch ball . . . 44

A.1 Construction of a controlled phase quantum oracle using a regular quantum oracle 46
A.2 Construction of a regular quantum oracle using a controlled phase quantum oracle 46

Abstract

Quantum computation is a field of computation theory that tries to find what can be computed, while taking into account the quantum nature of the physical world. The most celebrated achievement of quantum computation is Shor's algorithm, which factors large integers in polynomial time. Another achievement is Grover's unordered search algorithm, which finds a single "marked" element of a database in time which is in the order of the square root of the size of that database.

We commence the thesis with a brief introduction to that field of science. Then we present Grover's search algorithm, and a proof that it is the optimal search algorithm—it is better than any other algorithm, be it classical or quantum. The algorithm includes an initialization step and $O(\sqrt{N})$ iterations of selective phase inversions and Hadamard transforms.

The parameters of the algorithm have been generalized by various authors. Some have generalized the Grover Iterate, and some have generalized the initial state of the algorithm. We present these generalizations and provide an analysis of each of them in a uniform method. Then we show and discuss the most general extension to the Grover Iterate.

We present a new generalization of the initial state of the algorithm, in which it is allowed to be an arbitrary mixed quantum state. We show that even when the initial state is extremely mixed, there are cases where Grover's algorithm performs very well. We provide an approximation to the von Neumann entropy of pseudo-pure states, and we find that it grows smoothly with the level of mixedness of the pseudo-pure state. Combined with the previous result about the good performance of Grover's algorithm, our finding is in disagreement with Bose et al. We give a simple counter-example to their claim that for states with entropy larger than $\frac{1}{2}\log N$, Grover's algorithm is as bad as classical algorithms, and show where their mistake comes from.

We examine the usefulness of Grover's algorithm when initialized in a pseudo-pure state, and provide a measure for its effectiveness, including a threshold under which the algorithm is ineffective. We find that this threshold coincides with Braunstein et al.'s inseparability bound. This result may be considered as an evidence that entanglement is necessary for nontrivial quantum computation.

Chapter 1

Introduction

1.1 Quantum Computation Overview

This thesis resides in the realm of Quantum Computation—a relatively new field of science, combining Computation Theory and Quantum Mechanics. In the early 1980's, Richard Feynman pointed out that simulating a quantum mechanical system on a classical computer is a difficult task. It seems that no efficient polynomial reduction of quantum behavior to classical computation exists. The state of a quantum system comprising n 2-states subsystems belongs to a 2^n-dimensional complex space. Its evolution is controlled by a $(2^n \times 2^n)$-dimensional unitary matrix. Any approach taken to simulate its behavior, ended up with an exponential cost in terms of required time or acquired precision. Since some quantum systems can simulate each other efficiently, Feynman saw a hidden opportunity—maybe, quantum mechanics has a computational power not utilized by conventional computers.

Computation Theory is a field of computer science where computation models are designed and their power is then studied. For each computation model, computer scientists find a class of problems that it can solve—**RE**, **R** and $\mathbf{NP}^{\mathbf{NP}}$ are just few examples. An important question of Computation Theory is what can be computed realistically, with consideration of real-world limitations of time and space. The archetype of efficient and feasible computation model is the Polynomial-time Turing Machine. For a long time it was considered as the ultimate model of realistic computation. In fact, the strong version of the Church-Turing thesis asserts that "Any efficiently computable function can be computed by a Polynomial-time Turing

Machine". Later on, it has been noticed that if the machine is allowed to err, yet it produces the correct answer with some bounded probability, it seems to solve many additional problems efficiently. (By the way, as many other questions in Computation Theory, this question, too, is still open.)

The search for an ultimate computation model has led David Deutsch to ask what are the inherent physical limitation to the power of a real-world computation. Having asked that, he had devised a probably feasible, probably stronger computation model, that takes advantage of natural phenomena that previous models have ignored—Quantum Mechanics. Since a quantum system does not have to be in a specific state, and rather can be on a superposition of states, a Quantum Turing Machine could perform multiple calculations *simultaneously*.

In 1985 Deutsch gave the first demonstration of a task where a Quantum Turing Machine requires less steps to perform, compared to a classical Turing Machine. Together with Richard Jozsa, he later extended this task to a series of problems where a quantum computer has an advantage over a classical one. Yet the advantage was certainly sub-exponential, and the problem itself was neither interesting nor difficult.

Bernstein and Vazirani, followed by Daniel Simon [33], found problems which a quantum computer can solve efficiently, while a classical probabilistic computer cannot. The most dramatic spur in the field occurred when Peter Shor demonstrated in 1994 how a quantum computer could calculate the period of a function, and showed that utilizing this ability, it can solve the Factoring and Discrete Logarithm problems efficiently. Both problems are important and considered difficult to solve on a classical computer. Immediately afterwards, Lov Grover presented his quantum search algorithm, which we discuss in detail in this work. Its advantage was not as dramatic as the factoring algorithm's advantage, but it *proved* that for a large set of interesting problems, a quantum computer performs significantly better than a classical computer.

A serious obstacle to the implementation of a quantum computer was the question whether it can be built of simple gates, taken from a finite set of "building blocks". This has both experimental and theoretical consequences—with analogy to VLSI, building a complicated gate is unthinkable without reusing many instances of simple components. Respectively, if the simulation of a general quantum gate by simple ingredients is inefficient, one cannot consider the gates as a cheap resource. It was Deutsch again who provided in 1989 a 3-input 3-output universal gate, relying on the Toffoli universal gate for reversible computation. Later on, D. DiVincenzo showed [14] that a set of gates including one of many 2 bit gates and a single 1 bit phase rotation, is enough in order to achieve any unitary operation efficiently with reasonable (that is, polynomial) precision.

Another impediment to applied quantum computation is the inherent fragility of quantum-scale systems. Objects like photons, electrons, nuclei, atoms and molecules are vulnerable to external effects, such as changes in temperature, electro-magnetic field or vibrations. They have a tendency to emit energy spontaneously, change their state randomly, and *decohere*. It seemed that quantum information cannot be stored or managed in a controlled fashion. Similar difficulties in classical computation are tackled by measuring the stored information repeatedly and boosting problematic values. A complementary method is using error-correction codes. The first method is totally inappropriate to quantum information, since the act of measurement does exactly what we try to avoid: the original state is collapsed and destroyed. It was not at all certain whether the second method of error-correction code can be extended to the quantum regime, until in 1995 Peter Shor showed the existence of the first such code [31]. Later he described [30] how the error correction process can be performed fault-tolerantly.

Meanwhile, the study of Quantum Information per se had been developing. The basic ingredient of quantum information, the quantum bit or *qubit*, was defined, and interesting relations between classical and quantum information have been discovered. For example, classical information is easily duplicated, while the No-Cloning Theorem asserts that arbitrary quantum information cannot. Another fundamental result is the Holevo bound: A bit of quantum information comprises 2 independent real numbers, and a bit of classical information is only one binary digit. Yet no more than one classical bit can faithfully be encoded into and then extracted from a single qubit.

Therefore, it came as a surprise that in 1992 Charles Bennett and Stephen Wiesner found a method to transfer 2 bits encoded into one qubit. This result, called superdense coding, does not contradict Holevo, since in addition to the single qubit, it uses another resource: pre-shared pair of *entangled* qubits. Conversely, two classical bits and one pre-shared entangled pair can be used to transfer a single qubit, in a process called quantum teleportation. Amazingly, quantum teleportation is done without transfer of matter—nothing but information is exchanged—and without the sender having to know what is transferred. The question how these three kinds of resources are related (along with the need to quantify entanglement) is still open to date.

The great practical interest in Shor's algorithm stems from its ability to attack all prevalent public-key cryptosystems. A remedy to this is Charles Bennett and Gilles Brassard's quantum key exchange protocol, known as BB84. It allows two remote participants to exchange a private message secretly. Unlike classical key-exchange protocols, the security of BB84 does not assume anything on the computational power of the adversary, other than its existence in our quantum mechanical world.

1.2 Basic Ingredients of Quantum Computation

1.2.1 States

According to both classical and quantum viewpoints, a finite discrete physical system may be in one of N distinguishable *states*. In Quantum Theory, the ith state of these is denoted by $|i\rangle$. Much more importantly, Quantum Theory asserts that the system may be also in a *superposition* of states—a normalized linear combination as $|\psi\rangle = \sum_i \alpha_i |i\rangle$.

The most rudiment ingredient of classical computation models and realizations, is the *bit*—a system with two distinct states, conveniently called 0 and 1. Respectively, the basic component of quantum computation is the *qubit*—a system with two distinct states, $|0\rangle$ and $|1\rangle$. However, since it is a *quantum* bit, its state may be a superposition of these two distinguishable states: $|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle$. α_0 and α_1 are complex and satisfy the *normalization condition* $|\alpha_0|^2 + |\alpha_1|^2 = 1$. The state resides in a two-dimensional Hilbert space spanned by $|0\rangle$ and $|1\rangle$ over the complex number field $\mathbb{C}$.

A system consisting of n qubits is called a *quantum register*. If it is isolated from its environment, its state resides in a 2^n-dimensional Hilbert space, spanned by *the computation basis* $\{|i\rangle\}_{i=0}^{2^n-1}$. The qubits of the register are inter-connectable—without it, states like the "cat state" $\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|2^n - 1\rangle$ cannot be devised. Such states are called *entangled* states—they are multi-particle states and cannot be thought of as a combination of single qubit states.

The coefficient α_i is sometimes called the *amplitude* of $|i\rangle$. The description of a system by a state vector is slightly superfluous—the two vectors $|\psi\rangle$ and

$e^{i\theta}|\psi\rangle$ represent the same physical reality. These vectors are said to be equal up to a global phase.

1.2.2 Operations

The dynamics of the state of an isolated quantum register is ruled by unitary operations. A linear operation U is unitary iff its inverse exists and equals its Hermitian conjugate: $U^{\dagger}U = I$. Computation is no exception—every calculation applied to the quantum register has to be unitary. Fortunately, any classical computation can be extended efficiently to be unitary [15]. Moreover, every quantum computation may be realized as a series of applications of a 2-bit universal gate [14, 3].

Unitary operations are linear by definition. If an operation is known to produce the transformation

$$|i\rangle \xrightarrow{U_f} |f(i)\rangle$$

for all basis states $|i\rangle$, the transformation of any other state is well-defined:

$$|\psi\rangle = \sum_i \alpha_i |i\rangle \xrightarrow{U_f} \sum_i \alpha_i |f(i)\rangle.$$

All N values of $f(\cdot)$ are computed simultaneously at no extra cost. A delicate point, however, is that none of these values can be accessed with certainty. The function $f(\cdot)$ and the state $|\psi\rangle$ have to be chosen carefully in order to make use of this quantum parallelism.

The Hadamard Transform

A very useful operation in Quantum Computation is the Hadamard transform. On a single qubit it is defined by the matrix

$$H_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

expressed in the ordered basis $\{|0\rangle, |1\rangle\}$. Applying a Hadamard transform to an n-qubit register is defined as applying the single-qubit Hadamard transform to each of its qubits. Thus,

$$H_{2^n} = \frac{1}{\sqrt{2}} \begin{pmatrix} H_{2^{n-1}} & H_{2^{n-1}} \\ H_{2^{n-1}} & -H_{2^{n-1}} \end{pmatrix} = \underbrace{H_2 \otimes \cdots \otimes H_2}_{n}.$$

Equivalently, H_{2^n} may be defined as $\frac{1}{\sqrt{2^n}}\sum_{i=0}^{2^n}\sum_{j=0}^{2^n}(-1)^{i\cdot j}|i\rangle\langle j|$, where $i \cdot j$ denotes the inner product modulo 2 of the binary representations of i and j.

1.2.3 Measurement

In order to obtain information from a quantum system, we must apply a measurement to it. A measurement is usually an irreversible process, whose outcome cannot be predicted with certainty. If a quantum register in the state $|\psi\rangle = \sum_i \alpha_i|i\rangle$ undergoes a *complete measurement in the computation basis*, the result would be $|i\rangle$ with probability $|\alpha_i|^2$. Notice that for $e^{i\theta}|\psi\rangle$ the distribution of results is the same as for $|\psi\rangle$. This is in agreement with the assertion that a global phase has no physical meaning.

Our definition of measurement may seem excessively restrictive, since different kinds of measurement exist. For example, it is possible to measure only a few of the qubits, let the system evolve according to the outcome, and then measure the system again. It is also possible to perform the measurement in a basis other than the computation basis. However, it is well known [28] that a complete measurement in the computation basis is equivalent to the more general measurement methods, if we can add qubits to the register and can apply unitary operations to it.

1.2.4 Mixed States

For any superposition $|\psi\rangle = \sum_i \alpha_i|i\rangle$, there exists a unitary operation $U_{\psi\to 0}$ so that $U_{\psi\to 0}|\psi\rangle = |0\rangle$. Theoretically, if the parameters α_i are known, we could apply this $U_{\psi\to 0}$ to $|\psi\rangle$, and produce $|0\rangle$ with absolute certainty. This is why $|\psi\rangle$, and all the states mentioned in Subsection 1.2.1 are called *pure states*. However, in latter chapters of this work we discuss the evolution of a quantum register whose state is not completely known. At best we can describe its state as a probabilistic *ensemble* $\mathcal{E} = \{p_j, |\psi_j\rangle\}$ of pure states—it is in a pure state $|\psi_j\rangle$ with probability p_j. The *density matrix* notation summarizes this as the *mixed state*

$$\rho = \sum_j p_j|\psi_j\rangle\langle\psi_j|.$$

When a unitary operation is applied to a mixed state, it produces the transformation

$$\rho \xrightarrow{U} U\rho U^\dagger = \sum_j p_j U|\psi_j\rangle\langle\psi_j|U^\dagger.$$

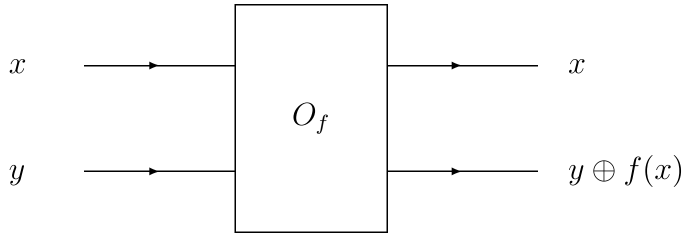

Figure 1.1: A classical reversible oracle
אורקל קלסי הפיך

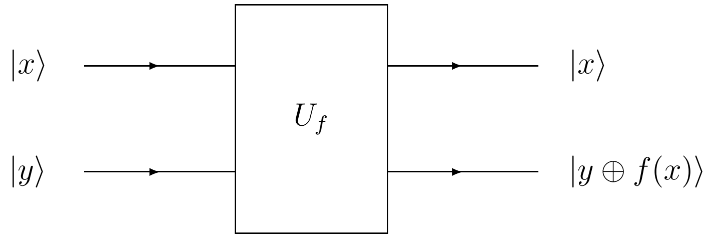

Figure 1.2: A regular quantum oracle
אורקל קוונטי רגיל

This implies that the time evolution of each $|\psi_j\rangle$ of the ensemble may be studied independently of other ensemble states.

A unitary operation cannot change the mixedness of a state—if ρ is pure ($\rho = |\psi\rangle\langle\psi|$), so is $U\rho U^\dagger$, and vice versa.

1.2.5 Classical and Quantum Oracles

A *classical oracle* of a function $f(\cdot)$ is a "black box" that when given a value x, computes $f(x)$. A special type of a classical oracle is the reversible oracle seen in Figure 1.1.

A *regular quantum oracle* is different from the reversible oracle[1] only in that it may be given a superposition of inputs $\sum_i a_i|i,0\rangle$, and produces a superposition of pairs $\sum_i a_i|i,f(i)\rangle$, as seen in Figure 1.2.

When f is binary, we define the *phase quantum oracle* as a black box that flips the phase of its input state $|x\rangle$ if and only if $f(x) = 1$. For

---

[1] Since quantum operations are unitary, a quantum oracle must be reversible.

