Reference Frames and Algorithms for Quantum Information Processing

by

Lana Sheridan

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Physics

Waterloo, Ontario, Canada, 2009

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The main results of this thesis fall in to two areas, firstly quantum reference frames as a resource for quantum computations and secondly quantum algorithms. The results relating to quantum references consider their scaling with a requirements to perform measurements, operations and computations with a certain fidelity. For the case of a directional frame, the central question considered is of how many operations or measurements can be performed with it before its fidelity falls below some threshold. This is found to scale as the square of the size of the reference frame under for a range of physically interesting cases. To prove that result a new general form for any rotationally invariant map. This could have many applications is comparing and classifying rotationally invariant behaviour of quantum systems. Phase references are also considered for the case of performing quantum computations under an energy conservation law. The restriction that the expected energy be conserved for large quantum computations is shown to be manageable in many different potential architectures. In the case of completing computations is an energy conserving subspace, the requirements for ancillas are sublinear in the number of qubits, and even in a circuit model implementation, the errors due to phase reference imperfections are shown to not limit the apparent algorithmic improvements of quantum computing over classical computing.

A quantum walk for the novel concept of two entangled walkers is proposed and analyzed. A modest improvement is found in the scaling of the expected separation of the walkers over the separable case. It illustrates the potential for making use of particle statistic behaviour in algorithms. Lastly, the relation between discrete and continuous time models of quantum computing is explored through the analysis of a new algorithm for simulating the Hamiltonian behaviour of a black box unitary operation. The scaling of the number of calls to the unitary required to obtain a simulation correct to within a parameter ϵ is found, as is a case where the efficiency of the algorithm is superior to directly applying the unitary repeatedly. Applications of the algorithm are considered.

Acknowledgements

I would like to thank my supervisor Michele Mosca for his guidance and wisdom throughout our years of working together. I also thank Mike and Ophelia Lazaridis, who gave me their support through the Mike and Ophelia Lazaridis Fellowship.

I thank my committee, Richard Cleve, Achim Kempf, Norbert Lütkenhaus, Terry Ruldolf, my external examiner, and Raymond Laflamme, my co-supervisor, for their constructive comments and questions.

For Chapter 3, I wish to acknowledge the contributions of my co-authors Jean-Christian Boileau, Martin Laforest, and Stephen Bartlett, in particular to J.-C. for inspiring the project and suggesting the form of the covariant map. Also, Terry Rudolph for an insight in proving the first theorem. I also thank Giacomo Mauro D'Ariano, Joseph Emerson, David Kribs, David Poulin and Peter Turner for helpful discussions.

For Chaper 4, I acknowledge the collaboration of Michele Mosca and helpful discussions with Raymond Laflamme and Norbert Lütkenhaus.

For Chapter 5, I acknowledge my co-authors Yasser Omar, Nikola Paunković, and Sugato Bose. I would like to thank Konrad Banaszek for his guidance, and Elham Kashefi and Tobias Schaetz for useful discussions.

For Chapter 6, I acknowledge collaboration with Dmitri Maslov and Michele Mosca. I thank Scott Aaronson, Patrick Hayden, Rolando Somma, and John Watrous for discussions about the relevance of this problem. I also thank Donny Cheung for helpful insights.

I would also like to acknowledge the pivotal role that the work of Trent Renzor played in the accomplishment of this thesis. I did not fall apart.

Most importantly, I would like to thank my parents: my mother who inspired me and my father who encouraged me, for as long as I can remember; and Bill, who knows that it was his love and support that pulled me through.

Table of Contents

List of Tables

List of Figures

Chapter 1

Overview

1.1 Introduction

This thesis spans two main topic areas, quantum reference frames and quantum algorithms. The work on reference frames stems from the question, what if the physical system which defines the measurement convention for a quantum system is itself quantized? This has implications not only for understanding the role of conservation laws in quantum theory, but also for large-scale quantum computation in which the scaling of the resources required for the computation are critical to whether it is a practical enterprise. The reference system in this case becomes a required resource necessary for maintaining coherence within the computer or for performing accurate measurements on parts of the quantum system. The choice to quantize the reference system is equivalent to moving the "Heisenberg cut" outward to include another system in the quantum mechanical description. This system might be used in measurement or defining operations that are normally chosen to be classical. The intuitive attraction of doing this is that it gives a natural interpretation of conservation laws in quantum systems: that conservation laws are rigid and apply individually to every term in a superposition. Further, this allows for reference information to be treated as a resource and a theory developed to describe reference requirements in different scenarios.

The quantum algorithms results relate to two different problems. The first is quantum walks. Random walks have long been of interest in a classical setting for solving a wide range of problems including modelling share prices, genetic drift, and Brownian motion. It is also used to sample at random from a set in Monte Carlo techniques [GG84, RC04] and to estimate the size of a set [BYG06] and in search

algorithms [LCC+02]. Quantum walks are the extension of this concept to quantum states of the walker and quantum operations for the evolution of the walk. The second problem is of simulating continuous time quantum evolution using access only to a discrete black box unitary operation. An algorithm is developed for achieving this, with some constraints. This is of theoretical interest for understanding the relation between the continuous and discrete models of quantum computing.

1.2 Summary of Work Presented

In Chapter 2 relevant background information is reviewed as a means of introducing the notation that will be used in the rest of this thesis. This falls into six sections, quantum mechanics and information, quantum optics, quantum angular momentum theory, an overview of relevant group theoretic concepts, applications of them to reference frames and angular momentum symmetries, and a consideration of reference frames as quantum systems and their relation to conservation laws and superselection rules.

In Chaper 3, the problem of a finite size reference frame is considered. The reference studied is a three-space direction reference. The reference is needed for some quantum computation task — either to define a measurement that conserves angular momentum, or to define an operation on a quantum system. It is treated as a limited resource which is consumed in the course of performing these tasks. Specifically, the effect of completing these tasks on a series of identical completely mixed quantum states drawn from a reservoir is considered. It is an extension and generalization of work by Bartlett *et al.* [BRST06]. A new general form is found for expressing all rotationally covariant maps on density matrices in terms of the angular momentum operators. This is used to derive relations that describe the change of the expectation value of the angular momentum along the z-direction for the quantum reference frame. There relations allow for the derivation of a set of physically significant conditions under which the fidelity of a task performed by a reference frame will decay with the square of the total angular momentum of the reference, j^2. Additionally, the definition of fidelity is generalized to show that this behaviour is not an artifact of a choice of definitions. Some examples are given to demonstrate the power of the proposed methods. Lastly, there is some consideration of the limitations and potentials of the results.

In Chapter 4, a survey is undertaken of results relating to the analysis of quantum reference frames and energy requirements. The aim is to explore what reference

frame resources are required for quantum information and in what contexts. New conclusions are drawn about how reference frames might be managed in a large quantum computer. Various methods for implementing gates while maintaining consistent definitions of relative phase are considered and compared for practicality and efficiency. The problem of maintaining phase references, and hence coherent qubits, is shown to be manageable and not a severe limitation to implementing quantum computing.

Chapter 5 is a consideration of a quantum walk algorithm. The case of a discrete time walk on a line is considered for the novel proposal of a pair of entangled walkers. This slight modification to originally proposed quantum walks gives a rise to a rich variety of interference effects that can be tuned based on the initial state of the two walkers. In particular, in the case of the antisymmetric state the expected coverage of the line is greater than for a pair of walkers in a product state. Potential implementations of such a walk and its applications are considered.

In Chapter 6, the question of how some fraction of an unknown unitary gate can be implemented on a quantum system is posed. Specifically, the goal is to implement a unitary raised to any real number exponent. A new algorithm is given for achieving this operation with a probability selected by the user and a number of calls to the unitary that depends on the chosen probability of success. An example is given where the exponent of the unitary to be applied is a positive integer, t, but the new algorithm runs with less than t calls to the unitary. The algorithm is considered in the context of different possible applications and the constraints that they may impose. Also, the implications of this work for the relation of discrete to continuous time models of quantum computing are discussed.

The final chapter summarizes the conclusions to be drawn from each chapter and and appendix containing a proof of Theorem 3.1 from Chapter 3 and Lemma B from Chapter 6.

Chapter 2

Background

2.1 Overview

In this chapter, notational conventions are introduced that are used in the remainder of this thesis through a brief review of relevant background material. First, quantum mechanics and quantum information are outlined. The concept of states, density operators, measurement operators, channels, qubits, and Pauli operations are introduced. Second, some relevant quantum optics theory is presented beginning with the quantum harmonic oscillator. Various useful relations involving creation and annihilation operators and coherent states are derived. This formalism will be useful in conjunction with Chapter 4. Third, quantum angular momentum and the notation commonly used by physicists to describe angular momentum states is given. This is relevant to Chapter 3. A brief introduction to relevant group theoretic concepts is provided. Lastly, some ideas regarding reference frames and their behaviour in a quantum setting are introduced.

2.2 Quantum Mechanics and Quantum Information

Quantum mechanics is a generalization of classical mechanics and can be arrived at from a suitable correction to classical Hamiltonian Mechanics [Dir58, Kem04]. The *Hamiltonian*, H, of a system is the energy operator that specifies how energy is distributed in a system. H is given by the sum of the kinetic, T, and potential,

V, energy contributions:[1]

$$H = T + V. \tag{2.1}$$

The equations of motion of a system governed by a Hamiltonian, H can be determined from this quantity alone using the *Poisson bracket* (the Lie bracket for a Poisson algebra).

The Poisson bracket $\{\cdot, \cdot\}$ for the positions and momenta is defined by the following:

$$\{x_i, p_j\} = \delta_{i,j}, \qquad \{x_i, x_j\} = \{p_i, p_j\} = 0$$

where the x_i and p_i are the position and momentum coordinates of a single particle in the system, with the index $i \in \{1, 2, ...d\}$, where d is the number of dimensions in the system (commonly, $d = 3$), and $\delta_{i,j}$ is the Kronecker delta. For the case of many particles, the subscript can also index the individual particles.

In general, for functions of these coordinates, $f(x_i, p_i, t)$ and $g(x_i, p_i, t)$ the Poisson bracket is given by

$$\{f, g\} = \sum_i^d \left(\frac{\mathrm{d}f}{\mathrm{d}x_i}\frac{\mathrm{d}g}{\mathrm{d}p_i} - \frac{\mathrm{d}f}{\mathrm{d}p_i}\frac{\mathrm{d}g}{\mathrm{d}x_i} \right). \tag{2.2}$$

For any function $f(x, p, t)$ which is polynomial in x_i and p_i the equations of motion have the form:

$$\frac{\mathrm{d}}{\mathrm{d}t} f(x_i, p_i, t) = \{f(x_i, p_i, t), H\} + \frac{\partial}{\partial t} f(x_i, p_i, t) \tag{2.3}$$

The quantum mechanical equations of motion are obtained by a deformation of the Poisson bracket to a modified Lie bracket [Gro46]:

$$\frac{\mathrm{d}f}{\mathrm{d}t} = -\frac{i}{\hbar}[f, H] + \frac{\partial f}{\partial t}, \tag{2.4}$$

where for matrix group representations of the operators f, H, *etc.*, the Lie bracket $[f, H]$ is simply the commutator $[f, H] = fH - Hf$. Now that f and H (and x and p) can be matrices, this commutator is not zero in general. Planck's constant, $h = 2\pi\hbar = 6.26 \times 10^{-34}$ Joules per second. In this thesis, subsequent chapters will adopt the units convention such that $\hbar = 1$.

This non-commutation has the consequence that there is a limited precision to

[1] This is usually true. It fails if the transformations defining the coordinate system depend on time, t.

which one can know the simultaneous values of pairs of variables. This is called the *Uncertainty Principle*. It can be expressed:

$$\Delta f \Delta g \geq \frac{1}{2} \left| \langle [f, g] \rangle \right|, \tag{2.5}$$

where Δf is the standard deviation of f, $\Delta f = \sqrt{\langle f^2 \rangle - \langle f \rangle^2}$, $\langle f \rangle$ is the expectation value of f, and likewise for Δg. Conjugate pairs of variables are pairs that have the largest magnitude value of the commutator, $i\hbar$, and hence the largest joint uncertainty. The variables x and p are one example, and another pair of such variables is energy, E, and time, t.

States in quantum mechanical systems can be expressed as vectors in a Hilbert space. A *Hilbert space* is an inner product space that is complete, meaning that every Cauchy sequence converges, with respect to the norm induced by the inner product, to a limit within the space. The evolution of quantum states in closed systems is described by unitary matrices. Unitary matrices obey the equation $U^\dagger U = \mathbb{I}$. In this thesis, the notation of physicists is employed, so $M^\dagger$ denotes the Hermitian conjugate of M. Such matrices have the property that they preserve the inner product between two vectors. This is easy to see. Two vectors $|\psi\rangle$ and $|\phi\rangle$ have the inner product $\langle\phi|\psi\rangle$. If the unitary operation U is applied to both vectors then they are mapped to new vectors give by $U|\psi\rangle$ and $U|\phi\rangle$ and the inner product becomes $\langle\phi| U^\dagger U |\psi\rangle = \langle\phi|\psi\rangle$.

Unitary evolution is not sufficient to describe what happens during the process of measurement. The state vector can be interpreted as a probability distribution, in the sense that the state vector is not observed directly, rather a measurement operation is performed on it and there is a single outcome for each event. Over many measurements of identically prepared systems, the state vector can be reconstructed.

Observables are Hermitian operators ($\hat{\Lambda} = \hat{\Lambda}^\dagger$) on the Hilbert space and represent possible measurements on the system. These measurement operators are applied to the state vector of the system, $|\psi\rangle$. The operators take the form $\hat{\Lambda} = \sum_i \lambda_i \Pi_i$ where Π_i are projectors onto subspaces of the system's vector space such that $\sum_i \Pi_i = 1$ and $\Pi_i \Pi_j = \delta_{i,j} \Pi_i$. The final state is renormalized so that after the measurement the state becomes

$$|\psi\rangle \xrightarrow[\text{outcome } \lambda_i]{\text{measurement}} \frac{\Pi_i |\psi\rangle}{\sqrt{\langle\psi| \Pi_i |\psi\rangle}} \tag{2.6}$$

if the outcome seen is λ_i. The probability for the outcome λ_i to be seen is given by