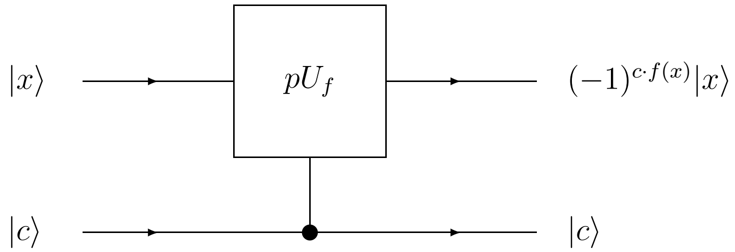

Figure 1.3: A controlled phase quantum oracle
אורקל פאזה קוונטי מותנה

every quantum operation there exists a *controlled* version, that performs the original operation if a control bit is $|1\rangle$ and acts as the identity operation otherwise. In particular, we consider the *controlled phase quantum oracle* outlined in Figure 1.3.

Regular quantum oracles and controlled phase quantum oracle are equivalent. Numerous sources such as [23], [27, page 249] or [28, page 277], show that a regular quantum oracle is reducible into an (uncontrolled) phase quantum oracle. However, we could not find any demonstration of the opposite direction of that reduction, although the equivalence is assumed in many of the proofs of the optimality of Grover's algorithm. Therefore, we devised our own proof, and provide it in Appendix A.

It is obvious that a quantum oracle of $f(\cdot)$ is not weaker than its classical counterpart—a quantum oracle fed with a single basis state (no superposition) easily simulates a classical oracle. Besides that, intuition tells us that since a quantum oracle executes multiple computations *simultaneously*, it is sharply *stronger* than a classical oracle.

1.3 Well-Known Quantum Algorithms

This section presents three famous quantum algorithms as examples of the potential advantage of quantum computers. The first two algorithms prove that a quantum oracle is indeed stronger than its classical counterpart.

1.3.1 Deutsch-Jozsa

This algorithm [12] solves an artificial problem, yet it is of interest since it was the first algorithm showing the excessive powers of a quantum oracle.

Assume we are presented with a phase quantum oracle U_f of a binary function $f : \{0,1\}^n \rightarrow \{0,1\}$. We are promised that $f(\cdot)$ is either constant ($\forall x : f(x) = c$) or balanced ($\big|\{x : f(x) = 0\}\big| = \big|\{x : f(x) = 1\}\big|$). The problem is to find whether f is constant or balanced.

The quantum algorithm to solve this problem requires an n-bit register and goes as follows:

1. Initialize the register to $H|0\rangle$ (the Hadamard transform applied to the zero state).
2. Apply the oracle U_f.
3. Apply the Hadamard transform H again.
4. Measure the result. Interpret $|0\rangle$ as constant, and anything else as balanced.

To understand the algorithm, we notice that in case the function is constant, the application of the oracle does not change the state of the register. If $c = 0$ no phase is flipped, and if $c = 1$ all phases are flipped, changing only the inconsequential global phase. Thus, the state of the register remains as the equal superposition of all states, and reapplication of H returns it to $|0\rangle$. In case the function is balanced, the probability to measure $|0\rangle$ can be verified to be zero.

The best classical algorithm for this problem is very similar to the following. Evaluate $f(x)$ on k randomly-selected x's. If they are all equal, answer "constant", and otherwise—answer "balanced". If the function is indeed constant, it never produces different values, and thus the result is correct. However, if the function is balanced, we might get k equal results in a row, and be mistaken with probability $\frac{1}{2^{k-1}}$. While the quantum algorithm required single oracle query to solve the problem with certainty, the classical algorithm requires more queries and solves the problem probabilistically.

1.3.2 Simon

Simon's algorithm [33] solves another artificial promise problem, but it does so with a spectacular speedup compared to classical counterparts. Assume we are presented with a quantum oracle U_f of a function $f : \{0,1\}^n \rightarrow \{0,1\}^n$ that is 2-to-1. We are assured that $\forall x : f(x) = f(x \oplus a)$ for some unknown constant a. The problem is to find that a.

The quantum algorithm to solve this problem requires two n-bit registers and goes as follows:

1. (a) Initialize the input register to $H|0\rangle$, and the output register to $|0\rangle$.
   (b) Apply the oracle U_f to the combined register.
   (c) Apply the Hadamard transform H to the input register.
   (d) Measure the input register. The result m_i satisfies $a \cdot m_i = 0$, where '$\cdot$' denotes the inner product mod 2.
2. Repeat these steps until acquiring n linearly independent m_i's. Extracting a from them is a straightforward polynomial classical process (Gauss-Jordan elimination) which requires no oracle queries.

The average number of repetition (and oracle queries) required to find the linearly independent set of m_i's is $O(n)$. An m with $m \cdot a = 1$ is never measured because of a special feature of the Hadamard transform—when $|x\rangle + |x \oplus a\rangle$ are transformed, their "odd" $|m\rangle$ elements cancel out. A classical algorithm requires $O(2^{n/2})$ oracle queries on average. Hence, in the case of this problem, quantum computation is exponentially faster than classical computation.

1.3.3 Shor

Shor's algorithm [32] solves two problem (Factorization and Discrete Logarithm) whose hardness is the core of the security of RSA [29] and Diffie-Hellman [13] cryptographic protocols, respectively. Shor showed that both the Factorization and the Discrete Logarithm problems are reducible to finding the period of a function. In the case of factorization, finding the prime factors p, q of N is equivalent to finding the period of $f_{N,a}(x) = a^x \bmod N$. (This is true for most $a \in \{2, \ldots N-1\}$.)

Further, Shor showed how a period can be found efficiently using a quantum computer. The key algorithm to perform this task is the Quantum Fourier Transform

$$\sum_x f(x)|x\rangle \overset{QFT}{\rightarrow} \sum_y \left(\frac{1}{\sqrt{N}} \sum_x e^{2\pi i x y/N} f(x) \right) |y\rangle.$$

When $|y\rangle$ is measured, the outcome is in the close vicinity of the period of $f(x)$ with high probability. Two other important aspects are answered by

Shor: how the initial distribution $\sum_x f(x)|x\rangle$ can be created efficiently for the given $f_{N,a}(x)$, and how to perform QFT efficiently. The overall complexity of Shor's algorithm is $O(\log^3 N)$ (polynomial in the number of bits), while the best known classical algorithm's complexity is $\Theta(e^{c(\log N)^{1/3}(\log\log N)^{2/3}})$ (super-polynomial in the number of bits).

1.4 The Structure of this Thesis

In Chapter 2 of this work we describe Grover's original search algorithm. We analyze it using the eigenstates of the Grover Iterate. We take great effort to prove the somewhat obvious classical lower bound, and provide one of the proofs of the optimality of Grover's algorithm.

In Chapter 3 we survey the various generalizations of the algorithm. We analyze them in a uniform method, and state what is the ultimate generalization. One of the results of our analysis of one of the generalizations was cited by [7], and we provide it in full details here.

Chapter 4 presents and analyzes a new generalization of the algorithm, where the quantum register is allowed to be initialized in an arbitrary mixed state. We provide an expression for the probability to measure a marked state as a function of time, and a measure for its effectiveness, including a threshold under which the algorithm is ineffective. We use our result to give a counter-example to one of the results of Bose et al. [8].

We conclude our work by explaining why it may be considered as a support to the common belief that entanglement is essential for true quantum computation.

In Appendix A we provide our proof of the equivalence of regular quantum oracles and phase quantum oracle. The proof we present of the optimality of the algorithm (and other proofs, too) makes use of this equivalence.

Chapter 2

Grover's Algorithm

Assume a binary function $f : \{0, \ldots, N-1\} \to \{0, 1\}$ such that

$$f(x) = \begin{cases} 1, & \text{if } x = k \\ 0, & \text{if } x \neq k \end{cases}$$

for some unknown k (which is selected uniformly in the range $\{0, ..., N - 1\}$). Assume further that we may use a classical oracle of $f(x)$. The search problem is to find this unknown marked value k. Our task is quite difficult—the best option is to try out all values in a random order. The expected number of oracle queries required to find the marked value is $\frac{N}{2}$. However, if we may use a *quantum* oracle, we can follow Grover's algorithm [18] and find the marked value after $\frac{\pi\sqrt{N}}{4}$ queries on average.

2.1 Classical Lower Bound

Let $\mathcal{A}_N$ be a classical search algorithm of size N. Let $E(\mathcal{A}_N)$ be the expected number of queries that $\mathcal{A}_N$ uses to find the marked value k.

Without loss of generality we assume that all the queries that $\mathcal{A}_N$ asks are different, as if two of the queries are identical, there exists an equivalent more efficient algorithm $\mathcal{B}$ that remembers the first answer and skips the second invocation of the query. Therefore, any lower bound we prove on $\mathcal{B}$ is true for $\mathcal{A}_N$, too.

By virtue of the fact that information about $f(\cdot)$ is accessible only through oracle queries, $\mathcal{A}_N$ can identify k only by querying $f(k)$. There is only one exception, once $N - 1$ failed queries are asked, the only untested value must be k since we know there is one. Therefore, $\mathcal{A}_N$ defines a series of N oracle queries, that is independent of the identity of k. Since the marked value k is selected uniformly from $\{0, ..., N-1\}$, the index i_k of the query $f(k)$ in the series of queries is also distributed uniformly. For $1 \leq t \leq N$, $P(i_k = t) = \frac{1}{N}$. $\mathcal{A}_N$ finds k after i_k queries, except for k such that $i_k = N$, where only $N-1$ queries are required. Thus, for any algorithm (deterministic or randomized), the expected number of required queries is at least

$$E(\mathcal{A}_N) = \sum_{k=1}^{N} P(i_k) i_k = \sum_{i_k=1}^{N-1} \frac{i_k}{N} + \frac{N-1}{N} = \frac{N+1}{2} - \frac{1}{N}.$$

2.2 The Quantum Algorithm

We assume along this thesis that the size of the search field $N = 2^n$ is an integral power of 2. A search problem with $f' : \{1, ..., N'\} \to \{0, 1\}$ where N' is not a power of 2, is easily reducible to a problem of size $N = 2^n$. All that has to be done is to define N as the next power of 2 and to define

$$f(x) = \begin{cases} f'(x), & 0 \leq x \leq N'-1 \\ 0, & N' \leq x \leq N-1 \end{cases} .$$

Since $N < 2N'$, the expected number of queries $\frac{\pi\sqrt{N}}{4} < \frac{\pi\sqrt{2N'}}{4}$ is at worst $\sqrt{2}$ times larger than the case of an integral power of 2.

In order to perform the algorithm we assume the existence of a quantum computer with a computation register of n qubits, and the existence of a quantum oracle for the function f. The algorithm is as follows:

1. Initialize the register to $H|0\rangle$. That is, reset all qubits to $|0\rangle$ and apply the Hadamard transform to each of them. This produces an equal superposition of all states in the computation basis $\frac{1}{\sqrt{N}} \sum_i |i\rangle$.

2. Repeat the following operation (named the *Grover Iterate Q*) $T = \frac{\pi\sqrt{N}}{4}$ times:

    (a) Rotate the marked state by a phase of π radians (I_f^{π}). This is done by a single application of the phase quantum oracle.

    (b) Apply the Hadamard transform on the register.

   (c) Rotate the $|0\rangle$ state by a phase of π radians (I_0^π).
   (d) Apply the Hadamard transform again.
   (e) Negate the total phase of the register (This step has no physical meaning, and only provides some aid for understanding).

3. Measure the resulting state.

2.3 Analysis

A thorough study of this algorithm appears in [19, 9]. The simplest analysis is done in vector notation. The algorithm is initialized with

$$|\psi_0\rangle = H|0\rangle = \frac{1}{\sqrt{N}}\sum_i |i\rangle, \tag{2.1}$$

and then

$$\begin{aligned} Q &= -HI_0^\pi H I_f^\pi \\ &= -H\left(I - 2|0\rangle\langle 0|\right) H \left(I - 2|k\rangle\langle k|\right) \\ &= -\left(I - 2H|0\rangle\langle 0|H\right)\left(I - 2|k\rangle\langle k|\right) \end{aligned} \tag{2.2}$$

is applied iteratively, where $|k\rangle$ is the marked state. Let us now define an orthonormal basis:

- $|k\rangle$ the marked state
- $|l\rangle = \frac{1}{\sqrt{N-1}}\sum_{i\neq k}|i\rangle$ (equal superposition of the unmarked states).
- Extend these two with additional $N-2$ orthonormal vectors.

It is easily verified that in this basis

$$Q = \begin{pmatrix} 1-\frac{2}{N} & 2\frac{\sqrt{N-1}}{N} & & & \\ -2\frac{\sqrt{N-1}}{N} & 1-\frac{2}{N} & & & \\ & & -1 & & \\ & & & \ddots & \\ & & & & -1 \end{pmatrix} \tag{2.3}$$

which is clearly a rotation matrix in the $(|k\rangle, |l\rangle)$ plane, with angle ω where $\cos\omega = 1 - \frac{2}{N}$, and a phase flip in the orthogonal subspace. For large N, ω

can be approximated as $\omega \approx \frac{2}{\sqrt{N}}$. The initial state

$$|\psi_0\rangle = \frac{1}{\sqrt{N}}|k\rangle + \frac{\sqrt{N-1}}{\sqrt{N}}|l\rangle \tag{2.4}$$

lies on the rotated plane, with angle $\phi \approx \frac{1}{\sqrt{N}}$ off the $|l\rangle$ axis. Thus,

$$|\psi_t\rangle = Q^t|\psi_0\rangle = \sin(\omega t + \phi)|k\rangle + \cos(\omega t + \phi)|l\rangle, \tag{2.5}$$

and the probability to measure the marked state

$$P(t) = |\langle k|\psi_t\rangle|^2 = \sin^2(\omega t + \phi) = \frac{1}{2} - \frac{1}{2}\cos(2\omega t + 2\phi) \tag{2.6}$$

reaches one when $2\omega t + 2\phi = \pi$, after $T \approx \frac{\pi\sqrt{N}}{4}$ iterations. Notice that T, the number of iterations needed to measure the marked state with certainty, is unlikely to be an integer. In the limit of large N, this is of no interest, since the $P(t)$ for the integer nearest to T is very close to 1. Yet the question of how to find the marked state with certainty [22, 25] *is* interesting to the physicists and engineers who are building the first small-scale implementations of the algorithm (and cf. Subsection 3.5.2).

2.3.1 Rotation about the Average

Along with the description of the algorithm as a rotation of a plane, there exists another description according to which the Grover Iterate $Q = -HI_0^\pi HI_f^\pi$ is broken into two operations: the oracle query I_f^π, and the rotation about the average operator $-HI_0^\pi H$. The latter operation is a rotation about the average due to the following. For arbitrary $|\psi\rangle = \sum_{i=0}^N a_i|i\rangle$ described in the computation basis, the amplitudes average is

$$\overline{a} = \frac{1}{N}\sum a_i = \frac{1}{N}\sum_{i=0}^{N}\langle i| \cdot \sum_{i=0}^{N} a_i|i\rangle = \frac{1}{\sqrt{N}}\langle 0|H|\psi\rangle. \tag{2.7}$$

To rotate all amplitudes about this average means to map

$$a_i \rightarrow a_i - 2(a_i - \overline{a}) = 2\overline{a} - a_i. \tag{2.8}$$

In vector notation it looks like

$$\langle i|\psi\rangle \rightarrow \frac{2}{\sqrt{N}}\langle 0|H|\psi\rangle - \langle i|\psi\rangle. \tag{2.9}$$

Notice that $\frac{1}{\sqrt{N}} = \langle i|H|0\rangle$, and therefore the mapping is

$$\langle i|\psi\rangle \rightarrow 2\langle i|H|0\rangle\langle 0|H|\psi\rangle - \langle i|\psi\rangle. \tag{2.10}$$

Since this is true for every $\langle i|$, it follows that the rotation about the average operation is after all $2H|0\rangle\langle 0|H - I = -HI_0^{\pi}H^{\dagger}$.

2.3.2 Punctuated Execution

One of the implications of the sinusoidal behavior of $P(t)$ is that when the $P(t)$ is near its maximal value, it doesn't improve much with every iteration[1]. Thus, we can improve the expected required number of queries if we are willing to cut short of the maximal probability and risk repeating the algorithm in case of failure. Consider executing the algorithm for t iterations. The expected number of queries required to measure a marked state is

$$\mathcal{T} = \frac{t}{P(t)} = \frac{t}{\sin^2(\omega t + \phi)}.$$

In the limit $N \gg 1$ we can find its minimum using derivative and numerical analysis. The minimum is reached when $2\omega t \approx \tan(\omega t)$ which is when $\omega t \approx 1.1656$ and

$$t_{\text{opt}} \approx \frac{0.742\pi}{4}\sqrt{N}.$$

At that point

$$\mathcal{T} \approx \frac{0.8785\pi}{4}\sqrt{N}, \tag{2.11}$$

which means a slight improvement comparing to the full-length execution.