$P(\lambda_i) = \langle\psi| \Pi_i |\psi\rangle$. Operators have eigenvalues, λ_i, that are the possible outcomes of the measurement, $\hat{\Lambda}$, and the eigenvalues have corresponding eigenvectors, $|\lambda_i\rangle$:

$$\hat{\Lambda} |\lambda_i\rangle = \lambda_i |\lambda_i\rangle , \tag{2.7}$$

which are the final states of the system after the measurement. Such operators' eigenvalues are real numbers:

$$\begin{aligned} \langle\lambda_i| \left(\hat{\Lambda} |\lambda_i\rangle\right) &= \left(\langle\lambda_i| \hat{\Lambda}\right) |\lambda_i\rangle \\ \langle\lambda_i|\lambda_i\rangle\lambda_i &= \lambda_i^* \langle\lambda_i|\lambda_i\rangle \\ \lambda_i &= \lambda_i^*. \end{aligned} \tag{2.8}$$

Also, (using equation (2.8)) the eigenvectors have the property:

$$\begin{aligned} \langle\lambda_j| \left(\hat{\Lambda} |\lambda_i\rangle\right) &= \left(\langle\lambda_j| \hat{\Lambda}\right) |\lambda_i\rangle \\ \langle\lambda_j|\lambda_i\rangle\lambda_i &= \lambda_j \langle\lambda_j|\lambda_i\rangle \\ &\Rightarrow (\lambda_i - \lambda_j)\langle\lambda_j|\lambda_i\rangle = 0. \end{aligned} \tag{2.9}$$

This implies that for *non-degenerate* eigenvalues ($\lambda_i \neq \lambda_j$ for $i \neq j$), $\langle\lambda_j|\lambda_i\rangle = 0$ for all $i \neq j$, meaning that all eigenvectors corresponding to non-degenerate eigenvalues of Hermitian operators are orthogonal. If the system is in an eigenstate of a particular measurement operator it will have no projection into any other eigenstate.

In the case of infinite dimensional Hilbert spaces, the state $|\psi\rangle$ can also be expressed as a function of a continuous variable such as position x is $\langle x|\psi\rangle = \psi(x)$. In this context, $\psi(x)$ is often called the *wavefunction* of the system. Transforms between continuous bases are accomplished by integrating. Note that:

$$\int_{x\in\mathbb{R}} |x\rangle \langle x| \, \mathrm{d}x = \mathbb{I}. \tag{2.10}$$

This simply implies that the basis change is unitary. So, to change from the position, x, basis to the momentum, p, basis:

$$\psi(p) = \langle p|\psi\rangle = \int \langle p|x\rangle\langle x|\psi\rangle \mathrm{d}x \tag{2.11}$$

$$= \frac{1}{\sqrt{2\pi\hbar}} \int e^{-\frac{ipx}{\hbar}} \psi(x) \mathrm{d}x. \tag{2.12}$$

We have also used $\langle p|x\rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{-ipx/\hbar}$. Note that this is a Fourier transform between x and $\frac{p}{\hbar} = k$. Continuous conjugate variables in quantum mechanics are all related by Fourier transforms.

The same ideas hold in the discrete case: conjugate variables are related by discrete Fourier transforms and all basis changes are described by unitary operators.

2.2.1 Quantum Information Conventions

Qubits are normalized two-vectors with complex elements. A qubit is one natural extension of the concept of the digital bit to quantum mechanical systems. A classical bit is a binary digit that can store the value 0 or the value 1. The next extension of this is a *p-bit* or probabilistic bit which can take any value in the range 0 to 1. This was introduced to characterize probabilistic computation — a class of computation that solves problems with some arbitrarily small constant probability of error, ϵ. Many problems can currently be solved much more efficiently if the user tolerates some small (and remember, ϵ can be chosen by the user) probability of error. If the bit is considered to be one of two points and the p-bit to be a point on a line between those points, then a qubit lives in and on a sphere whose poles are the states $|0\rangle$ and $|1\rangle$, which can be written as the normalized complex two-element vectors $\binom{1}{0}$ and $\binom{0}{1}$. This large freedom for the qubit state occurs because terms in the quantum state superpositions are allowed to have complex coefficients, *i.e.*, in general the matrix elements can be complex.

It can be seen that this mapping from normalized vectors in $\mathbb{C}^2$ ignoring global phases[2] to S^2 the surface of the unit sphere in three dimensions. A point on the sphere's surface is given by a pair of Euler angles (ϕ, θ) (the radius of the sphere is 1) where $\phi \in [0, 2\pi)$ is an angle in the x, y plane and $\theta \in [0, \pi)$ is the declination from the z axis. A qubit can also be fully specified by two angles (χ, ξ), since any qubit state can be written:

$$\begin{bmatrix} \cos(\chi/2) \\ \sin(\chi/2)e^{i\xi} \end{bmatrix}$$

where $\xi \in [0, 2\pi)$ and $\chi \in [0, \pi)$, so the mapping that takes qubit states to states on the surface of the unit sphere is simply $\xi \mapsto \phi$ and $\chi \mapsto \theta$.

The extreme points on the sphere of the three axes (x, y, and z) are frequently used as bases for measurement and the operators corresponding to these three bases

[2]Global phases are ignored because there is no physical measurement that will distinguish them.

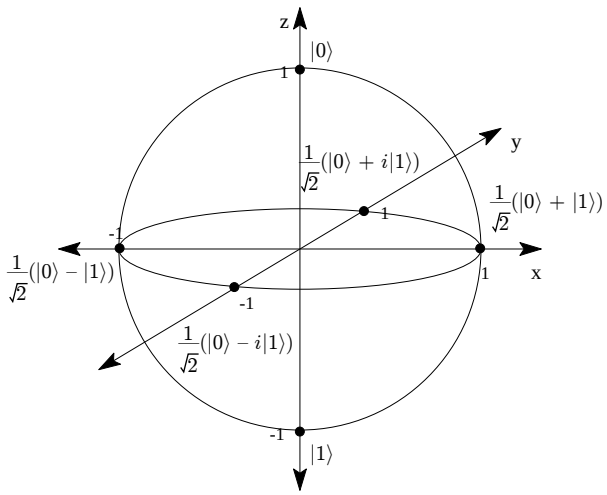


Figure 2.1: The Bloch sphere.

are the *Pauli operators* and written $\hat{X}, \hat{Y}, \hat{Z}$, or are sometimes defined with factors of $\frac{\hbar}{2}$ and then called the *Pauli spin matrices* and denoted $\hat{\sigma_x}, \hat{\sigma_y}, \hat{\sigma_z}$. ($\sigma_k = \frac{\hbar}{2}K$, where $k = \{x, y, z\}$.)

The $\{|0\rangle, |1\rangle\}$-basis is often called the "computational basis" and these two states are the eigenstates of the Pauli $\hat{Z}$ matrix:

$$\hat{Z} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{2.13}$$

Define

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{2.14}$$

The eigenvalues are $+1$ and -1 respectively.

The Pauli Y matrix is

$$\hat{Y} = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \tag{2.15}$$

The eigenvectors of Y are

$$|i+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i\,|1\rangle) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} \text{ and} \tag{2.16}$$

$$|i-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - i\,|1\rangle) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix}. \tag{2.17}$$

These states form the sphere's poles on the y-axis.

The states on the surface of the sphere along the x-axis are

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \text{ and} \tag{2.18}$$

$$|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \tag{2.19}$$

which are the eigenstates of the Pauli X matrix is given by:

$$\hat{X} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{2.20}$$

States on the surface of the sphere are called *pure states* and can be written as a vector, $|\psi\rangle$. However, there are also valid states which are incoherent mixtures of such pure states, called *mixed state*. These states are not on the surface of the Bloch sphere, but describe the interior of the sphere. It is useful to have a means of expressing both types of states with a single mathematical object. For this purpose, the *density matrix*, ρ, is introduced. For a pure state $|\psi\rangle$, ρ is formed by taking the outer product of $|\psi\rangle$ with itself: $\rho = |\psi\rangle \langle\psi|$.

Now consider the case where there is a $\frac{1}{2}$ probability of being in the state $|0\rangle$ and a $\frac{1}{2}$ probability of being in the state $|1\rangle$. This has the density matrix:

$$\hat{\rho} = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{2.21}$$

The terms on the diagonals correspond to probabilities. The off-diagonal elements have no corresponding physical interpretation in the classical world. The process by which the off-diagonals become zero is called *decoherence*. For example, if we perform a measurement and discard the outcome, then the pure state $|\psi\rangle$ is updated to a mixed state to reflect our uncertainty of the outcome:

$$|\psi\rangle \longrightarrow \rho = \sum_i \Pi_i |\psi\rangle \langle\psi| \Pi_i. \tag{2.22}$$

For any density matrix, pure or mixed, the trace is always 1. This ensures that the total probability summed over all possible outcomes is 1. For a pure state density matrix $\mathrm{Tr}[\rho^2] = 1$. All density matrices have the further properties that they are Hermitian: $\rho = \rho^\dagger$, so that measurements on the state produce real-

number-valued outcomes, and they are positive-semidefinite: $\langle v|\,\rho\,|v\rangle \geq 0$ where $|v\rangle$ is any normalized vector, so that the probability of any outcome occurring is never negative.

Measurements and operations can also be generalized from the case described in the previous section. Firstly, observable measurements, Λ, can be performed on density operators, ρ. This is represented mathematically as $\sum_i \Pi_i \rho \Pi_i$ with $P(\lambda_i) = \frac{\Pi_i \rho \Pi_i}{\mathrm{Tr}[\Pi_i \rho]}$ and the expectation value is given by $\langle \lambda \rangle_\rho = \mathrm{Tr}[\Lambda \rho]$. A more general concept of measurement is the POVM or positive operator-valued measure. A POVM has elements E_i which obey $\sum_i E_i = \mathbb{I}$ and the probabilities of the various outcomes of the POVM are $P(\lambda_i) = \mathrm{Tr}[E_i \rho]$. In the special case that the E_i are projectors, this is an observable, but they need not be. There is no requirement of orthogonality, they only must sum to the identity and each be positive semi-definite operators so that the probabilities of the outcomes are always positive and sum to one. Physically, a POVM can always be realized by coupling to another system and performing a projective measurement over both systems where the dimension of the ancilla system the equal to the number of elements describing the POVM. This follows from Naimark's Dilation Theorem [Nai43].

2.3 Quantum Optics

2.3.1 Quantum Harmonic Oscillator

An important model, which shall be used as a model for an optical cavity, is the quantum harmonic oscillator. A vibrating spring is an example of such an oscillator. The Hamiltonian is

$$H = \frac{1}{2} m \left(\frac{\mathrm{d}x}{\mathrm{d}t} \right)^2 + \frac{1}{2} k x^2 \tag{2.23}$$

where k is the spring constant. For our current purposes, we shall re-write this in an entirely equivalent form (noting that $p = m\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)$, since momentum is the generator of translations in position), now with quantum operators for x, p, and H:

$$H = \frac{p^2}{2m} + \frac{1}{2} m \omega^2 x^2 \tag{2.24}$$

where ω is the circular frequency of oscillation. Such a system resides in a quadratic potential well. This can be seen from the potential term of the Hamiltonian which is the second term of equation (2.24), which is written in the form, $H = T + V$,

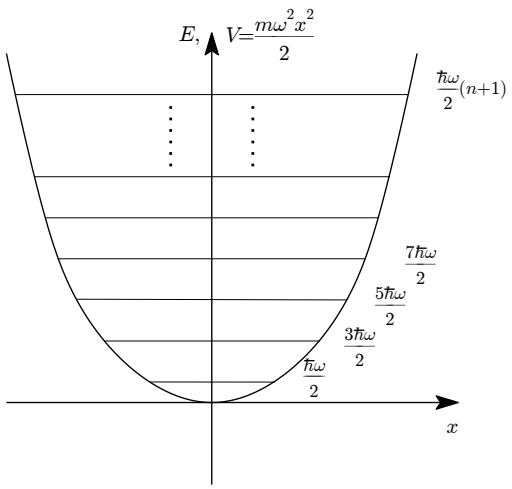


Figure 2.2: The form of the harmonic potential well and its energy eigenvalues.

where T is the kinetic energy and V is the potential energy. The shape of the quadratic curve shown in Figure 2.2. The Schrödinger equation for this system is

$$\left(\frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2\right)|n\rangle = E_n\,|n\rangle\,, \tag{2.25}$$

where the $|n\rangle$ are the energy eigenstates and the E_n are the energy eigenvalues. The eigenstates include a lowest-energy ground state $|0\rangle$ and all state $|n\rangle$ where n is a positive integer. These states form a countably infinite orthonormal basis for the Hilbert space, which is the space L^2 of all square-integrable functions, and has an inner product defined by $\langle f|g\rangle = \int \big(f(x)\big)^* g(x)\mathrm{d}\mu(x)$. The forms of these states as functions of the continuous variable for position of a particle in a Harmonic potential will be considered later in this section.

The energy levels can be found by solving the Schrödinger equation, imposing a parabolic potential, $\frac{1}{2}m\omega^2x^2$. However, it is easier, more useful, and more intuitively helpful, to instead make a change of operators in the Hamiltonian. Let us define

$$a = \sqrt{\frac{m\omega}{2\hbar}}\left(x + \frac{ip}{m\omega}\right) \tag{2.26}$$

$$a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\left(x - \frac{ip}{m\omega}\right) \tag{2.27}$$

noting that $a^\dagger$ is the conjugate of a. Also note that x and p are measurement operators, and therefore are Hermitian, so $x = x^\dagger$ and $p = p^\dagger$. Consider the value

of the commutator of these operators.

$$\begin{aligned}
[a, a^\dagger] &= \frac{m\omega}{2\hbar}\left[x + \frac{ip}{m\omega}, x - \frac{ip}{m\omega}\right] \\
&= \frac{1}{2\hbar}(-i[x,p] + i[p,x]) \\
&= \frac{1}{2\hbar}(-i(i\hbar) + i(-i\hbar)) \\
&= 1. \qquad (2.28)
\end{aligned}$$

From this it can be seen that for any eigenstate $|n\rangle$

$$\langle n| \left[a, a^\dagger\right] |n\rangle = \langle n| \left(aa^\dagger - a^\dagger a\right) |n\rangle = \langle n|n\rangle = 1, \qquad (2.29)$$

which implies

$$\langle n| aa^\dagger |n\rangle = \langle n| a^\dagger a |n\rangle + 1. \qquad (2.30)$$

The operator $a^\dagger a$ is called the *number operator*, N, and has a special property, which shall be considered momentarily. Explicitly in terms of x and p, it is

$$\begin{aligned}
N = a^\dagger a &= \frac{m\omega}{2\hbar}\left(x^2 + \left(\frac{p}{m\omega}\right)^2\right) + \frac{i}{2\hbar}[x,p] \\
&= \frac{1}{\hbar\omega}\left(\frac{p^2}{2m} + \frac{m\omega^2 x^2}{2}\right) + \frac{i}{2\hbar}[x,p] \\
&= \frac{1}{\hbar\omega}H - \frac{1}{2}. \qquad (2.31)
\end{aligned}$$