2.4 Algorithm Optimality

Recently before the quantum search algorithm has been devised, Bennett, Bernstein, Brassard, and Vazirani [5] implicitly proved that a lower bound to the number of oracle queries required for unordered quantum search is $O(\sqrt{N})$. Thus, Grover's algorithm was known to be asymptotically optimal since its birth. This, of course, did not deter [9, 17, 35, 8, 4, 2] and others from providing additional proofs. Some of the proofs vary only in the extent of their similarity to the first one [5], their simplicity, and their tightness. Other approach the question of finding a lower bound in a more general way.

---

[1] This was first noted by Boyer et al. in [9], and later discussed in great detail by Gingrich et al. in [16].

In this section we shall restate Grover's proof [17] which appears to be the simplest. We (assisted by [27]) extend that proof slightly to include probabilistic search algorithms, as was first done by Boyer et al. in [9].

At first we describe the most general search algorithm $\mathcal{A}$: It may contain settings of qubits, unitary operations, oracle queries, and measurements. Without loss of generality we may assume that $\mathcal{A}$ begins with initialization of a quantum register, continues with unitary evolution of that register (including oracle queries), and ends with a measurement. Thus, any search algorithm $\mathcal{A}$ includes initialization to some oracle-independent state $|\psi(0)\rangle$ and unitary evolution of the form $U_T O U_{T-1} ... U_1 O U_0$, where O is an oracle query and $U_0 ... U_T$ are oracle-independent unitary operations. The algorithm concludes with a measurement. For $\mathcal{A}$ to be useful it has to find the marked state with a bounded probability by some p, no matter what is the identity of the marked state. For simplicity of the proof we assume $p = \frac{1}{2}$.

Consider different executions $\mathcal{A}_i$ of $\mathcal{A}$, each with a different marked state i. Let $|\psi_i(t)\rangle$ be the state of the quantum register after $U_t O_i ... U_1 O_i U_0$ is applied, when the marked state is i. O_i denotes an oracle query when the marked state is i. Let $|\psi_{null}(t)\rangle$ be the state of the quantum register after $U_t ... U_0$ is applied—with all oracle queries replaced by calls to the null oracle.

Definition 1 *The* Euclidean distance *between two states* $|\phi\rangle = \sum_i \alpha_i |i\rangle$ *and* $|\psi\rangle = \sum_i \beta_i |i\rangle$ *is defined by* $\| |\phi\rangle - |\psi\rangle \|^2 \triangleq \sum_i |\alpha_i - \beta_i|^2$.

Definition 2 *The spread of an execution* $\mathcal{A}_i$ *after* t *oracle queries is* $\Delta_i^2(t) \triangleq \big\| |\psi_i(t)\rangle - |\psi_{null}(t)\rangle \big\|^2$.

Definition 3 *The spread of* $\mathcal{A}$ *after* t *oracle queries is* $\Delta^2(t) \triangleq \sum_{i=0}^{N-1} \Delta_i^2(t)$.

We prove the following 3 lemmas about the spread:

Lemma 1 *The initial spread is zero.*

The proof is straightforward: before the first step, all executions are equal to the null execution, and

$$\Delta^2(0) = \sum_i \big\| |\psi_i(0)\rangle - |\psi_{null}(0)\rangle \big\|^2 = 0$$

Lemma 2 *In order to meet the bounded probability of success, the spread has to be* $\Omega(N)$.

Proof We know that the algorithm is successful with probability higher than $\frac{1}{2}$. That is, for every execution, when the register is finally measured at time T, the marked state is found with probability $> \frac{1}{2}$:

$$|\langle\psi_i|i\rangle|^2 \geq \frac{1}{2}. \tag{2.12}$$

(Throughout this proof we discuss only the last step and therefore drop the (T) qualifier for clarity.) The global phase of the final state of each execution has no physical meaning, and thus may be assumed to satisfy $\langle\psi_i|i\rangle = |\langle\psi_i|i\rangle|$. Therefore, we may state that

$$\begin{aligned} \alpha^2(T) &\triangleq \sum_i \| |\psi_i\rangle - |i\rangle \|^2 \\ &= \sum_i 2 - 2\,\mathrm{Re}\langle\psi_i|i\rangle = \sum_i 2 - 2|\langle\psi_i|i\rangle| \end{aligned}$$

and from (2.12),

$$\leq 2N - \sqrt{2}N = (2-\sqrt{2})N. \tag{2.13}$$

Using Lagrange multipliers we can show that for real a_i's and b_i's and under the constraint of $\sum_i |a_i|^2 + |b_i|^2 = 1$, the expression

$$\sum_i a_i \leq \sqrt{N}$$

reaches its maximum when $a_i = \frac{1}{\sqrt{N}}$ and $b_i = 0$. Therefore, for every state, and specifically for $|\psi_{null}\rangle$

$$\begin{aligned} \beta^2(T) &\triangleq \sum_i \left\| |\psi_{null}\rangle - |i\rangle \right\|^2 \\ &= 2N - 2\,\mathrm{Re}\sum_i \langle i|\psi_{null}\rangle \\ &\geq 2N - 2\sqrt{N} = 2N(1 - \frac{1}{\sqrt{N}}). \end{aligned} \tag{2.14}$$

Using Cauchy-Schwarz inequality (that states that $\mathrm{Re}\langle x|y\rangle \leq \|x\|^2 \|y\|^2$) and by definition,

$$\begin{aligned}
\Delta^2(T) &= \sum_i \left\| |\psi_i\rangle - |\psi_{null}\rangle \right\|^2 \\
&= \sum_i \left\| |\psi_i\rangle - |i\rangle + |i\rangle - |\psi_{null}\rangle \right\|^2 \\
&= \alpha^2 - 2\,\mathrm{Re}\sum_i \langle \psi_i - i|\psi_{null} - i\rangle + \beta^2 \\
&\geq \alpha^2 - 2\sum_i \left\| |\psi_i\rangle - |i\rangle \right\| \left\| |\psi_{null} - |i\rangle\rangle \right\| + \beta^2,
\end{aligned}$$

and applying Cauchy-Schwarz inequality again we arrive at

$$\begin{aligned}
&\geq \alpha^2 - 2\sqrt{\alpha^2\beta^2} + \beta^2 \\
&= (\beta - \alpha)^2.
\end{aligned}$$

From (2.13) and (2.14) we obtain

$$\begin{aligned}
\alpha &\leq \sqrt{N}\sqrt{2-\sqrt{2}} < \sqrt{N}, \\
\beta &\geq \sqrt{2}\sqrt{N}\sqrt{1 - \frac{1}{\sqrt{N}}} \\
&> \frac{2}{\sqrt{3}}\sqrt{N} \text{ (for } N \geq 9)
\end{aligned}$$

and conclude that

$$\begin{aligned}
\Delta^2(T) &= (\beta - \alpha)^2 \\
&> N(\frac{2}{\sqrt{3}} - 1)^2 = O(N).
\end{aligned}$$

■

Lemma 3 *The spread grows not faster than* $o(t^2)$.

Proof At first we notice that

$$\begin{aligned}
\Delta_i^2(t) &= \left\| |\psi_i(t)\rangle - |\psi_{null}(t)\rangle \right\|^2 \\
&= \left\| U_i \left[O_i|\psi_i(t-1)\rangle - |\psi_{null}(t-1)\rangle \right] \right\|^2 \\
&= \left\| O_i|\psi_i(t-1)\rangle - |\psi_{null}(t-1)\rangle \right\|^2 \\
&= 2 - 2\,\mathrm{Re}\langle \psi_{null}(t-1)|O_i|\psi_i(t-1)\rangle
\end{aligned}$$

since the Euclidean distance is invariant under unitary transformations. Thus,

$$\begin{aligned}\Delta_i^2(t) - \Delta_i^2(t-1) &= 2\,\mathrm{Re}[\langle\psi_{null}(t-1)|\psi(t-1)\rangle \\ &\qquad -\langle\psi_{null}(t-1)|O_i|\psi_i(t-1)\rangle] \\ &= 2\,\mathrm{Re}\textstyle\sum_j\langle\psi_{null}(t-1)|j\rangle[\langle j|\psi(t-1)\rangle \\ &\qquad\quad -\langle j|O_i|\psi(t-1)\rangle].\end{aligned} \tag{2.15}$$

Without loss of generality (Cf. Appendix A) we assume that O_i is a phase oracle, that is,

$$\langle j|O_i|\psi\rangle = \begin{cases} \langle j|\psi\rangle, & \text{if } j \neq i \\ -\langle j|\psi\rangle, & \text{if } j = i \end{cases}, \tag{2.16}$$

and therefore

$$\begin{aligned}\Delta_i^2(t) - \Delta_i^2(t-1) &= 4\,\mathrm{Re}\langle\psi_{null}(t-1)|i\rangle\langle i|\psi_i(t-1)\rangle \\ &\leq 4|\langle i|\psi_{null}(t-1)\rangle||\langle i|\psi_i(t-1)\rangle|.\end{aligned}$$

After subtracting and adding $\langle i|\psi_{null}\rangle$, and dropping the $(t-1)$ qualifier, this becomes

$$= 4|\langle i|\psi_{null}\rangle|\ |\langle i|\psi_i\rangle - \langle i|\psi_{null}\rangle + \langle i|\psi_{null}\rangle|,$$

which is (by the triangle inequality),

$$\begin{aligned}&\leq 4|\langle i|\psi_{null}\rangle|\ \ [|\langle i|\psi_i\rangle - \langle i|\psi_{null}\rangle| + |\langle i|\psi_{null}\rangle|] \\ &= 4|\langle i|\psi_{null}\rangle|\ |\langle i|\psi_{null}\rangle - \langle i|\psi_i\rangle| + 4|\langle i|\psi_{null}\rangle|^2.\end{aligned}$$

Now notice that for any a, b and any positive λ, $(\lambda|a|-|b|)^2 \geq 0$ and therefore

$$4|a||b| \leq 2\lambda|a|^2 + \frac{2}{\lambda}|b|^2. \tag{2.17}$$

Summing over all i's, and applying this inequality,

$$\begin{aligned}\Delta^2(t) - \Delta^2(t-1) &= \sum_i \Delta_i^2(t) - \Delta_i^2(t-1) \\ &\leq 4\sum_i |\langle i|\psi_{null}\rangle|^2 + 2\lambda\sum_i |\langle i|\psi_{null}\rangle|^2 \\ &\quad + \frac{2}{\lambda}\sum_i |\langle i|\psi_{null}\rangle - \langle i|\psi_i\rangle|^2\end{aligned}$$

and since $|\langle i|\psi_{null}\rangle - \langle i|\psi_i\rangle|^2 < \Delta_i^2(t-1)$,

$$\leq 4 + 2\lambda + \frac{2}{\lambda}\Delta^2(t-1). \tag{2.18}$$

Substituting for $\lambda = \Delta(t-1)$, this becomes

$$\Delta^2(t) - \Delta^2(t-1) \leq 4\Delta(t-1) + 4,$$

which is equivalent to

$$\Delta(t) \leq \Delta(t-1) + 2.$$

Since $\Delta(0) = 0$, this means that

$$\Delta^2(t) \leq 4t^2.$$

■

Together, these lemmas infer that $\mathcal{A}$ requires $\Omega(\sqrt{N})$ queries in order to fulfill its task. The first proof following this structure appeared as one of the weaknesses of quantum computing in Bennett et al.'s paper [5], before Grover's algorithm has been devised. Zalka's proof [35] follows that structure, too, and gives the exact number of queries required to obtain certain probability p.

Bose, Rallan and Vedral [8] give a similar, yet different proof. Instead of following the evolution of the spread of the algorithm, they follow the entropy of the quantum register (See Eq. 4.8). They prove three properties of the entropy:

1. The initial entropy is zero, since the initial state is known to be $H|0\rangle$.
2. At the end of the algorithm the entropy must be $\log N$, since the final state is the marked state, and there are N different possible marked states occurring with equal probability $\frac{1}{N}$. (Bose et al. do not discuss probabilistic algorithms, where the final state does not have 1-to-1 mapping with the marked state.)
3. The entropy can grow by no more than $\frac{3}{\sqrt{N}} \log N$ with each oracle query.

And again, the conclusion is that $\Omega(\sqrt{N})$ queries are required.

Chapter 3

Generalizations

Grover's original algorithm described in Section 2.2 is quite restrictive. In order to make it more useful, to study its resistance to noise and to learn where its power stems from, the algorithm was generalized by many people. In this chapter we survey all known generalizations, and find the most general one.

3.1 Many Marked States

The simplest generalization of Grover's algorithm is to consider a function where the number of x's satisfying $f(x) = 1$ is $r > 1$. Not surprisingly, Boyer et al. show in [9] that finding one of the r marked x's is easier when r is large, and the difficult search problem is when $r \ll N$. Grover's original algorithm requires almost no changes in order to solve the problem of multiple marked states. The only element that changes is the required number of iterations—with $r > 1$ it is $T = \frac{\pi}{4}\sqrt{\frac{N}{r}}$. The analysis of the algorithm is very similar, too, yet we have to be more careful while selecting the convenient orthonormal basis. The N-dimensional Hilbert space of the register states may be broken into two subspaces: $\mathcal{K}$ of r dimensions, spanned by the marked states, and $\mathcal{L}$ of $N - r$ dimensions, spanned by the unmarked states. The first basis element for $\mathcal{K}$ is

$$|k\rangle \triangleq \frac{1}{\sqrt{r}} \sum_{i \in M} |i\rangle, \tag{3.1}$$

where M is the set of all marked states. The rest of the basis is any orthonormal extension denoted by $\{|k_1\rangle \dots |k_{r-1}\rangle\}$. A basis for $\mathcal{L}$ is

$$|l\rangle \triangleq \frac{1}{\sqrt{N-r}} \sum_{i \notin M} |i\rangle. \tag{3.2}$$

and any orthonormal basis extending it, denoted by $\{|l_1\rangle \dots |l_{N-r-1}\rangle\}$. In the ordered basis $(|k\rangle, |l\rangle, |k_1\rangle \dots |k_{r-1}\rangle, |l_1\rangle \dots |l_{N-r-1}\rangle)$,

$$Q = \begin{pmatrix} 1-\frac{2r}{N} & 2\frac{\sqrt{r(N-r)}}{N} & & & & & & \\ -2\frac{\sqrt{r(N-r)}}{N} & 1-\frac{2r}{N} & & & & & & \\ & & \overbrace{1 \qquad\quad}^{r-1} & & & & & \\ & & & \ddots & & & & \\ & & & & 1 & & & \\ & & & & & \overbrace{-1 \qquad\qquad}^{N-r-1} & & \\ & & & & & & \ddots & \\ & & & & & & & -1 \end{pmatrix} \tag{3.3}$$

Ignoring the partial phase flip, Q is a 2-dimensional rotation with angle ω, where $\cos\omega = 1 - \frac{2r}{N}$. The initial state lies with angle ϕ to the $|l\rangle$ axis, where

$$\cos\phi = \langle 0|H|l\rangle = \sqrt{\frac{N-r}{N}}.$$

For $r \ll N$, $\omega \approx 2\sqrt{\frac{r}{N}}$ and $\phi \approx \sqrt{\frac{r}{N}}$. Thus, the probability to measure a marked space is again

$$P(t) = \frac{1}{2} - \frac{1}{2}\cos(2\omega t + 2\phi). \tag{3.4}$$

Notice that the best time to measure, $T \approx \frac{\pi}{4}\sqrt{\frac{N}{r}}$ depends on r. This is why Gilles Brassard compares quantum search to baking a soufflé—if you don't know when it is ready, you cannot enjoy it.

3.2 Unknown Number of Marked States

Another interesting question addressed by [9] is what happens when r is unknown. Is it then impossible to bake the soufflé? They found out that it is possible, and with no asymptotic penalty—it requires only $O(\sqrt{\frac{N}{r}})$ queries. All that is needed to accomplish that, is to start with a single application of the Grover Iterate. If the marked state is not found, repeat the algorithm with a number of iterations which is $\frac{6}{5}$ times the previous number. Stop when the number of iteration has reached $\frac{\pi}{4}\sqrt{N}$.

We give here a proof for a weaker result: When $r \ll N$, the expected number of queries to find a marked state is $O(\sqrt{N})$. Choose $t \in_R \{1 \ldots \sqrt{N}\}$ and run t iterations of Grover's algorithm. Since $P(t)$ behaves like a smooth cosine (3.4), the expected value of $P(t)$ is

$$\frac{\int_0^{\sqrt{N}} P(t')dt'}{\sqrt{N}} = \frac{1}{2} + \frac{\sin 2\omega\sqrt{N}}{4\omega\sqrt{N}} > \frac{1}{2} - \frac{1}{4\omega\sqrt{N}} = \frac{1}{2} - \frac{1}{8\sqrt{r}} > \frac{3}{8} > \frac{1}{3}. \quad (3.5)$$

Thus, no more than 3 choices of random t are required on average, and a marked state is found in $O(\sqrt{N})$ queries.

3.3 Arbitrary Pure Initial State

If Grover's search algorithm is used as a procedure by another algorithm, it might be necessary to avoid its initialization step. Even if the initialization is performed, gate imperfection or external noise might cause the outcome to differ from the exact $H|0\rangle$ state. Rather, it may well be some general pure state $|\psi_0\rangle$, which is a superposition of marked states and unmarked states. The first to address this problem were Biham et al. in [6], yet here we take another path taken by [16], and similar to [10].

Let us represent $|\psi_0\rangle$ in the orthonormal basis of Section 3.1 (which is independent of the initial state)

$$\begin{aligned} |\psi_0\rangle &= |k\rangle\langle k|\psi_0\rangle + |l\rangle\langle l|\psi_0\rangle + \sum_{i=1}^{r-1} |k_i\rangle\langle k_i|\psi_0\rangle + \sum_{i=1}^{N-r-1} |l_i\rangle\langle l_i|\psi_0\rangle \\ &= A_k e^{i\theta_k}|k\rangle + A_l e^{i\theta_l}|l\rangle + \sqrt{r}\sigma_k|\psi_{0k}\rangle + \sqrt{N-r}\sigma_l|\psi_{0l}\rangle \qquad (3.6) \end{aligned}$$

where $A_k e^{i\theta_k}$ and $A_l e^{i\theta_l}$ are defined as the unique polar representation of $\langle k|\psi_0\rangle$ and $\langle l|\psi_0\rangle$, respectively. σ_k, $|\psi_{0k}\rangle$, σ_l and $|\psi_{0l}\rangle$ are defined such that $\sqrt{r}\sigma_k|\psi_{0k}\rangle$ and $\sqrt{N-r}\sigma_l|\psi_{0l}\rangle$ are the unique representation of their respective terms as norm and normalized state.

The Grover Iterate is, of course, independent of the initial state. Thus, Q keeps the form of a 2-dimensional rotation matrix (3.3). This means that (except for signs) only the projections of $|\psi_0\rangle$ on $|k\rangle$ and $|l\rangle$ are affected by Q, and that the state of the quantum register as a function of time is

$$\begin{aligned}
|\psi(t)\rangle &= Q^t|\psi_0\rangle \\
&= \sqrt{r}\sigma_k|\psi_{0k}\rangle + (-1)^t\sqrt{N-r}\sigma_l|\psi_{0l}\rangle + A_k e^{i\theta_k}Q^t|k\rangle + A_l e^{i\theta_l}Q^t|l\rangle \\
&= \sqrt{r}\sigma_k|\psi_{0k}\rangle + (-1)^t\sqrt{N-r}\sigma_l|\psi_{0l}\rangle \\
&\quad + \left(A_k e^{i\theta_k}\cos\omega t + A_l e^{i\theta_l}\sin\omega t\right)|k\rangle \\
&\quad + \left(A_l e^{i\theta_l}\cos\omega t - A_k e^{i\theta_k}\sin\omega t\right)|l\rangle \qquad (3.7)
\end{aligned}$$

Let $P_{\mathcal{K}}$ be the operator of projection on the marked subspace. The probability to measure a marked state as a function of the number of Grover iterations t is the squared amplitude of $P_{\mathcal{K}}|\psi(t)\rangle$:

$$\begin{aligned}
P(t) &= \left|P_{\mathcal{K}}|\psi(t)\rangle\right|^2 \\
&= r\sigma_k^2\langle\psi_{0k}|\psi_{0k}\rangle + \frac{A_k^2+A_l^2}{2} - \frac{1}{2}\left|A_k^2e^{2i\theta_k} + A_l^2e^{2i\theta_l}\right|\cos(2\omega t + 2\phi_{\psi_0}) \\
&= \langle P_{\psi_0}\rangle - \Delta P_{\psi_0}\cos(2\omega t + 2\phi_{\psi_0}). \qquad (3.8)
\end{aligned}$$

where

$$\begin{aligned}
\tan 2\phi_{\psi_0} &= \frac{2A_kA_l\cos(\theta_l-\theta_k)}{A_k^2-A_l^2}, \\
\langle P_{\psi_0}\rangle &= r\sigma_k^2 + \frac{A_k^2+A_l^2}{2},
\end{aligned}$$

and

$$\Delta P_{\psi_0} = \frac{1}{2}\left|A_k^2e^{2i\theta_k} + A_l^2e^{2i\theta_l}\right|.$$

The subscripts ψ_0 denote that the values depend on the initial state. ω is independent of the initial state, and as explained in Section 3.1, may be approximated by $\omega = 2\sqrt{\frac{r}{N}}$. The probability to measure the marked state reaches its maximum after $T = \frac{\pi-2\phi}{2\omega}$ iterations.

Our results are in agreement with previous works. For example, by our definition

$$\sqrt{r}\sigma_k|\psi_{0k}\rangle = P_{\mathcal{K}}(|\psi_0\rangle - |k\rangle\langle k|\psi_0\rangle).$$

This means that

$$\sigma_k^2 = \sigma_k^2 \langle \psi_{0k} | \psi_{0k} \rangle = \frac{1}{r} \sum_{i \in M} |\langle i | \psi_0 \rangle - \langle i | k \rangle \langle k | \psi_0 \rangle|^2 = \frac{1}{r} \sum_{i \in M} |\langle i | \psi_0 \rangle - \bar{k}|^2 \quad (3.9)$$

is the variance of $\langle i|\psi_0\rangle$'s. This is exactly the definition of σ_k in [6]. Another example is the original case studied by Grover, for which we find $\left\langle P_{H|0\rangle} \right\rangle = \Delta P_{H|0\rangle} = \frac{1}{2}$ and $\phi_{H|0\rangle} \approx 0$. The maximum probability (≈ 1) is reached after $T = \frac{\pi\sqrt{N}}{4}$ iterations.

3.4 Amplitude Amplification

Lov Grover [20, 21] reported that his algorithm works well when the Hadamard transform is replaced by almost any other unitary operation. Brassard, Høyer, Mosca and Tapp took a slightly different approach in [10], where they call this idea "Amplitude Amplification". They consider a search problem with r marked states, and assume a quantum algorithm $\mathcal{A}$ that, when initialized by $|0\rangle$, finds one of the marked states with probability

$$W_k \triangleq \sum_{i \in M} |\langle i | \mathcal{A} | 0 \rangle|^2$$

and fails with probability $W_l \triangleq 1 - W_k$. They showed that Grover's algorithm may be used to amplify this amplitude if it is initialized with $\mathcal{A}|0\rangle$, and the Grover Iterate is modified to $Q = -\mathcal{A} I_0^\pi \mathcal{A}^\dagger I_f^\pi$. After applying Q iteratively $\frac{\pi}{4\sqrt{W_k}}$ times, a marked state is found with almost certainty. In this perspective, the Hadamard transform in the original algorithm is a "blind guess" $O(N)$ algorithm. Gingrich et al. [16] combined this generalization with the arbitrary initial state. We follow their analysis, since it hardly differs from that of Section 3.3. We start by defining

$$|k\rangle \triangleq \frac{1}{\sqrt{W_k}} \sum_{i \in M} |i\rangle \langle i | \mathcal{A} | 0 \rangle, \quad (3.10)$$

and

$$|l\rangle \triangleq \frac{1}{\sqrt{W_l}} \sum_{i \notin M} |i\rangle \langle i | \mathcal{A} | 0 \rangle. \quad (3.11)$$

$\{|k_i\rangle\}_{i=1}^{r-1}$ and $\{|l_i\rangle\}_{i=1}^{N-r-1}$ are defined accordingly to complete an orthonormal basis. In this basis Q continues to be a rotation matrix, but now with angle $\omega \approx 2\sqrt{W_k}$. Similarly to (3.6) we represent $|\psi_0\rangle$ in this new basis

$$\begin{aligned}
|\psi_0\rangle &= |k\rangle\langle k|\psi_0\rangle + |l\rangle\langle l|\psi_0\rangle + \sum_{i=1}^{r-1} |k_i\rangle\langle k_i|\psi_0\rangle + \sum_{i=1}^{N-r-1} |l_i\rangle\langle l_i|\psi_0\rangle \\
&= A_k e^{i\theta_k}|k\rangle + A_l e^{i\theta_l}|l\rangle + \sqrt{W_k}\sigma_k|\psi_{0k}\rangle + \sqrt{W_l}\sigma_l|\psi_{0l}\rangle
\end{aligned}$$

and re-define $A_k e^{i\theta_k}$, $A_l e^{i\theta_l}$, σ_k, $|\psi_{0k}\rangle$, σ_l and $|\psi_{0l}\rangle$ accordingly. Notice that now

$$\sigma_k^2 = \frac{1}{W_k}\sum_{i\in M} |\langle i|\mathcal{A}|0\rangle|^2 \left|\frac{\langle i|\psi_0\rangle}{\langle i|\mathcal{A}|0\rangle} - \frac{1}{W_k}\sum_{j\in M}\langle 0|\mathcal{A}|j\rangle\langle j|\psi_0\rangle\right|^2$$

is the weighted variance of $\langle i|\psi_0\rangle$'s, as defined in [7, Eq. (3.48)].

$P(t)$ still adheres to Eq. (3.8), only that now $\langle P_{\psi_0}\rangle = W_k\sigma_k^2 + \frac{A_k^2+A_l^2}{2}$.

3.5 General Rotations

The original Grover Iterate $-HI_0H^\dagger I_f$ has been further generalized to $-UI_s^\beta U^\dagger I_f^\gamma$ by Biham et al. [7]. That is, they use some arbitrary transformation U instead of Hadamard (this is equivalent to $\mathcal{A}$ in Section 3.4), arbitrary "pivot" state $|s\rangle$ instead of $|0\rangle$, and arbitrary phase rotations (β and γ) instead of phase inversion. Actually, replacing $|0\rangle$ with $|s\rangle$ is superfluous. Let

$$X_{s0} = I - |0\rangle\langle 0| - |s\rangle\langle s| + |0\rangle\langle s| + |s\rangle\langle 0|$$

be the swap operation between $|0\rangle$ and $|s\rangle$. For any U and $|s\rangle$ there exists some $V = UX_{s0}$ such that

$$\begin{aligned}
VI_0^\beta V^\dagger &= V\left(I - (1-e^{i\beta})|0\rangle\langle 0|\right)V^\dagger \\
&= U\left(I - (1-e^{i\beta})X_{s0}|0\rangle\langle 0|X_{s0}^\dagger\right)U^\dagger \\
&= U\left(I - (1-e^{i\beta})|s\rangle\langle s|\right)U^\dagger \\
&= UI_s^\beta U^\dagger.
\end{aligned}$$

Their analysis of the generalized algorithm has been done in the method of recursion equations, which they also used in [6]. We show that the vector notation analysis, which is used throughout this work, is applicable too.

When represented in the orthonormal basis of the previous section, the generalized Grover Iterate is

$$Q = \begin{pmatrix} (M) & \overbrace{\hphantom{-e^{i\gamma}\quad\ddots\quad -e^{i\gamma}}}^{r-1} & & & & & \\ & -e^{i\gamma} & & & & & \\ & & \ddots & & & & \\ & & & -e^{i\gamma} & \overbrace{\hphantom{-1\quad\ddots\quad -1}}^{N-r-1} & & \\ & & & & -1 & & \\ & & & & & \ddots & \\ & & & & & & -1 \end{pmatrix}. \tag{3.12}$$

The topmost 2-dimensional component of Q (in the $\{|k\rangle, |l\rangle\}$ subspace) is

$$\begin{aligned} M &= \begin{pmatrix} (1-e^{i\beta})e^{i\gamma}W_k - e^{i\gamma} & \sqrt{W_k W_l}(1-e^{i\beta}) \\ \sqrt{W_k W_l}e^{i\gamma}(1-e^{i\beta}) & (1-e^{i\beta})W_l - 1 \end{pmatrix} \\ &= \begin{pmatrix} a & b\sqrt{\frac{W_k}{W_l}} \\ c\sqrt{\frac{W_l}{W_k}} & d \end{pmatrix}. \end{aligned} \tag{3.13}$$

Since the elements of M are potentially complex, it cannot be interpreted immediately as a rotation matrix. In order to analyze Q, we first diagonalize M with $M = SDS^\dagger$, and write (as in [7]),

$$D = \begin{pmatrix} \lambda_+ & \\ & \lambda_- \end{pmatrix} = \begin{pmatrix} e^{i\omega_+} & \\ & e^{i\omega_-} \end{pmatrix},$$

where

$$\omega_\pm = \pi + \frac{\beta+\gamma}{2} \pm \omega$$

and

$$\cos\omega = W_k \cos\frac{\beta+\gamma}{2} + W_l \cos\frac{\beta-\gamma}{2}. \tag{3.14}$$

The eigenvectors of M, and their conjugates, are

$$|\Psi_+\rangle = \frac{\sqrt{\frac{W_k}{W_l}}b|k\rangle + (\lambda_+ - a)|l\rangle}{\sqrt{\frac{W_k}{W_l}|b|^2 + |\lambda_+ - a|^2}},$$

$$\begin{aligned}
|\Psi_-\rangle &= \frac{\sqrt{\frac{W_k}{W_l}}b|k\rangle + (\lambda_- - a)|l\rangle}{\sqrt{\frac{W_k}{W_l}|b|^2 + |\lambda_- - a|^2}}, \\
\langle\Psi_+| &= \frac{\sqrt{\frac{W_k}{W_l}}b^*\langle k| + (\lambda_+^* - a^*)\langle l|}{\sqrt{\frac{W_k}{W_l}|b|^2 + |\lambda_+ - a|^2}} \\
&= \frac{\sqrt{\frac{W_k}{W_l}|b|^2 + |\lambda_+ - a|^2}}{\lambda_- - \lambda_+}\left(\frac{\lambda_- - a}{b\sqrt{\frac{W_k}{W_l}}}\langle k| - \langle l|\right),
\end{aligned}$$

and

$$\begin{aligned}
\langle\Psi_-| &= \frac{\sqrt{\frac{W_k}{W_l}}b^*\langle k| + (\lambda_-^* - a^*)\langle l|}{\sqrt{\frac{W_k}{W_l}|b|^2 + |\lambda_- - a|^2}} \\
&= \frac{\sqrt{\frac{W_k}{W_l}|b|^2 + |\lambda_- - a|^2}}{\lambda_- - \lambda_+}\left(-\frac{\lambda_+ - a}{b\sqrt{\frac{W_k}{W_l}}}\langle k| + \langle l|\right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
Q^t|\psi_0\rangle &= e^{i\omega_+ t}|\Psi_+\rangle\langle\Psi_+|\psi_0\rangle + e^{i\omega_- t}|\Psi_-\rangle\langle\Psi_-|\psi_0\rangle \\
&\quad + (-1)^t e^{i\gamma t}\sqrt{W_k}\sigma_k|\psi_{0k}\rangle + (-1)^t\sqrt{W_l}\sigma_l|\psi_{0l}\rangle \qquad (3.15)
\end{aligned}$$

and its projection on the marked subspace is

$$\begin{aligned}
P_{\mathcal{K}}Q^t|\psi_0\rangle &= e^{i\omega_+ t}|k\rangle\langle k|\Psi_+\rangle\langle\Psi_+|\psi_0\rangle + e^{i\omega_- t}|k\rangle\langle k|\Psi_-\rangle\langle\Psi_-|\psi_0\rangle \\
&\quad + (-1)^t e^{i\gamma t}\sqrt{W_k}\sigma_k|\psi_{0k}\rangle.
\end{aligned}$$

The probability to measure a marked state, which is the squared amplitude of this projection, is

$$\begin{aligned}
P(t) &= \left|e^{i\omega_+ t}\langle k|\Psi_+\rangle\langle\Psi_+|\psi_0\rangle + e^{i\omega_- t}\langle k|\Psi_-\rangle\langle\Psi_-|\psi_0\rangle\right|^2 + W_k\sigma_k^2 \\
&= |\langle k|\Psi_+\rangle\langle\Psi_+|\psi_0\rangle|^2 + |\langle k|\Psi_-\rangle\langle\Psi_-|\psi_0\rangle|^2 + W_k\sigma_k^2 \\
&\quad + \ \mathrm{Re}\, e^{i2\omega t}\langle k|\Psi_+\rangle\langle\Psi_+|\psi_0\rangle\langle\psi_0|\Psi_-\rangle\langle\Psi_-|k\rangle \\
&= W_k(z_1^2 + z_2^2 + \sigma_k^2) - 2W_k z_1 z_2\cos(2\omega t + 2\phi), \qquad (3.16)
\end{aligned}$$

where z_1, z_2, ϕ_1 and ϕ are real and satisfy

$$z_1\sqrt{W_k}e^{i\phi_1} = \langle k|\Psi_+\rangle\langle\Psi_+|\psi_0\rangle$$

and

$$-z_2\sqrt{W_k}e^{i(\phi_1-2\phi)} \quad = \quad \langle k|\Psi_-\rangle\langle\Psi_-|\psi_0\rangle.$$

We conclude that

$$P(t) = \langle P\rangle - \Delta P\cos(2\omega t + 2\phi),$$

where $\langle P\rangle = W_k(z_1^2 + z_2^2 + \sigma_k^2)$ and $\Delta P = 2W_k z_1 z_2$.