So, the Hamiltonian can be expressed as

$$H = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right). \qquad (2.32)$$

The form of this expression is very important. It implies that any eigenstates of H (the $|n\rangle$) are also eigenstates of $N = a^\dagger a$. The eigenvalues of N are given by the equation

$$N|n\rangle = n|n\rangle. \qquad (2.33)$$

Then the energy levels of this system are given by

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right). \qquad (2.34)$$

Equation (2.30) can now be written as

$$\langle n| a\hat{a}^\dagger |n\rangle = n + 1, \tag{2.35}$$

which means that the state $a^\dagger |n\rangle$ must be of the form

$$a^\dagger |n\rangle = \sqrt{n+1}\,|\phi\rangle \tag{2.36}$$

where $|\phi\rangle$ is a normalized state that can now be found. Finding the expectation value of the number operator N for the state $\sqrt{n+1}\,|\phi\rangle$ to determine if it is an eigenstate and what its eigenvalue is:

$$\begin{aligned}
\left(\langle\phi|\sqrt{n+1}\right) N \left(\sqrt{n+1}\,|\phi\rangle\right) &= \left(\langle n|\, a\right)\left(a^\dagger a\right)\left(a^\dagger |n\rangle\right) \\
&= \langle n|\left(aa^\dagger\right)\left(aa^\dagger\right)|n\rangle \\
&= \langle n|\left(a^\dagger a + 1\right)\left(a^\dagger a + 1\right)|n\rangle \\
&= \langle n|\left(a^\dagger a\right)^2 |n\rangle + 2\left(\langle n|\, a^\dagger a\,|n\rangle\right) + 1 \\
&= n^2 + 2n + 1 \\
(n+1)\,\langle\phi|\, N\,|\phi\rangle &= (n+1)^2.
\end{aligned}$$

Therefore,

$$\langle\phi|\, N\,|\phi\rangle = (n+1) \Rightarrow |\phi\rangle = |n+1\rangle \tag{2.37}$$

(where superpositions of energy eigenstates with an expectation value of $n+1$ are excluded because they could not give $aa^\dagger |n\rangle = (n+1)\,|n\rangle$ which equation (2.35) asserts and we know$\langle n|n'\rangle = 0$ for $n \neq n'$ since N is Hermitian, see equation 2.9) and we conclude

$$a^\dagger |n\rangle = \sqrt{n+1}\,|n+1\rangle\,. \tag{2.38}$$

Similar logic also gives

$$a\,|n\rangle = \sqrt{n}\,|n-1\rangle\,. \tag{2.39}$$

The operators a and $a^\dagger$ are called the *annihilation* and *creation* operators respectively. They have the useful property that they transfer the system from one energy eigenstate to the one immediately above it (for the case of the creation operator) or the one immediately below it (for the case of the annihilation operator). Because of this property they are also sometimes called ladder operators. This can be interpreted physically in the following way. Consider a laser cavity, which is a tube with a nearly prefect reflecting mirror at each end. It may contain some fixed number (assuming the cavity is lossless) of photons, which are understood to be

electromagnetic field excitations, and the creation operator can be used to describe mathematically adding a single photon to the cavity or the annihilation operator to describe removing one. Alternatively, the harmonic oscillator can be thought of as an abstract energy bath that changes its energy eigenstate by emitting or absorbing photons. The harmonic oscillator is the most commonly used model for describing the interaction of quantum fields with each other and with matter.

Assume that the harmonic potential has a finite minimum value, and therefore there must be some ground energy state of the system, call it $|0\rangle$. Such a state cannot evolve to another energy state under the action of the annihilation operator, so it must be that:

$$a\,|0\rangle = 0. \tag{2.40}$$

The energy associated with this state is clearly

$$E_0 = \langle 0|\,\hbar\omega\left(N + \frac{1}{2}\right)|0\rangle = \frac{\hbar\omega}{2}. \tag{2.41}$$

Note that this is not zero. If one has an optical cavity with no photons in it, the "vacuum field" within the cavity still contains non-zero energy.

The expectation values of observables are not the only quantities of interest. The expected fluctuations – the variance – of these variables can also be found. For an arbitrary operator A, the variance of the outcome of a measurement A for a state ψ is denoted $\langle(\Delta\hat{A})^2\rangle_\psi$ and given by:

$$\langle(\Delta A)^2\rangle_\psi = \langle\psi|\,\hat{A}^2\,|\psi\rangle - \langle\psi|\,\hat{A}\,|\psi\rangle^2\,. \tag{2.42}$$

Using an electric field of the form

$$E = \lambda(a + a^\dagger), \tag{2.43}$$

(where λ is just a constant) the fluctuation of this field operator for a state of $|n\rangle$ photons is

$$\langle(\Delta E)^2\rangle_n = \langle n|\,E^2\,|n\rangle - \langle n|\,E\,|n\rangle^2\,. \tag{2.44}$$

$$\begin{aligned}
\langle(\Delta E)^2\rangle_n &= \langle n|\, E^2\, |n\rangle - \langle n|\, E\, |n\rangle^2 \\
&= \lambda^2 \langle n|\, (a+a^\dagger)^2\, |n\rangle - \lambda^2 \langle n|\, (a+a^\dagger)\, |n\rangle^2 \\
&= \lambda^2 \langle n|\, (a^2 + aa^\dagger + a^\dagger a + (a^\dagger)^2)\, |n\rangle - \lambda^2 \left(\langle n|\, a\, |n\rangle + \langle n|\, a^\dagger\, |n\rangle\right)^2 \\
&= \lambda^2(0 + \langle n|\, aa^\dagger\, |n\rangle + \langle n|\, a^\dagger a\, |n\rangle + 0) - \lambda^2(0+0)^2 \\
&= \lambda^2(n+1+n) \\
&= \lambda^2(2n+1).
\end{aligned}$$

Notice that when $n = 0$

$$\langle(\Delta E)^2\rangle_{n=0} = \lambda^2$$

Even in a vacuum electric field, the field fluctuations are non-zero. That is, a measurement of the electric field sometimes finds that it has a non-zero value, even though the field is in a ground state.

The energies of the excited states, E_n, are of course

$$E_n = \langle n|\, H\, |n\rangle = \langle 1|\, \hbar\omega \left(N + \frac{1}{2}\right) |1\rangle = \frac{(2n+1)\hbar\omega}{2}. \tag{2.45}$$

The form of the energy eigenstates can also be found in terms of the quadratures, x and p. Using the Hamiltonian expressed in the form $H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2$ and noting that $p = -i\frac{\mathrm{d}}{\mathrm{d}x}$, it can be re-expressed as a differential operator in x:

$$\begin{aligned}
H &= \frac{1}{2m}\left(-i\hbar\frac{\mathrm{d}}{\mathrm{d}x}\right)^2 + \frac{1}{2}m\omega^2 x^2 \\
&= -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{1}{2}m\omega^2 x^2.
\end{aligned}$$

For $E = E_0$, $\psi_0(x) = \frac{1}{\sqrt[4]{\pi}} e^{-\frac{m\omega}{2\hbar}x^2}$ is the normalized ground eigenstate.