3.5.1 δ-sensitivity of ΔP

With some algebra we obtain that

$$\begin{aligned} z_1 &= \left|\frac{1}{\lambda_- - \lambda_+}\left((\lambda_- - a)\frac{\langle k|\psi_0\rangle}{\sqrt{W_k}} - \frac{\langle l|\psi_0\rangle}{\sqrt{W_l}}b\right)\right| \\ z_2 &= \left|\frac{1}{\lambda_- - \lambda_+}\left((\lambda_+ - a)\frac{\langle k|\psi_0\rangle}{\sqrt{W_k}} - \frac{\langle l|\psi_0\rangle}{\sqrt{W_l}}b\right)\right|. \end{aligned}$$

In order to discuss our result we define the difference between the two rotation angles

$$\delta \triangleq \gamma - \beta.$$

The two extreme cases of $\delta = 0$ and $\delta = O(1)$ were discussed by Biham et al. [7]. Here we discuss the intermediate case where $0 < \delta \ll 1$. One of our results was cited in [7], and we would like to present the full details here. Assuming $W_k \ll 1$ (otherwise the algorithm is not of much use), we can approximate (3.14) by

$$\begin{aligned} \cos\omega &\approx W_k\cos\beta(1 - \frac{\delta^2}{2}) - W_k\frac{\delta}{2} + (1 - W_k)(1 - \frac{\delta^2}{8}) \\ \omega &\approx \sqrt{2W_k\left(1 - \cos\beta + \frac{\delta}{2}\sin\beta\right) + \frac{\delta^2}{4}}. \end{aligned} \tag{3.17}$$

By definition $\omega_\pm = \pi + \beta + \frac{\delta}{2} \pm \omega$, and through approximation of (3.13),

$$\begin{aligned} a &= -e^{i\gamma} + O(W_k) \approx e^{i(\gamma+\pi)}, \\ b &= 1 - e^{i\beta} + O(W_k) \end{aligned}$$

and

$$|b|^2 \approx 2\left(1 - \cos\beta\right).$$

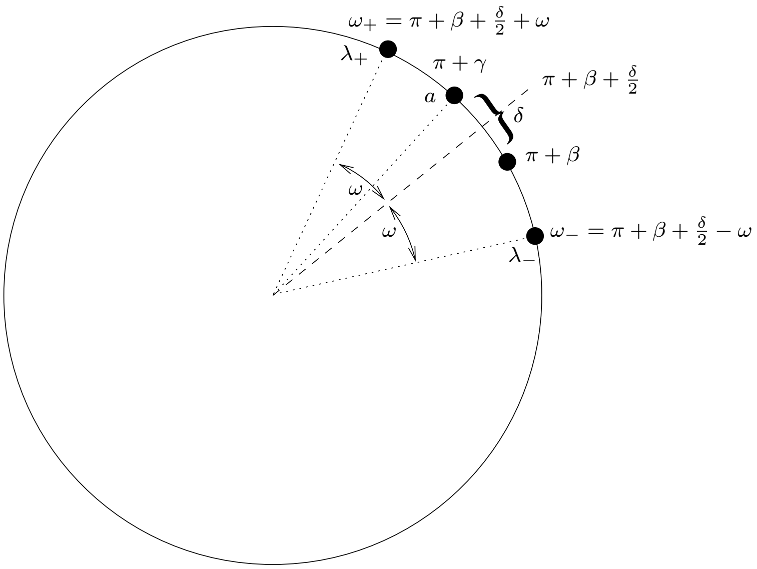
$\omega_+ = \pi + \beta + \frac{\delta}{2} + \omega$
λ_+
$\pi + \gamma$
a
$\pi + \beta + \frac{\delta}{2}$
δ
$\pi + \beta$
ω
ω
$\omega_- = \pi + \beta + \frac{\delta}{2} - \omega$
λ_-

Figure 3.1: The relation between β, γ, $\omega_\pm$, and δ

היחס בין β, γ, $\omega_\pm$ **ו-**δ

The approximations

$$\begin{aligned} \lambda_- - a &\approx -\left(\omega + \frac{\delta}{2}\right) e^{i\beta}, \\ \lambda_+ - a &\approx -\left(\omega - \frac{\delta}{2}\right) e^{i\beta}, \end{aligned}$$

and

$$|\lambda_- - \lambda_+| \approx 2\omega$$

are affirmed looking at Figure 3.1. Finally we can deal with ΔP, using k and l as shorthand for $\frac{\langle k|\psi_0\rangle}{\sqrt{W_k}}$ and $\frac{\langle l|\psi_0\rangle}{\sqrt{W_l}}$, respectively:

$$\begin{aligned} \Delta P &= 2W_k z_1 z_2 \\ &\approx \frac{W_k}{2\omega^2} \left|\left(\omega + \frac{\delta}{2}\right) e^{i\beta} k + bl\right| \left|\left(\omega - \frac{\delta}{2}\right) e^{i\beta} k + bl\right| \\ &= \frac{W_k}{2\omega^2} \left|\left(ke^{i\beta}\omega - bl\right)^2 - \frac{\left(\delta k e^{i\beta}\right)^2}{4}\right| \\ &= \frac{W_k}{2\omega^2} \left|b^2 l^2 - 2blke^{i\beta}\omega + k^2 e^{2i\beta}\left(\omega^2 - \frac{\delta^2}{4}\right)\right| \end{aligned}$$

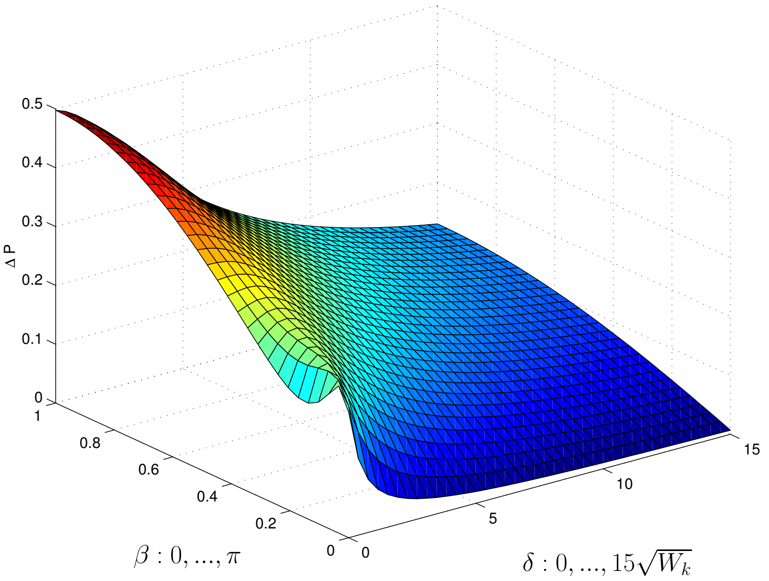

Figure 3.2: ΔP as a function of β and δ, with constant W_k
ΔP כפונקציה של β ו-δ, כאשר W_k קבוע

$$\begin{aligned} &\approx \frac{W_k}{2\omega^2}\left|b^2l^2 - 2blke^{i\beta}\omega + k^2e^{2i\beta}2W_k(1-\cos\beta)\right| \\ &< \frac{W_k}{2\omega^2}\left(|b|^2(1+|k|^2W_k) + 2|b||k|\omega\right) \\ &< \frac{W_k}{2\omega^2}\left(4(1+1) + 4\frac{\omega}{\sqrt{W_k}}\right) \end{aligned}$$

and since by (3.17) $\omega > \frac{\delta}{2}$,

$$< \frac{16W_k}{\delta^2} + \frac{8\sqrt{W_k}}{\delta} \qquad (3.18)$$

which means that if $\delta \gg \sqrt{W_k}$, the algorithm hardly changes the probability. Figure 3.2 illustrates this behavior of ΔP. The special case of (3.18) for $r = 1$, $W_k = 1/N$ and $|\psi_0\rangle = U|0\rangle$, was already discussed in [24].

3.5.2 Finding a Marked State with Certainty

The probability to measure a marked state as a function of time is given by a smooth cosine form (3.16). However, the first experimental implementations of Grover's algorithm use a small number of qubits which means that $\omega = O(2^{-n/2})$ is not that small, and it is probable that the optimal time to measure the register falls far between two iterations. An interesting outcome of the arbitrary rotation generalization, is that one can carefully set β and γ so that the optimal time is exactly an integer. Algebraically, we search for β, γ, $|\psi_0\rangle$ and an integer t such that

$$P(t) = W_k(z_1^2 + z_2^2 + \sigma_k^2) - 2W_k z_1 z_2 \cos(2\omega t + 2\phi) = 1.$$

We do not solve this equation here. Recently, Høyer [22] and Long et al. [25] published two special solutions of this equation.

3.6 The Ultimate Generalization

In this section we discuss the most general iterative quantum process R^t. We show that any such R is a generalized Grover Iterate, where there may be multiple "pivot" states with different β for each of them. Since R may be any unitary operation, this is the ultimate generalization conceivable.

Let f define an arbitrary number of marked states, and let γ be an arbitrary rotation angle. Since *any* unitary operation has a unitary diagonalization [34], there exist U, a set of states S and a corresponding set of angles β such that $RI_f^{-\gamma} = UI_S^{\vec{\beta}}U^\dagger$, where $I_S^{\vec{\beta}}$ rotates the phase of each of the states in S by a possibly different angle. Let us now define

$$Q \triangleq R = UI_S^{\vec{\beta}}U^\dagger I_f^\gamma$$

According to the previous section, we know how R^t operates on arbitrary input if S include a single state. If only we knew to analyze Q when S includes several states and $\vec{\beta}$ includes different elements, we would know how to analyze *any* iterative quantum process.

However, the analysis of such ultra-generalized algorithm has proved to be difficult. To understand why, we consider A and B, two unitary operations over the vector space $\mathbb{C}^N$. Let $\{|a_i\rangle\}_{i=1}^{n_a}$ and $\{|b_j\rangle\}_{j=1}^{n_b}$ be the eigenvectors of A and B, respectively, whose eigenvalues are not 1. We may extend these two sets into a complete basis with $\{|c_m\rangle\}_{m=1}^{N-n_a-n_b}$, unless for some j $|b_j\rangle \in \text{span}(\{|a_i\rangle\}_{i=1}^{n_a})$, in which case we would need more $|c_m\rangle$'s. Now let us examine the operation AB. Since $AB|c_m\rangle = A|c_m\rangle = |c_m\rangle$, it has $N - n_a - n_b$ eigenvectors which are easy to find and whose eigenvalue is 1. Its other $n_a + n_b$ eigenvectors are most likely to have different eigenvalues, and are usually much harder to find. Before we understood this elementary linear

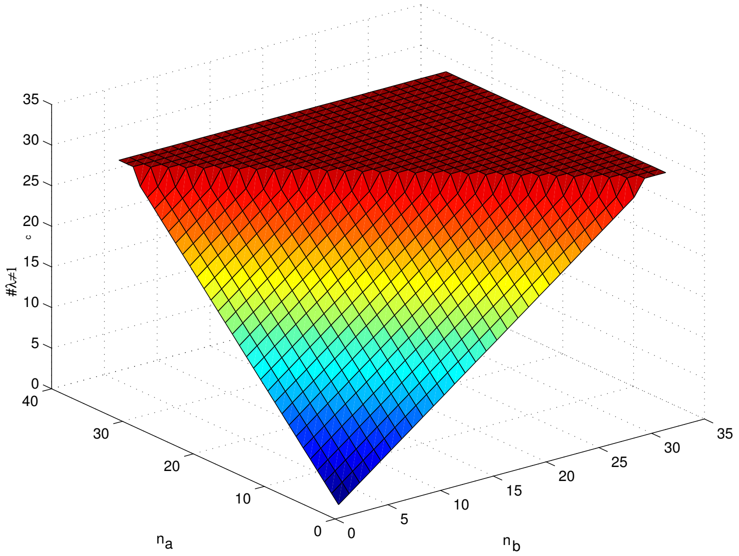

Figure 3.3: The number of eigenvalues which are not 1 as a function of n_a and n_b

מספר הערכים העצמיים שאינם 1 כפונקציה של n_a ו-n_b

algebra fact, we checked it with Matlab and provide a graph in Figure 3.3. In the context of this section, $A = UI_S^{\vec{\beta}}U^\dagger$, $n_a = |S|$, $B = I_f^\gamma$ and $n_b = r$. When $|S| = 1$, the problem degenerates into the simple rotation with partial phase rotations seen in (3.15). However, in general it remains a complex multi-dimensional motion.

Chapter 4

Initialization with a Mixed State

In this chapter we study the case where the original Grover Iterate, as defined in Section 2.2, is applied to a quantum register that is initialized to an arbitrary mixed state. This work is the first rigorous discussion of this case. In addition, our study extends and corrects a result from [8], and provides a simple approximation to the entropy of a pseudo-pure state (4.9). Our generalization can be easily combined with the generalizations from Chapter 3.

4.1 Arbitrary Mixed Initial State

A mixed state arises when one cannot describe the state of a quantum system deterministically, no matter what basis one chooses (Cf. Subsection 1.2.4). Such a state appears very often when a quantum system is entangled with its environment, while the environment cannot be accessed or manipulated.

Extending the argument of Section 3.3, the initial state of the quantum register might not be pure, due to external noise, decoherence or previous manipulations. Instead, the initial state may be some general mixed state $\mathcal{E}$. Given the description of $\mathcal{E}$ as an ensemble, all we can say is that the register is in the pure state $|\psi_i\rangle$ with probability p_i (for all i's).

When the Grover algorithm is applied to the register whose state is $|\psi_i\rangle$, the probability to measure the marked state is $P_i(t)$. The probability for the register to be in that state is p_i. Thus, the total probability to measure the

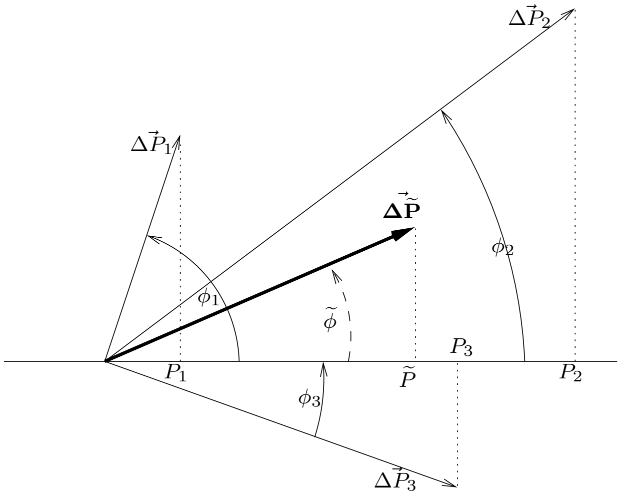
$\Delta\vec{P}_2$
$\Delta\vec{P}_1$
$\vec{\Delta\widetilde{\mathbf{P}}}$
ϕ_2
ϕ_1
$\widetilde{\phi}$
P_3
P_1
$\widetilde{P}$
P_2
ϕ_3
$\Delta\vec{P}_3$

Figure 4.1: The differences $P_i - \langle P_i \rangle$ as projections of a rotating $\Delta\vec{P}_i$
ההפרשים $P_i - \langle P_i \rangle$ כהיטלים של $\Delta\vec{P}_i$ מסתובבים

marked state is the weighted average

$$\begin{aligned}\widetilde{P}(t) &= \sum_i p_i P_i(t) \\ &= \sum_i p_i \left(\langle P_i \rangle - \Delta P_i \cos\left(2\omega t + 2\phi_i\right)\right). \end{aligned} \tag{4.1}$$

The functions

$$P_i(t) - \langle P_i \rangle = -\Delta P_i \cos\left(2\omega t + 2\phi_i\right)$$

share a sinusoidal form, differing in amplitude and phase, but not in frequency. They may be thought of as the projections of vectors rotating in frequency ω, as exemplified in Figure 4.1. Therefore, their weighted sum (the Center of Mass of the vectors in the figure) is a sinusoidal function with the same frequency:

$$\widetilde{P}(t) = \widetilde{\langle P \rangle} - \widetilde{\Delta P} \cos\left(2\omega t + 2\widetilde{\phi}\right) \tag{4.2}$$

where

$$\widetilde{\langle P \rangle} = \sum_i p_i \langle P_i \rangle, \tag{4.3}$$

$$\widetilde{\Delta P} = \sqrt{\left(\sum_i p_i \Delta P_i \cos 2\phi_i\right)^2 + \left(\sum_i p_i \Delta P_i \sin 2\phi_i\right)^2} \tag{4.4}$$

and

$$\tan 2\widetilde{\phi} = \frac{\sum_i p_i \Delta P_i \sin(2\phi_i)}{\sum_i p_i \Delta P_i \cos(2\phi_i)}. \tag{4.5}$$

The probability to measure the marked state reaches its maximum value

$$\widetilde{P}_{max} = \widetilde{\langle P \rangle} + \widetilde{\Delta P} \tag{4.6}$$

after $T = \frac{\pi - 2\widetilde{\phi}}{2\omega}$ iterations.

If the algorithm is repeated until success with T iterations each time, the expected total time to measure a marked state is

$$\mathcal{T}_Q = \frac{\pi - 2\widetilde{\phi}}{2\omega \widetilde{P}_{max}} = \frac{\pi - 2\widetilde{\phi}}{4\widetilde{P}_{max}}\sqrt{N}$$

since the number of repetition until success is distributed geometrically with parameter $\widetilde{P}_{max}$. If this value is significantly smaller than the classical expected time $\mathcal{T}_C = N/2$, then the quantum algorithm has an advantage. Quantitatively, the quantum algorithm is faster by a factor of

$$\frac{\mathcal{T}_C}{\mathcal{T}_Q} = \frac{N\omega \widetilde{P}_{max}}{\pi - 2\widetilde{\phi}} = \frac{2\widetilde{P}_{max}\sqrt{N}}{\pi - 2\widetilde{\phi}}. \tag{4.7}$$

Of course, the constant factors in this expression have no real meaning until we know the relative "clock speed" of quantum computers. Notice that this definition of the quantum advantage counts only oracle queries and does not take into account the cost of repeated initialization of the register. In cases where the initialization is costly, arefind measure should be used.

4.2 Examples

4.2.1 Pure Initial State

When the arbitrary mixed state is chosen to be pure, the summations are degenerate and the results of [6] are regained. For example, if the initial state is the original $\mathcal{E} = \{p = 1, H|0\rangle\}$, the original Grover case is found. If $\mathcal{E} = \{p = 1, |k\rangle\}$ (where $|k\rangle$ is the marked state), then $\widetilde{\langle P\rangle} = \widetilde{\Delta P} = \frac{1}{2}$ and $\widetilde{\phi} = \frac{\pi}{2}$. An interesting known property of the Grover algorithm is that for all states orthogonal to both $|k\rangle$ and $H|0\rangle$, $\widetilde{\langle P\rangle} = \widetilde{\Delta P} = 0$.

4.2.2 Pseudo-Pure Initial State

Ensembles where a pure state $|\psi\rangle$ appears with probability $\epsilon + \frac{1-\epsilon}{N}$ and any state orthogonal to it appears with equal probability of $\frac{1-\epsilon}{N}$ are called pseudo-pure mixed states. They are written more conveniently as

$$\rho_{\epsilon\text{-}pure} = \frac{1-\epsilon}{N} I + \epsilon|\psi\rangle\langle\psi|.$$

Notice that $0 \leq \epsilon \leq 1$ is a measure of the purity of ρ: when $\epsilon = 0$ it is totally mixed, and when $\epsilon = 1$ it is totally pure. It is easy to see that in the limit of large N, $\widetilde{\langle P\rangle} = \epsilon \langle P_\psi\rangle$, $\widetilde{\Delta P} = \epsilon \Delta P_\psi$ and $\widetilde{\phi} = \phi_\psi$. For example, for

$$\rho_{\frac{1}{\log N}\text{-}pure} = \frac{1}{N}\left(1 - \frac{1}{\log N}\right) I + \frac{1}{\log N} H|0\rangle\langle 0|H,$$

we obtain $\widetilde{\langle P\rangle} = \widetilde{\Delta P} = \frac{1}{2\log N}$ and $\widetilde{\phi} = 0$. Notice that although ρ is extremely mixed, the quantum advantage is of factor $\frac{2\sqrt{N}}{\pi \log N} = O(\sqrt{N})$.

4.2.3 Initial State Where m of the Qubits Are Mixed

Let us study the case where the register is initialized to

$$\rho_{m\text{-}mix} = \frac{1}{2^m} \sum_{i=0}^{2^m - 1} H|i\rangle\langle i|H.$$

This state may occur if the m least significant qubits of the register are totally mixed before the first Hadamard transform is applied. Since all $H|i\rangle$ are orthogonal to $H|0\rangle$ (except for $H|0\rangle$ itself) and they are almost orthogonal to $|k\rangle$ (since $|\langle k|H|i\rangle|^2 = \frac{1}{N}$), the evolution of $\rho_{m\text{-}mix}$ is governed by $\{p = 2^{-m}, H|0\rangle\}$ and we obtain $\widetilde{\langle P\rangle} = \widetilde{\Delta P} = \frac{1}{2^{m+1}}$ and $\widetilde{\phi} = 0$. The quantum advantage is of factor $\frac{2\sqrt{N}}{2^m \pi}$: large m would render the algorithm useless.

4.3 Algorithm Usefulness and Entropy

The von Neumann entropy of a mixed state ρ is defined as

$$S(\rho) = -\mathrm{tr}\rho \log \rho. \tag{4.8}$$

Bose et al. [8] presented a new model for quantum computation and laid out a new proof for the optimality of the Grover algorithm (Cf. Section 2.4). However, one of their results was the following: if the Grover algorithm is initiated with a mixed state ρ, such that

$$S(\rho) \geq \frac{1}{2} \log N,$$

the algorithm would have no advantage compared to the classical case. This is in disagreement with our findings: Bose et al. have a mistake regarding the entropy of the *classical* search problem[1]. This mistake does not invalidate the other results in their paper.

A counter-example to their claim is the state $\rho_{\frac{1}{\log N}\text{-}pure}$ as defined above. The entropy of the pseudo-pure state $\rho_{\epsilon\text{-}pure}$ is

$$\begin{aligned}
S(\rho_{\epsilon\text{-}pure}) &= S\left(\frac{1-\epsilon}{N} I_N + \epsilon|0\rangle\langle 0|\right) \\
&= -\sum_{1}^{N-1} \frac{1-\epsilon}{N} \log \frac{1-\epsilon}{N} \\
&\quad -\frac{1+(N-1)\epsilon}{N} \log \frac{1+(N-1)\epsilon}{N} \\
&= -(N-1)\frac{1-\epsilon}{N} \log \frac{1-\epsilon}{N} \\
&\quad -\frac{1+(N-1)\epsilon}{N} \log \frac{1+(N-1)\epsilon}{N},
\end{aligned}$$

and for large N, where $N/(N-1) \approx 1$,

$$\begin{aligned}
&\approx -(1-\epsilon) \log \frac{1-\epsilon}{N} - \left(\frac{1}{N} + \epsilon\right) \log \left(\frac{1}{N} + \epsilon\right) \\
&= (1-\epsilon) \log N - (1-\epsilon) \log(1-\epsilon) - \left(\frac{1}{N} + \epsilon\right) \log \left(\frac{1}{N} + \epsilon\right) \\
&= (1-\epsilon) \log N - \ell, \qquad (4.9)
\end{aligned}$$

---

[1] Just above their Equation (10), they say that classical search can change entropy by $\log \sqrt{N}$ in $\sqrt{N}$ steps. This is true for a search field of size $\sqrt{N}$, but wrong for the question in matter where the search field is of size N.

where $\ell = (1-\epsilon)\log(1-\epsilon) + \left(\frac{1}{N}+\epsilon\right)\log\left(\frac{1}{N}+\epsilon\right) \in [0, 0.88)$ for any $0 \le \epsilon \le 1$ and any $N \ge 2$. For $\epsilon = \frac{1}{\log N}$, we obtain $\ell \approx 0$ and

$$S(\rho_{\frac{1}{\log N}\text{-}pure}) \approx \left(1 - \frac{1}{\log N}\right)\log N = \log N - 1.$$

This entropy is almost maximal. However, as noted above, the Grover algorithm outperforms any classical algorithm, even when it is initialized with this state.

Entropy is not a good measure for the usefulness of the Grover algorithm. For almost every value of entropy, there exist states that are applicable as initializers and states that are not. For example (for $n = \log N$), $S(\rho_{(n-1)\text{-}mix}) = \log N - 1 = S(\rho_{\frac{1}{\log N}\text{-}pure})$, but when initialized in $\rho_{(n-1)\text{-}mix}$, the Grover algorithm is as bad as guessing the marked state. Another example may be given using the pure states $H\left|0\right\rangle\left\langle 0\right|H$ and $H\left|1\right\rangle\left\langle 1\right|H$. With the first, Grover arrives at the marked state with quadratic speed-up, while the second is practically unchanged by the algorithm.

It seems as if Bose et al. had in mind only mixed states of the type of Subsection 4.2.3 ($\rho_{m\text{-}mix}$). For these states, Grover's algorithm is faster when

$$\frac{2\sqrt{N}}{2^m\pi} > 1$$

which means also

$$m < \frac{1}{2}\log N + 1 - \log\pi \approx \frac{1}{2}\log N$$

which in turn is almost identical to [8, Eq. (10)].

Chapter 5

Summary and Conclusions

In this work we have given a brief introduction to Quantum Computation. We presented Grover's quantum search algorithm, and presented a proof that it is the optimal search algorithm—better than any other, be it classical or quantum. Later we surveyed the various generalizations of this algorithm, some of which generalize the operations of the algorithm, and some generalize its initial state. We have analyzed each of the generalization using the method of eigenvector analysis of the Grover Iterate. We stated what is the ultimate generalization thinkable for the operations of the algorithm. Based on the work of Biham et al., we generalized the initial state of the algorithm further, so we can deal with arbitrary mixed initial states.

Particularly, we studied how the algorithm behaves when initialized with a pseudo-pure initial state. We have shown that Grover's algorithm is better than any classical algorithm in some exponentially mixed state. Braunstein et al. [11] showed that mixed-enough states are separable. Had we found a separable state with better-than-classical behavior, it would have had resounding implications: it would have implied that entanglement is not essential for non-trivial quantum computation. However, a careful examination of [11] shows that our results provides some evidence to the contrary.

Figure 5.1 is a cross-section of the Bloch sphere. Braunstein et al. showed that when $\epsilon \geq \frac{1}{1+\sqrt{N}}$, inseparable states are appearing. Notice that according to Subsection 4.2.2, the quantum advantage of Grover's algorithm which is initialized with the pseudo-pure state $\rho_{\epsilon\text{-}pure} = \frac{1-\epsilon}{N} + \epsilon H|0\rangle\langle 0|H$ is of factor $2\epsilon\sqrt{N}/\pi$. From this we learn that a necessary condition for a quantum speed-up is $\epsilon > \frac{\pi}{2\sqrt{N}}$. These two thresholds coincide up to a constant factor. This result may be considered as an evidence (although not a proof) that

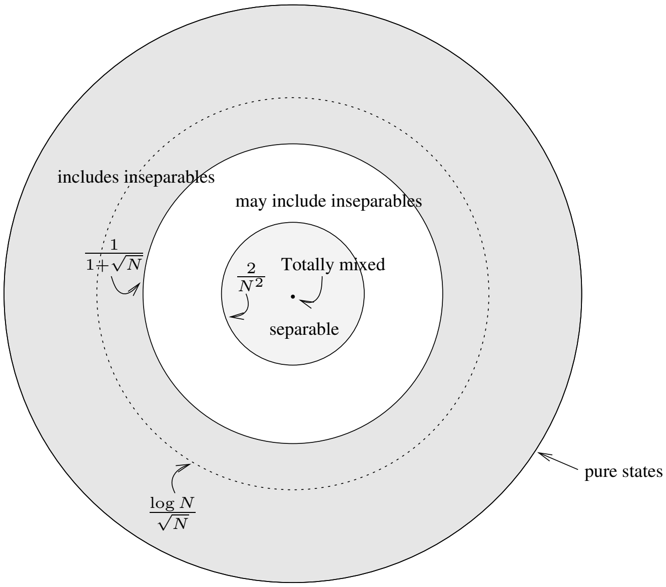
includes inseparables
may include inseparables
$\frac{1}{1+\sqrt{N}}$
$\frac{2}{N^2}$
Totally mixed
separable
pure states
$\frac{\log N}{\sqrt{N}}$

Figure 5.1: Braunstein et al.'s separability bounds on the Bloch ball

חסמי הפריקות של בראונשטיין ושותפיו על כדור בלוך

inseparability *is* necessary for nontrivial quantum computation.

Amazingly or not, we observe that a similar result appears when considering Simon's algorithm (Cf. Section 1.3.2), which performs well only when initialized with the pure state $H|0\rangle$. If it is initialized with the $\rho = \frac{1-\epsilon}{N} + \epsilon H|0\rangle\langle 0|H$ pseudo-pure state and repeated until success, the expected number of queries needed is $O(\frac{\log N}{\epsilon})$. Since the best classical solution requires $O(\sqrt{N})$ queries, the quantum advantage is of factor $\frac{\epsilon\sqrt{N}}{\log N}$. Again, this implies that if $\epsilon < \frac{1}{1+\sqrt{N}}$, the quantum advantage vanishes.

Appendix A

Oracle Equivalence

The regular quantum oracle and the controlled phase quantum oracle of a binary function $f(\cdot)$ were described in Subsection 1.2.5 and illustrated in figures 1.2 and 1.3. In this appendix, we prove that these two oracle models are equivalent. Let us define them rigorously, by describing how each operates on every element of the computation basis. The regular quantum oracle is the operation

$$U_f : |x, y\rangle \to |x, y \oplus f(x)\rangle,$$

and the controlled phase quantum oracle is

$$c\text{-}pU_f : |x, c\rangle \to (-1)^{c \cdot f(x)} |x, c\rangle.$$

Figure A.1 demonstrates the reduction $U_f \geq c\text{-}pU_f$, where H denotes the Hadamard transform of a single qubit. Inspired by [26, page 113], we added the control bit to the known construction of an uncontrolled phase oracle using a regular oracle. Figure A.2 demonstrates the reduction $c\text{-}pU_f \geq U_f$. We apply the Hadamard transform to the control line of $c\text{-}pU_f$ before and after the oracle query. Each transform cancels out its respective Hadamard transform from the previous construction, and returns the oracle to its simple form.