And, in general for $E = E_n$ we obtain the equation

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} f_n(x) - 2xs\frac{\mathrm{d}}{\mathrm{d}x} f_n(x) + 2ns f_n(x) = 0 \tag{2.46}$$

where $s = \frac{m^2\omega^2}{\hbar^2}$ and let $\psi_n(x) = f_n(x) e^{-\frac{s}{2}x^2}$. This is the definition for the "physicist" Hermite polynomials (with an extra factor of s, which is usually set to 1 by choosing dimensionless parameters). The Hermite polynomials are also defined by:

$$H_n(x) = (-1)^n e^{-x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} e^{-x^2}. \tag{2.47}$$

Including s, this makes the general eigenstates:

$$\psi_n(x) = \sqrt{\frac{1}{2^n n!}\sqrt{\frac{s}{\pi}}} H_n(\sqrt{s}x) e^{-\frac{s}{2}x^2}. \tag{2.48}$$

2.3.2 Coherent States

The energy eigenstates $|n\rangle$ are sometimes called Fock states. It is also possible to define a state which is an eigenstate of the annihilation operator a.

$$a|\alpha\rangle = \alpha|\alpha\rangle \tag{2.49}$$

$$\begin{aligned} a|\alpha\rangle &= \sum_n c_n a|n\rangle \\ &= \sum_n c_n\sqrt{n}|n-1\rangle = \alpha\sum c_n|n\rangle \end{aligned}$$

where the second equality follows from equation 2.49, which implies that

$$c_n\sqrt{n} = c_{n-1}\alpha \tag{2.50}$$

and solving this recurrence relation gives

$$c_n = \frac{\alpha^n}{\sqrt{n!}}c_0. \tag{2.51}$$

So, $|\alpha\rangle = c_0\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle$ and c_0 is a normalization constant. Since $\sum_{n=0}^{\infty}\frac{(|\alpha|^2)^n}{n!} = e^{|\alpha|^2}$ it must be that the value of $c_0 = e^{-\frac{|\alpha|^2}{2}}$. This means that the state can be writen:

$$|\alpha\rangle = e^{-\frac{|\alpha|^2}{2}}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle = e^{-\frac{\bar{n}}{2}}\sum_{n=0}^{\infty}\sqrt{\frac{\bar{n}^n}{n!}}e^{i\phi n}|n\rangle \tag{2.52}$$

where the second equality comes from defining $\alpha = \sqrt{\bar{n}}e^{i\phi}$ where $\bar{n}$ is a real number. Noticing that $\frac{\bar{n}^n}{n!}$ is the form of the Poisson distribution with mean $\bar{n}$ implies that is the mean excitation number of the state is $\bar{n}$. The argument ϕ is the phase of the coherent state.

Coherent states have equal uncertainty in both quadratures, x and p, and achieve the Heisenberg limit for joint uncertainty. (See Figure 2.3.) For conve-

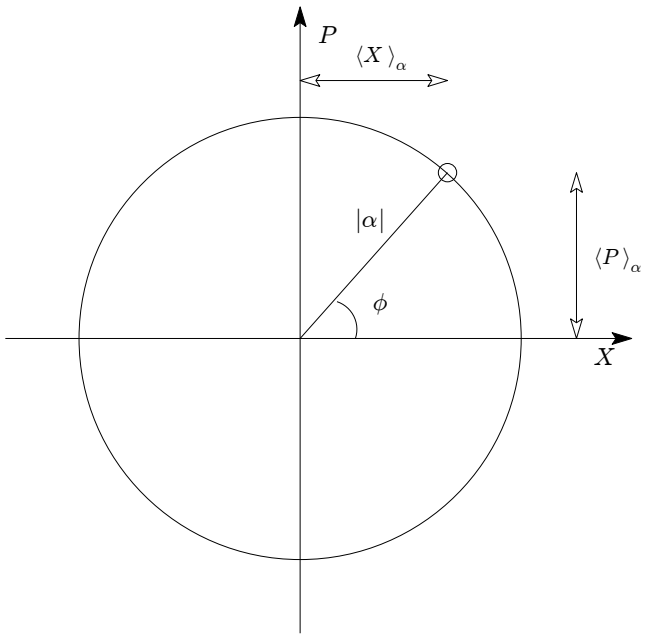


Figure 2.3: A phase space diagram of the coherent state, $|\alpha\rangle$. Such a state has a Gaussian profile and equal spread in position and momentum. The expectation values of these *quadrature* observables, the projections of the magnitude of the state onto the axes, is illustrated.

nience, consider the dimensionless definitions of the quadratures $X = \sqrt{\frac{m\omega}{2\hbar}}\, x = \sqrt{\frac{1}{2}}(a + a^\dagger)$ and $P = \sqrt{\frac{1}{2\hbar m\omega}}\, p = \sqrt{\frac{i}{2}}(\alpha - \alpha^*)$

$$\begin{aligned} \langle\alpha|\, X\, |\alpha\rangle &= \langle\alpha| \sqrt{\frac{1}{2}}(a + a^\dagger)\, |\alpha\rangle \\ &= \sqrt{\frac{1}{2}}(\langle\alpha|\, a\, |\alpha\rangle + \langle\alpha|\, a^\dagger\, |\alpha\rangle) \\ &= \sqrt{\frac{1}{2}}(\alpha + \alpha^*) \end{aligned} \tag{2.53}$$

and similarly for P, $\langle\alpha|\, P\,|\alpha\rangle = \sqrt{\frac{i}{2}}(\alpha - \alpha^*)$. Also,

$$\begin{aligned}
\langle\alpha|\, X^2\,|\alpha\rangle &= \langle\alpha|\, \frac{1}{2}(a + a^\dagger)^2\,|\alpha\rangle \\
&= \frac{1}{2}(\langle\alpha|\, aa\,|\alpha\rangle + \langle\alpha|\, aa^\dagger\,|\alpha\rangle + \langle\alpha|\, a^\dagger a\,|\alpha\rangle + \langle\alpha|\, a^\dagger a^\dagger\,|\alpha\rangle) \\
&= \frac{1}{2}(\alpha^2 + \langle\alpha|\, (a^\dagger a + 1)\,|\alpha\rangle + |\alpha|^2 + (\alpha^*)^2) \\
&= \frac{1}{2}(\alpha^2 + 1 + 2|\alpha|^2 + (\alpha^*)^2). \qquad (2.54)
\end{aligned}$$

Therefore,

$$\begin{aligned}
\Delta X^2 &= \langle X^2\rangle - \langle X\rangle^2 \qquad (2.55) \\
&= \frac{1}{2}. \qquad (2.56)
\end{aligned}$$

And for P, $\langle\alpha|\, P^2\,|\alpha\rangle = \frac{1}{2}(-\alpha^2 + 1 + 2|\alpha|^2 - (\alpha^*)^2)$:

$$\begin{aligned}
\Delta P^2 &= \langle P^2\rangle - \langle P\rangle^2 \qquad (2.57) \\
&= \frac{1}{2}. \qquad (2.58)
\end{aligned}$$

As already surmised, the mean number of excitations $\langle\alpha|\, N\,|\alpha\rangle$ is $\bar{n} = |\alpha|^2$. The variance is

$$\begin{aligned}
\Delta N^2 &= \langle N^2\rangle - \langle N\rangle^2 \\
&= \langle\alpha|\, a^\dagger a a^\dagger a\,|\alpha\rangle - (|\alpha|^2)^2 \\
&= \langle\alpha|\, a^\dagger (a^\dagger a + 1) a\,|\alpha\rangle - |\alpha|^4 \\
&= |\alpha|^4 + |\alpha|^2 - |\alpha|^4 \\
&= |\alpha|^2 = \bar{n} \qquad (2.59)
\end{aligned}$$

meaning that the mean and the variance are equal, as should be the case for a Poissonian distribution.

There is no observable operator for phase, ϕ, the variable conjugate to excitation number, n. A observable operator should be Hermitian with real eigenvalues corresponding to the measurement outcomes and associated eigenvectors that span the state space. It is not possible to construct such an operator, but since a coherent state is a minimum-uncertainty state, the variance should be $\frac{1}{4\bar{n}}$ – the inverse of the variance for N times the Heisenberg limit ($\frac{1}{2}$) squared, which in these units

is $\frac{1}{4}$.

One other point to notice is that coherent states are also intuitive states for indicating phase. Though true coherent states are states of infinite dimensional Hilbert spaces, for large values of $\bar{n}$ cropping the summation at some value of n, say $n = N$, they are very close to states which are tensor products of qubits in identical states of the form $(|0\rangle + e^{i\phi}|1\rangle)$. That is as $n \to \infty$ the state

$$(|0\rangle + e^{i\phi}|1\rangle)^{\otimes N}$$

approaches an excitation distribution equivalent to

$$e^{-N/4} \sum_{n=0}^{\infty} \sqrt{\frac{(N/2)^n}{n!}} e^{in\phi} |n\rangle$$

where $|n\rangle$ denotes the symmetric bosonic state of n excitations and $|1\rangle$ denotes the excited state of a single qubit. Here, $\bar{n} = N/2$. This result is due to the Central Limit Theorem [Boa83] which gives the convergence of the Poisson distribution (the distribution of the number of excitations in the coherent state) and the binomial distribution (the distribution of the number of excitations in the tensor-product state) to the Gaussian distribution as the number of samples and averages become large.

2.4 Quantum Angular Momentum

Angular momentum is defined as the cross product between the momentum p and r, the vector to the origin of the rotation:

$$L = p \times r, \tag{2.60}$$

so it is a momentum about a fixed point. Classical systems can have two kinds of angular momentum: *extrinsic*, or *orbital*, angular momentum, which is due to a massive object moving around a point in space, and *intrinsic* angular momentum, which is the momentum a body with extent has by virtue of rotating about its centre of mass. Like classical systems, quantum mechanical systems can have angular momentum, and two different kinds: *orbital* angular momentum, and *spin* angular momentum. Spin angular momentum is considered to be a form of intrinsic angular momentum, but particles such as electrons have non-zero spin, despite the fact that they are believed to be point particles. Specifically, the radius of the electron has

been shown to be less than 10^{-3}fm by high energy scattering experiments. A simple analysis of orders of magnitude reveals that even if the electron had physical extent on this scale or smaller, in order to have the value of intrinsic angular momentum it is found to have, the electron would need to be rotating so rapidly that the outside edges of the electron would travel faster than the speed of light, as is considered presently:

The magnetic moment $\mu = \frac{1}{2}gs$ is proportional to the electron spin, s. The electronic magnetic moment can also be written in terms of the Bohr magneton $\mu_B = \frac{e\hbar}{2m_e}$:

$$\mu_e = \pm\frac{1}{2}g\mu_B \tag{2.61}$$

where g is the gyromagnetic ratio, which for electron spin is slightly greater than 2. This means that we can use the Bohr magneton as the electron's magnetic moment in this rough calculation. The value of the Bohr magneton is

$$\mu_B = 9.27 \times 10^{-24}\text{J T}^{-1}. \tag{2.62}$$

The largest magnetic moment for a charged body occurs if the charge is all located as far from the centre of mass (the centre of rotation) as possible. Therefore, to be as permissive as possible, we shall suppose that the charge of the electron, rather than being evenly distributed, is all located at the outer edge of the electron. Magnetic moment of an orbiting charge, q, is given by the expression

$$\mu_=\frac{1}{2}q\mathbf{r} \times \mathbf{v}. \tag{2.63}$$

The radius is perpendicular to the direction of the velocity, so we find, for the velocity of the electron's outer edge, the rough expression

$$v \approx \frac{2\mu_B}{e} \tag{2.64}$$

which gives $v \approx 11 \times 10^{13}\text{m s}^{-1}$. This is much greater than the speed of light, $c = 3.00 \times 10^{8}\text{m s}^{-1}$.

It is clear from this that even though spin is an intrinsic degree of freedom of quantum systems, it is not due to rotation of the body about its centre in the classical sense.

Orbital, $\mathbf{L}$, and spin, $\mathbf{S}$, angular momenta are vector quantities which sum to give the *total angular momentum* of the quantum system, $\mathbf{J} = \mathbf{L} + \mathbf{S}$. All of these angular momenta are quantized.

We can also consider the direction of the total angular momentum relative to a z axis defined by our measurement apparatus. The z-projection of the total angular momentum, m, is also quantized and can take the $2J+1$ different values $m \in \{-2J-1, -2J+1, ..., 0, ..., 2J-1, 2J+1\}$. Equally, the measurement device could be configured to define an x or y axis. The operators associated with the x, y, and z projections of angular momentum are typically denoted J_x, J_y, and J_z and obey the commutation relations

$$[J_i, J_j] = i\epsilon_{i,j,k} J_k, \tag{2.65}$$

where $\epsilon_{i,j,k}$ is the Levi-Civita symbol which is $+1$ for $i = x, j = y, k = z$ and any cyclic permutation of those, -1 for any other permutation of x, y, and z, and zero if any of i, j, and k are equal. All of J_x, J_y, and J_z commute with the total angular momentum operator $J^2 = J_x^2 + J_y^2 + J_z^2$, in particular

$$[J^2, J_z] = [J_x^2, J_z] + [J_y^2, J_z] = J_x(-iJ_y) + J_y(iJ_x) + (-iJ_y)J_x + (iJ_x)J_y = 0. \tag{2.66}$$

So, J^2 and J_z are compatible observables, and we can express a joint eigenbasis for them. Following standard notation, let us denote the eigenstates $|J, m\rangle$ which have the eigenvalue equations:

$$J_z |J, m\rangle = m |J, m\rangle \tag{2.67}$$

$$J^2 |J, m\rangle = j(j+1) |J, m\rangle , \tag{2.68}$$

where j can take any non-negative integer or half-integer value.