Initially we check the reduction $U_f \geq c\text{-}pU_f$ for the two possible values of c, and for arbitrary value of x:

$$\begin{aligned} |x, 0\rangle &\overset{H}{\to} \frac{1}{\sqrt{2}} \left(|x, 0\rangle + |x, 1\rangle\right) \\ &\overset{U_f}{\to} \frac{1}{\sqrt{2}} \left(|x, f(x)\rangle + |x, 1 \oplus f(x)\rangle\right) \end{aligned}$$

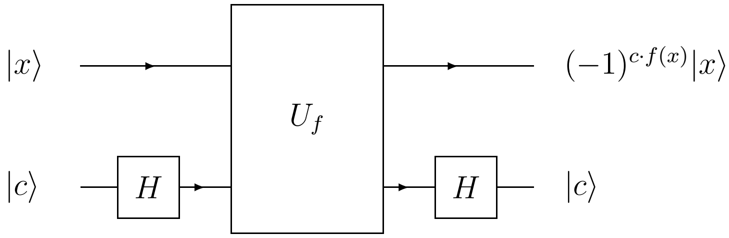
$|x\rangle$
$|c\rangle$
H
U_f
H
$(-1)^{c\cdot f(x)}|x\rangle$
$|c\rangle$

Figure A.1: Construction of a controlled phase quantum oracle using a regular quantum oracle

בניית אורקל פאזה מותנה באמצעות אורקל קוונטי רגיל

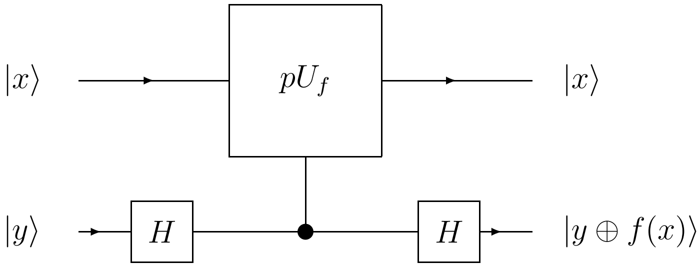
$|x\rangle$
pU_f
$|x\rangle$
$|y\rangle$
H
H
$|y \oplus f(x)\rangle$

Figure A.2: Construction of a regular quantum oracle using a controlled phase quantum oracle

בניית אורקל קוונטי רגיל באמצעות אורקל פאזה קוונטי

$$
\begin{aligned}
&= \frac{1}{\sqrt{2}}\left(|x,0\rangle + |x,1\rangle\right) \\
&\xrightarrow{H} |x,0\rangle = (-1)^{0\cdot f(x)}|x,0\rangle,
\end{aligned}
$$

and

$$
\begin{aligned}
|x,1\rangle &\xrightarrow{H} \frac{1}{\sqrt{2}}\left(|x,0\rangle - |x,1\rangle\right) \\
&\xrightarrow{U_f} \frac{1}{\sqrt{2}}\left(|x,f(x)\rangle - |x,1\oplus f(x)\rangle\right) \\
&= \frac{(-1)^{f(x)}}{\sqrt{2}}\left(|x,0\rangle - |x,1\rangle\right) \\
&\xrightarrow{H} (-1)^{1\cdot f(x)}|x,1\rangle.
\end{aligned}
$$

Similarly, the opposite reduction is checked for the two possible values of y, and for arbitrary $|x\rangle$:

$$
\begin{aligned}
|x,0\rangle &\xrightarrow{H} \frac{1}{\sqrt{2}}\left(|x,0\rangle + |x,1\rangle\right) \\
&\xrightarrow{c\text{-}pU_f} \frac{1}{\sqrt{2}}\big(|x,0)\rangle + (-1)^{f(x)}|x,1\rangle\big) \\
&\xrightarrow{H} |x,0\oplus f(x)\rangle,
\end{aligned}
$$

and

$$
\begin{aligned}
|x,1\rangle &\xrightarrow{H} \frac{1}{\sqrt{2}}\left(|x,0\rangle - |x,1\rangle\right) \\
&\xrightarrow{c\text{-}pU_f} \frac{1}{\sqrt{2}}\big(|x,0\rangle - (-1)^{f(x)}|x,1\rangle\big) \\
&\xrightarrow{H} |x,1\oplus f(x)\rangle.
\end{aligned}
$$

Bibliography

[1] Dorit Aharonov, Alexei Kitaev, and Noam Nisan. Quantum circuits with mixed states. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, pages 20–30, 1998. Also in quant-ph/9806029.

[2] Andris Ambianis. Quantum lower bounds by quantum arguments. In *Proceedings of the 32nd annual ACM Symposium on Theory of computing*, pages 636–643, 21–23 May 2000. Also in quant-ph/0002066.

[3] Adriano Barenco. A universal two-bit gate for quantum computation. *Proc. R. Soc. London A*, June 1995. Also in cond-mat/9505016.

[4] Robert Beals, Harry Buhrman, Richard Cleve, Michele Mosca, and Roland de Wolf. Quantum lower bounds by polynomials. In *Proceedings of the 39th IEEE Conference on Foundations of Computer Science*, pages 352–361, 1998. Also in quant-ph/9802049.

[5] Charles H. Bennett, Ethan Bernstein, Gilles Brassard, and Umesh Vazirani. Strengths and weaknesses of quantum computing. *SIAM Journal on Computing*, 26(5):1510–1523, October 1997. Also in quant-ph/9701001, as manuscript since 1994.

[6] Eli Biham, Ofer Biham, David Biron, Markus Grassl, and Daniel A. Lidar. Generalized Grover search algorithm for arbitrary initial amplitude distribution. *Phys. Rev. A*, 60(4):2742–2745, October 1999. Also in quant-ph/9711070.

[7] Eli Biham, Ofer Biham, David Biron, Markus Grassl, Daniel A. Lidar, and Daniel Shapira. Analysis of generalized Grover's quan-

tum search algorithms using recursion equations. *Phys. Rev. A*, 63:012310, January 2001. Also in quant-ph/0010077.

[8] Sougato Bose, Luke Rallan, and Vlatko Vedral. Communication capacity of quantum information. *Phys. Rev. Lett.*, 85(25):5448–5451, 18 December 2000. Also in quant-ph/0003072.

[9] Michel Boyer, Gilles Brassard, Peter Høyer, and Alain Tapp. Tight bounds on quantum searching. *Fortsch. Phys.*, 46:493–506, 1998. Also in quant-ph/9605034.

[10] Gilles Brassard, Peter Høyer, Michele Mosca, and Alain Tapp. Quantum amplitude amplification and estimation. In quant-ph/0005055, 2000.

[11] Samuel L. Braunstein, Carlton M. Caves, Richard Jozsa, Noah Linden, Sandu Popescu, and Rüdiger Schack. Separability of very noisy mixed states and imlications for NMR quantum computing. *Phys. Rev. Lett.*, 83(5):1054, August 1999. Also in quant-ph/9811018.

[12] David Deutsch and Richard Jozsa. Rapid solution of problems by quantum computation. *Proceedings of the Royal Society of London, Series A*, A439:553–558, 1992.

[13] Whitfield Diffie and Martin E. Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, 22(6), November 1976.

[14] David P. DiVincenzo. Two-bit gates are universal for quantum computation. *Phys. Rev. A*, 51(2):1015–1022, February 1995. Also in cond-mat/9407022.

[15] Edward Fredkin and Tommaso Toffoli. Conservative logic. *International Journal of Theoretical Physics*, 21(3/4):219–253, 1982.

[16] Robert M. Gingrich, Colin P. Williams, and Nicolas J. Cerf. Generalized quantum search with parallelism. *Phys. Rev. A*, 61(5):052313, 2000. Also in quant-ph/9904049.

[17] Lov K. Grover. How fast can a quantum computer search. In quant-ph/9809029.

[18] Lov K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, pages 212–219, New York, 22–24 May 1996. ACM Press.

[19] Lov K. Grover. Quantum mechanics helps in searching for a needle in a haystack. *Phys. Rev. Lett.*, 79(2):325–328, July 1997. Also in quant-ph/9706033.

[20] Lov K. Grover. Quantum computers can search rapidly by using almost any transformation. *Phys. Rev. Lett.*, 80(19):4329–4332, May 1998. Also in quant-ph/9712011.

[21] Lov K. Grover. Quantum search on structured problems. In *QCQS: NASA International Conference on Quantum Computing and Quantum Communications, QCQS*. LNCS, 1998. Also in quant-ph/9802035.

[22] Peter Høyer. Arbitrary phases in quantum amplitude amplification. *Phys. Rev. A*, 62:052304, 2000.

[23] Richard Jozsa. Searching in Grover's algorithm. In quant-ph/9901021, January 1999.

[24] Gui Lu Long, Yan Sung Li, Wei Lin Zhang, and Chang Cun Tu. Dominant gate imperfection in Grover's quantum search algorithm. *Phys. Rev. A*, 61(4):042305, April 2000.

[25] Gui Lu Long, Li Xiao, and Yang Sun. General phase matching condition for quantum searching. In quant-ph/0107013.

[26] Michele Mosca. *Quantum Computers Algorithms.* PhD thesis, Wolfson College, University of Oxford, 1999.

[27] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, 2000.

[28] John Preskill. Lecture notes for Physics 229: Quantum information and computation. http://www.theory.caltech.edu-/people/preskill/ph229/notes/book.ps, September 1998.

[29] Ronald L. Rivest, Adi Shamir, and Leonard M. Adleman. A method for obtaining digital signatures and public key cryptosystems. *CACM*, 21(2), February 1978.

[30] Peter W. Shor. Fault-tolerant quantum computation. In *Proceedings, 37th Annual Symposium on Fundamentals of Computer Science*, pages 56–65, 1995.

[31] Peter W. Shor. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A*, 52(4):R2493–R2496, October 1995.

[32] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997.

[33] Daniel R. Simon. On the power of quantum computation. *SIAM Journal on Computing*, 26(5):1474–1483, 1997.

[34] Vladimir I. Smirnov. *Linear Algebra and Group Theory*, pages 180–184. Dover Publications, Inc., New York, 1970.

[35] Christof Zalka. Grover's quantum searching algorithm is optimal. *Phys. Rev. A*, 60(4):2746–2751, October 1999. Also in quant-ph/9711070.

אלגוריתם החיפוש הקוונטי של גרובר ומצבים מעורבים

דן קניגסברג

אלגוריתם החיפוש הקוונטי של גרובר ומצבים מעורבים

חיבור על מחקר

לשם מילוי חלקי של הדרישות לקבלת תואר
מגיסטר למדעים במדעי המחשב

דן קניגסברג

הוגש לסנט הטכניון – מכון טכנולוגי לישראל
חשון ה'תשס"ב חיפה אוקטובר 2001

המחקר נעשה בהנחיית פרופ' אלי ביהם בפקולטה למדעי המחשב.

אני מודה לאלי ביהם על הנחייתו המסורה לאורך המחקר.

אבקש להודות לטל מור על שיחות פוריות והפניות למאמרים מועילים, לגלעד, זיו, יובל, פיליפ וארי על חברתם בהפסקות הקפה, לצפריר ולנדב על התמיכה המקוונת בענייני יוניקס/לאטך/עברית, לעומר על עזרתו במטלאב, וליוליה על התבנית הראשונה של תיזה זו. עידו ביקש במפורש שלא להזכיר בכתובים את הערותיו האינטיליגנטיות, ולכן לא אודה לו.

תודה מיוחדת שלוחה לאביטל על אהבתה ועידודה, להורי שהביאוני עד הלום, לסבי שלימד אותי קריאה וחשבון, ולסבתי שוויתרה על ביקורים רבים באשמת עבודה זו.

אני מודה לקרן ע"ש מונטי וברתה טייסון ולטכניון על התמיכה הכספית הנדיבה בהשתלמותי.

תוכן עניינים

תקציר **1**

1 מבוא **3**
1.1 חישוב קוונטי - מבט על 3
1.2 מרכיבים בסיסיים של חישוב קוונטי 6
1.2.1 מצבים 6
1.2.2 פעולות 7
1.2.3 מדידה 8
1.2.4 מצבים מעורבים 8
1.2.5 אורקלים קלסיים וקוונטיים 9
1.3 אלגוריתמים קוונטיים מוכרים 10
1.3.1 דויטש-ג'וזה 10
1.3.2 סיימון 11
1.3.3 שור 12
1.4 מבנה התיזה 13

2 אלגוריתם גרובר **14**
2.1 סף תחתון קלסי 14
2.2 האלגוריתם הקוונטי 15
2.3 ניתוח 16
2.3.1 סיבוב סביב הממוצע 17
2.3.2 ביצוע מקוצר 18
2.4 אופטימליות האלגוריתם 18

3 הכללות **24**
3.1 מצבים מסומנים רבים 24
3.2 כמות לא ידועה של מצבים מסומנים 25
3.3 מצב התחלתי שרירותי 26

3.4 הגברת אמפליטודה 28
3.5 סיבובים כלליים 29
3.5.1 רגישות ΔP ל-δ 32
3.5.2 מציאת מצב מסומן בוודאות 34
3.6 ההכללה האולטימטיבית 35

4 אתחול במצב מעורב 37
4.1 מצב התחלתי מעורב שרירותי 37
4.2 דוגמאות 39
4.2.1 מצב התחלתי טהור 39
4.2.2 מצב התחלתי פסידו-טהור 40
4.2.3 מצב התחלתי בו m מהביטים מעורבים לחלוטין 40
4.3 יעילות האלגוריתם ואנטרופיה 41

5 סיכום ומסקנות 43

א שקילות אורקלים 45

תקציר בעברית ה

רשימת איורים

1.1 אורקל קלסי הפיך 9
1.2 אורקל קוונטי רגיל 9
1.3 אורקל פאזה קוונטי מותנה 10

3.1 היחס בין β, γ, $\omega_\pm$ ו-δ 33
3.2 ΔP כפונקציה של β ו-δ, כאשר W_k קבוע 34
3.3 מספר הערכים העצמיים שאינם 1 כפונקציה של n_a ו-n_b 36

4.1 ההפרשים $P_i - \langle P_i \rangle$ כהיטלים של $\Delta\vec{P}_i$ מסתובבים 38

5.1 חסמי הפריקות של בראונשטיין ושותפיו על כדור בלוך 44

א.1 בניית אורקל פאזה מותנה באמצעות אורקל קוונטי רגיל 46
א.2 בניית אורקל קוונטי רגיל באמצעות אורקל פאזה קוונטי 46

תקציר

חישוביות קוונטית היא שדה של תורת החישוביות, אשר מנסה למצוא את מה ניתן לחשב ביעילות, תוך התחשבות באופי הקוונטי של העולם הפיסיקלי. בראשית שנות השמונים של המאה בה התחלתי ללמוד תואר שני, נוכח ריצ'רד פיינמן שסימולציה של מערכות קוונטיות באמצעות מחשב קלסי היא מטלה קשה. נראה היה שאין רדוקציה פולינומית מתהליכים קוונטיים לפעולה חישובית. המצב הקוונטי של מערכת המורכבת מ־n תת־מערכות דו־מצביות שייך למרחב וקטורי מרוכב בן 2^n מימדים, והתפתחותו נקבעת בידי מטריצה אוניטרית שגודלה $2^n \times 2^n$. כל ניסיון לערוך סימולציה של ההתפתחות, הסתכם בעלות אקספוננציאלית מבחינת הזמן הדרוש או מידת הדיוק המתקבלת. כיוון שיש מערכות קוונטיות המדמות זו את זו ביעילות, פיינמן ראה בכך הזדמנות חבויה – ייתכן שתורת הקוונטים מעניקה יכולת שאינה מנוצלת בידי המחשבים הקונבנציונלים.

תורת החישוביות היא תחום של מדעי המחשב, בו מתוכננים מודלים חישוביים וכוחם נבחן. עבור כל מודל מוגדרת קבוצת הבעיות אותה הוא יכול לפתור – $\mathbf{RE}$, $\mathbf{R}$ ו־$\mathbf{NP}^{\mathbf{NP}}$ הן כמה דוגמאות. שאלה חשובה בתורת החישוביות היא מה ניתן לחישוב בפועל, תוך התייחסות למגבלות של זמן ומקום בעולם הממשי. אב־טיפוס למודל חישובי יעיל ובן־ביצוע, הוא מכונת טורינג הפועלת בזמן פולינומי. זמן רב היא נחשבה למודל האולטימטיבי עבור חישוב ממשי. הגרסה החזקה של תיזת צ'רץ'־טורינג מנסחת זאת במפורש באומרה ש"כל פונקציה הניתנת לחישוב יעיל, ניתן לחישוב על־ידי מכונת טורינג בזמן פולינומי". מאוחר יותר הוצע מודל חישובי על־פיו למכונה מותר לקבל גם קלט אקראי ולטעות בהסתברות מסויימת. מודל זה מאפשר, כנראה, לפתור בעיות נוספות ביעילות

(השאלה האם טענה זו נכונה עודנה פתוחה, בדומה לשאלות רבות בתורת החישוביות).

החיפוש אחר מודל חישובי אולטימטיבי הוביל את דייויד דויטש לשאול מה הן המגבלות הפיסיקליות המובנות בתהליך חישובי ממשי. כתוצאה מכך, הוא הגדיר מודל בן־ביצוע לכאורה וחזק יותר לכאורה מהמודלים הקיימים, אשר מנצל תופעות טבעיות מהן מודלים קודמים התעלמו. בשנת 1985 הציג דויטש את המטלה הראשונה אותה פותרת מכונת טורינג קוונטית בפחות צעדים לעומת מכונת טורינג קלסית. יחד עם ריצ'רד ג'וזה הוא הרחיב מטלה זו לסדרת בעיות בהן למחשב קוונטי יש יתרון על־פני מחשב קלסי, אולם יתרון זה הינו תת־אקספוננציאלי והבעיה עצמה אינה קשה או מעניינת במיוחד.

ברנשטיין ווזיראני, ואחריהם סיימון, מצאו בעיות אותן מחשב קוונטי יכול לפתור ביעילות, בעוד מחשב קלסי (אפילו הסתברותי) אינו יכול. אולם ההישג המפורסם ביותר של החישוביות הקוונטית הוא האלגוריתם של שור (Peter Shor), אשר מפרק מספר לגורמיו הראשוניים ומחשב לוגריתם בדיד בזמן פולינומי. הישג נוסף הוא אלגוריתם החיפוש הבלתי־ממוין של גרובר (Lov Grover). אלגוריתם זה מצליח למצוא איבר בודד בתוך בסיס־נתונים בלתי־ממוין, בזמן שהוא בסדר גודל של השורש הריבועי של גודל בסיס־הנתונים.

מכשול משמעותי בפני מימוש של מחשב קוונטי היתה השאלה האם ניתן לבנותו משערים פשוטים השייכים לקבוצה סופית של "אבני בניין". זו שאלה חשובה, הן מבחינה פרקטית (לא ניתן לחשוב על ייצור תעשייתי ללא אלמנט בסיסי פשוט) הן מבחינה תיאורטית (כי אם סימולציה של שער קוונטי כללי באמצעות מרכיבים פשוטים איננה יעילה, לא ניתן לראות אותו כמשאב זול.) היה זה שוב דויטש שסיפק שער אוניברסלי בן 3 כניסות ויציאות. מאוחר יותר הראה דוינצ'נזו שדי להשתמש בשער מסויים בן 2 כניסות בצירוף סיבוב פאזה של ביט בודד, על־מנת לדמות כל חישוב בדיוק סביר.

מכשול חמור נוסף בפני חישוב קוונטי מציאותי, היא שבירותן המובנית של מערכות בסדר־הגודל הקוונטי. חלקיקים כגון פוטונים, אלקטרונים, גרעינים, אטומים ופרודות, הינם רגישים להפרעות חיצוניות כגון שינויי טמפרטורה, שדה אלקטרו־מגנטי או רעידות. הם

נוטים לפלוט אנרגיה באופן ספונטני ולשנות את מצבם באורח אקראי. נראה היה שלא ניתן לשמור מידע קוונטי ולטפל בו באופן מבוקר. בחישוב קלסי מתמודדים עם קשיים דומים באמצעות מדידה ושמירה רציפה של המידע, תוך רענון ערכים בעייתים. אמצעי משלים הוא שימוש בקודים לתיקון שגיאות. האמצעי הראשון לחלוטין אינו מתאים למידע קוונטי, כיוון שפעולת המדידה עושה בדיוק את אשר אנו רוצים להימנע מפניו: המצב המקורי קורס ונהרס. לא ברור היה אם ניתן להרחיב את האמצעי השני עבור מידע קוונטי, עד אשר ב־1995 הדגים פיטר שור את קיומו של קוד תיקון שגיאות קוונטי ראשון. מאוחר יותר הוא הסביר כיצד לבצע את תהליך תיקון השגיאות באופן חסין לשגיאות (fault-tolerant).

בינתיים התפתח המחקר של מידע קוונטי כמושג עצמאי. הוגדר המרכיב הבסיסי שלו, הביט הקוונטי (או קיוביט), ונתגלו קשרים מעניינים בין מידע קוונטי וקלסי. למשל, קל לשכפל מידע קלסי, בעוד משפט אי־השכפול קובע שלא ניתן לעשות זאת עבור מידע קוונטי שרירותי. מסקנה יסודית אחרת הוא חסם חולבו (Holevo): ביט של מידע קוונטי מוגדר ע"י שני מספרים ממשיים בלתי־תלויים, בעוד ביט קלסי מודר ע"י ספרה בינארית בודדת. אולם לא ניתן לקודד בנאמנות יותר מביט קלסי אחד באמצעות ביט קוונטי.

לפיכך היתה זו הפתעה, כאשר ב־1992 צ'רלס בנט וסטיבן ויזנר מצאו דרך להעביר 2 ביטים המקודדים לתוך קיוביט אחד. תוצאה זו הקרויה קידוד על־צפוף (superdense coding) אינה סותרת את חולבו, מכיוון שבנוסף לקיוביט המועבר, היא מנצלת משאב נוסף: זוג קיוביטים סבוכים (entangled) ששני המתקשרים חלקו מראש. לחלופין, ניתן להשתמש בזוג סבוך כזה ובשני ביטים קלסיים על־מנת להעביר קיוביט שרירותי, בתהליך הקרוי טלפורטציה קוונטית. באורח מדהים למדי, הטלפורטציה מתבצעת ללא מעבר של חומר – שום דבר פרט למידע אינו מוחלף בין המתקשרים. השאלה כיצד שלושת סוגי המשאבים הללו מתייחסים זה לזה עודנה פתוחה (כמו גם הצורך לכמת סביכות).

העניין הפרקטי הרב שעורר האלגוריתם של שור נובע מיכולתו לתקוף את צפני המפתח הציבורי העיקריים שבשימוש. תרופה לחולשה זו היא הפרוטוקול הקוונטי להחלפת מפתח סודי של צ'רלס בנט וז'יל ברסרד, המכונה BB84. הוא מאפשר לשני משתמשים מרוחקים

לשלוח זה לזה הודעה פרטית בסודיות. בניגוד לפרוטוקולים קלסיים מקבילים, בטיחותו של BB84 איננה מותנית בכוחו החישובי של התוקף אלא רק בכך שהוא קיים בעולם הקוונטי שלנו.

אנו פותחים תיזה זו במבוא קצר לחישוביות קוונטית. אנו מציגים את המרכיבים הבסיסיים של שדה מדעי זה – מהו מצב של מערכת קוונטית, כיצד מצב זה יכול להשתנות, כיצד ניתן למדוד אותו, וכן מהם מצבים קוונטים מעורבים (mixed states). על־פי המודל המקובל, מערכת קוונטית המשמשת לחישוב מורכבת ממספר סיביות קוונטיות (קיוביטים, qubits). קבוצה של כמה סיביות כאלה נקראת רגיסטר קוונטי. בכל רגע נתון, מצב טהור (pure state) של רגיסטר בן n קיוביטים, הוא וקטור יחידה במרחב הילברט בן $N = 2^n$ מימדים. במרחב זה מוגדר בסיס פורש הנקרא "בסיס החישוב" (computation basis), ואיבריו מסומנים ב־$|0\rangle$ עד $|N-1\rangle$. מצב של הרגיסטר הוא קומבינציה לינארית של איברי הבסיס (*סופרפוזיציה* שלהם), כאשר לכל איבר מתאים מקדם מרוכב המוגדר ע"י אמפליטודה ופאזה. ההתפתחות בזמן של מצב של מערכת קוונטית סגורה היא אוניטרית (unitary), כלומר היא לינארית, הפיכה, ומשמרת־הסתברות.

ניתן למדוד את הרגיסטר בבסיס החישוב על־מנת ללמוד על מצבו. תוצאת המדידה תהיה אחד מאברי הבסיס, וההסתברות למדוד איבר בסיס מסויים היא ריבוע האמפליטודה המתאימה לו. עם זאת, עבור כל מצב טהור ניתן לתכנן ניסוי (פעולה אוניטרית ואחריה מדידה) אשר יאשר בוודאות האם הרגיסטר נמצא בו. בניגוד לכך, לא ניתן לעשות זאת עבור רגיסטר הנמצא ב*מצב מעורב* (mixed state) – תוצאתו של כל ניסוי תהיה אקראית במידה מסויימת. רגיסטר יכול להגיע למצב מעורב אם הוא חלק ממערכת גדולה יותר שהיתה במצב טהור, אך שאר המערכת הפכה לבלתי־נגישה (למשל, התפזרה לכל עבר). לקראת סוף פרק 1, אנו מסבירים מהו אורקל קוונטי רגיל של פונקציה ומהו אורקל פאזה קוונטי שלה, ומוכיחים (בנספח א') שהם שקולים זה לזה.

בפרק 2 אנו מציגים את אלגוריתם החיפוש של גרובר, ואת אחת ההוכחות שהוא אלגוריתם החיפוש האופטימלי עד כדי מקדם כפלי קבוע (אלגוריתם גרובר דורש פחות

קריאות אורקל לעומת כל אלגוריתם אחר, קוונטי או קלסי). האלגוריתם המקורי של גרובר מניח שמתוך N מספרים, רק עבור ערך (בלתי־ידוע) אחד, k, מתקיים $f(k) = 1$, ואילו עבור כל $x \neq k$ מתקיים $f(x) = 0$. k נקרא הערך המסומן ע"י הפונקציה f. האלגוריתם מניח כי נתון לנו אורקל קוונטי של הפונקציה בלבד, ולא ידוע לנו שום מידע נוסף עליה. האלגוריתם פועל על רגיסטר קוונטי בן $\log N$ ביטים, ומתחיל באתחולו של הרגיסטר למצב $H|0\rangle$ (סופרפוזיציה אחידה של כל המצבים מ־$|0\rangle$ עד $|N-1\rangle$). לאחר מכן, האלגוריתם מבצע $O(\sqrt{N})$ איטרציות. כל איטרציה כוללת קריאה לאורקל, טרנספורם הדמרד, היפוכי פאזה למצב $|0\rangle$ וטרנספורם הדמרד נוסף. (טרנספורם הדמרד, H, הוא פעולה קוונטית שימושית ביותר המתוארת בקצרה בגוף העבודה.) לבסוף, יש לבצע מדידה של הרגיסטר הקוונטי, ובוודאות כמעט מוחלטת תוצאת המדידה תהיה k. לפיכך, אלגוריתם גרובר מאפשר לזהות את k לאחר $O(\sqrt{N})$ קריאות אורקל, בעוד שבאמצעים קלסיים לא ניתן לעשות זאת בפחות מ־$O(N)$ קריאות.

הפרמטרים השונים של האלגוריתם הוכללו בידי מחברים רבים. חלק מההכללות מכלילות את האיטרטור, וחלקן – את המצב ההתחלתי של האלגוריתם. האלגוריתם הורחב כך שיטפל גם בפונקציות המגדירות כמה ערכים מסומנים, ונבדק כיצד הוא מתנהג כאשר הוא מאותחל במצב שרירותי. הוכח שניתן להשתמש בו להאצת היוריסטיקות חיפוש, ונחקרה התנהגותו כאשר היפוכי הפאזה מוחלפים בסיבובי פאזה בזוית שרירותית. אנו סוקרים הכללות אלה בפרק 3, ומנתחים אותן בשיטה אחידה. עבור כל אחת ואחת מהן אנו מגדירים בסיס נוח, מחשבים כיצד משתנה המצב של הרגיסטר הקוונטי במהלך האלגוריתם, ומגיעים לביטוי עבור ההסתברות למדוד ערך מסומן בכל איטרציה. מצאנו מהי ההכללה האולטימטיבית של האיטרטור של גרובר, ואנו דנים בכך בסוף הפרק.

בפרק 4 אנו מציגים הכללה חדשה של המצב ההתחלתי של האלגוריתם, בה הוא רשאי להיות מעורב באופן שרירותי. שוב, אנו נותנים ביטוי עבור ההסתברות למדוד מצב מסויים כפונקציה של מספר האיטרציות. גם כאשר המצב ההתחלתי מעורב, לפונקציה זו יש צורת קוסינוס, וכאשר המצב מעורב מאוד, האפליטודה של הקוסינוס די חלשה. אולם

אנו מראים כי אפילו כאשר המצב ההתחלתי מעורב במיוחד, יש מקרים בהם האלגוריתם מתפקד היטב – טוב יותר מכל אלגוריתם קלסי. דוגמה לכך היא המצב הפסידו־טהור (pseudo-pure state) $\rho_{\epsilon\text{-}pure} = \frac{1-\epsilon}{N} + \epsilon H|0\rangle\langle 0|H$ כאשר $\epsilon = \frac{1}{\log N}$. כאשר האלגוריתם מאותחל במצב זה, וחוזרים עליו שוב ושוב עד להצלחה, המצב המסומן יימצא לאחר $O(\frac{\sqrt{N}}{\log N})$ קריאות אורקל בממוצע.

אנו מספקים קירוב לאנטרופיית פון נוימן (von Neumann entropy) של מצב פסידו־טהור, ומצאנו שהיא גדלה באופן רציף עם מידת העירוב של המצב הפסידו־טהור. עד כדי ביט אחד, $S(\rho_{\epsilon\text{-}pure}) \approx (1-\epsilon)\log N$. קירוב זה, בצירוף תוצאתינו הקודמת לגבי תפקודו המוצלח של האלגוריתם, אינו תואם תוצאה של בוזה ושותפיו (Bose, Rallan ו־Vedral). המצב $\rho_{\frac{1}{\log N}\text{-}pure}$ הוא דוגמה נגדית פשוטה לטענתם שכאשר האלגוריתם מאותחל במצב שבו האנטרופיה גדולה מ־$\frac{1}{2}\log N$, האלגוריתם אינו טוב יותר מחיפוש קלסי. כמו כן, אנו מסבירים היכן חלה טעותם.

אנו בוחנים את השימושיות של אלגוריתם גרובר כאשר הוא מאותחל במצב פסידו־טהור, ומספקים מדד עבור יעילותו. כאשר האלגוריתם מאותחל ב־$\rho_{\epsilon\text{-}pure}$ הוא מהיר פי $O(\epsilon\sqrt{N})$ מכל אלגוריתם קלסי. על־פי מדד זה, קיים סף שמתחתיו האלגוריתם הקוונטי הופך למיותר – כאשר $\epsilon < o\left(\frac{1}{\sqrt{N}}\right)$. גילינו שסף זה חופף לחסם אי־הפריקות (inseparability bound) של בראונשטיין ושותפיו (Braunstein, Carlton, Caves, Jozsa, Linden, Popescu ו־Schack). על־פי החסם שלהם, כאשר $\epsilon > \frac{1}{1+\sqrt{N}}$, קיימים מצבים פסידו־טהורים שלא ניתן לפרקם למכפלה טנזורית של מצבי קיוביטים בודדים. מצבו של רגיסטר קוונטי כזה נקרא בלתי־פריק, או סבוך (inseparable, entangled). ניתן לראות תוצאה זו כעדות לכך שסביכות (אי־פריקות, entanglement) הינה הכרחית עבור חישוב קוונטי בלתי־טריוויאלי.