In the same way that ladder operators are constructed for the energy levels of a harmonic operator, it is useful to find raising and lowering operators for m for a fixed value of j. We require

$$J_+ |J, m\rangle = k_{j,m} |J, m+1\rangle . \tag{2.69}$$

The operator J_z does not change the state $|J, m\rangle$, but a combination of the operators J_x and J_y can. Note that

$$[J_z, J_x + iJ_y] = J_x + iJ_y \tag{2.70}$$

and

$$[J_x - iJ_y, J_z] = J_x - iJ_y. \tag{2.71}$$

From equation (2.70)

$$\begin{aligned} J_x + iJ_y \left|J,m\right\rangle &= [J_z, J_x + iJ_y] \left|J,m\right\rangle & (2.72)\\ &= (J_z - m)(J_x + iJ_y) \left|J,m\right\rangle & (2.73)\end{aligned}$$

so, for this equation to hold, $(J_x + iJ_y)\left|J,m\right\rangle$ must be an eigenstate of J_z with eigenvalue $m+1$, which means

$$(J_x + iJ_y)\left|J,m\right\rangle = k_{j,m}\left|J,m+1\right\rangle. \tag{2.74}$$

The factor $k_{j,m}$ can be found by considering the inner product

$$\left\langle J,m\right| (J_x - iJ_y)(J_x + iJ_y)\left|J,m\right\rangle = \left\langle J,m\right| (J^2 - J_z^2 - J_z)\left|J,m\right\rangle = j(j+1) - m^2 - m. \tag{2.75}$$

So,

$$J_+\left|J,m\right\rangle = \sqrt{(j-m)(j+m+1}\left|J,m+1\right\rangle. \tag{2.76}$$

where

$$J_+ = J_x + iJ_y. \tag{2.77}$$

This is the raising operator that we hoped to find. A similar calculation from equation (2.71) gives

$$(J_x - iJ_y)\left|J,m\right\rangle = \sqrt{(j+m)(j-m+1}\left|J,m-1\right\rangle, \tag{2.78}$$

which gives the expression for the lowering operator:

$$J_- = J_x - iJ_y. \tag{2.79}$$

These expressions also indicate how J_x and J_y act on the state $\left|J,m\right\rangle$:

$$\begin{aligned} J_x\left|J,m\right\rangle &= \frac{1}{2}(J_+ + J_-)\left|J,m\right\rangle = \frac{1}{2}(\left|J,m+1\right\rangle + \left|J,m-1\right\rangle) & (2.80)\\ J_y\left|J,m\right\rangle &= \frac{i}{2}(J_- - J_+)\left|J,m\right\rangle = \frac{i}{2}(\left|J,m-1\right\rangle - \left|J,m+1\right\rangle). & (2.81)\end{aligned}$$

Since the total angular momentum of a system is quantized, when two systems of non-zero angular momentum combine there are rules governing how the angular momentum vectors can add to produce a new system with new total angular momentum. For example consider a pair of systems with total angular momentum $\mathbf{j_1}$ and $\mathbf{j_2}$ which will be combined under the tensor product to give a single system of

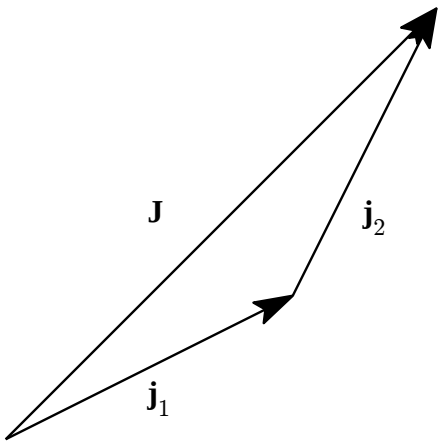


Figure 2.4: Combining two systems with angular momentum $\mathbf{j_1}$ and $\mathbf{j_2}$ to form a joint system with total angular momentum $\mathbf{J}$.

angular momentum $\mathbf{J} = \mathbf{j_1} + \mathbf{j_2}$ and let $j_1 = |\mathbf{j_1}| \geq j_2$. The allowed values of the outcome of a measurement of the total angular momentum of the joint system are the integer or half-integer values:

$$J = \{j_1 - j_2, j_1 - j_2 + 1, ..., j_1 + j_2\}, \tag{2.82}$$

where for $k = \{x, y, z\}$

$$J_k = j_{1k} \otimes \mathbb{I} + \mathbb{I} \otimes j_{2k}. \tag{2.83}$$

The amplitudes for each of these possible outcomes depends on the relative directions of the systems' individual angular momentum vectors and are given by the Clebsch-Gordan coefficients $\langle j_1 m_1, j_2 m_2 | JM \rangle$ [CS51, Hec00]:

$$|j_1 m_1, j_2 m_2\rangle = \sum_j \langle JM | j_1 m_1, j_2 m_2 \rangle \, |JM\rangle \tag{2.84}$$

and since the sum is only over j (M, m_1, and m_2 are fixed) these coefficients can always be made real, and this is the universal convention. Then, in the case that the direction of the total joint angular momentum defines the z axis, there is a

general expression [Hec00]:

$$\langle j_1 m_1, j_2 (J - m_1)_2 | JJ \rangle = (-1)^{j_1 - m_1} \times \sqrt{\frac{(j_1 + m_1)!(j_2 + J - m_1)!(j_1 + j_2 - J)!(2J+1)!}{(j_1 - m_1)!(j_2 - J + m_1)!(j_2 - j_1 + J)!(j_1 - j_2 + J)!(j_1 + j_2 + J + 1)!}}. \tag{2.85}$$

In other cases, the particular expressions required can be calculated from recursion relations derived by applying the raising operator J_+ to both sides of equation 2.84.

2.5 Group Theory

Group theory is a powerful formalism for exploiting symmetries in the analysis of physical systems. It will be used in this thesis to describe the symmetries that result from lacking certain reference frames in Chapters 4 and 3.

A *group* is formed from a set and an operation that combines two element to form a third. The set should be closed, associative, have an identity element, and each element should have an inverse element.

Groups can consist of (or have representations as) sets of matrices.

A *Lie group* is a group that is also a differentiable manifold, so that it can describe continuous symmetries, and the group operation and inverse map for an element should be differentiable. A matrix Lie group is a closed subgroup of the general linear matrix group, $GL(n;\mathbb{C})$, the group of all $n \times n$ invertible matrices.

A unitary matrix is a matrix U such that $UU^\dagger = \mathbb{I}$. The unitary group $U(n)$ is the group of all $n \times n$ unitary matrices. The group $SU(n)$ is the subset of $U(n)$ such that all the elements have determinant one. The orthogonal group $O(n)$ contains all $n \times n$ matrices that are orthogonal, *i.e.*, $OO^{\mathrm{T}} = \mathbb{I}$ and the subgroup with determinant one is denoted $SO(n)$. These are all matrix Lie groups.

A homomorphism is a map between two algebraic structures that preserves that structure. Such a map is a *Lie group homomorphism* if it is a map from a Lie group G to a Lie group H that is a group homomorphism and is continuous. If such a map is also one-to-one and onto and its inverse map is continuous, then it is a *Lie group isomorphism*, and the groups G and H can be written $G \simeq H$.

A *Lie algebra*, $\mathfrak{g}$, of a matrix Lie group, G, is the set of all matrices A, such that

$e^{iAt} \in G$ for all $t \in \mathbb{R}$. [3]

For the case of the group $SU(n)$, the group of all $n \times n$ unitary matrices with determinate 1, it must be that

$$U^{-1} = U^{\dagger} \Rightarrow e^{-iAt} = (e^{iAt})^{\dagger}. \tag{2.86}$$

Expanding the exponential function as a Taylor series, since the operation of taking the adjoint is additive, it can be seen that $(e^{iAt})^{\dagger} = e^{-iA^{\dagger}t}$, so

$$A = A^{\dagger}, \tag{2.87}$$

which gives that the Lie algebra $\mathfrak{su}(n)$ is composed of Hermitian matrices. In the setting of quantum mechanics, this shows that Hamiltonians, which are the generators of unitary transformations, are Hermitian operators.

A finite-dimensional matrix *representation* of a Lie group, G, is an homomorphism, Π from G to the general linear matrix group $GL(n;\mathbb{C})$ on some dimension n. If the mapping is one-to-one, then it is a *faithful representation*. A representation is *irreducible* if it has no non-trivial invariant subspaces. A subspace, S, of a vector space, V, is non-trivial if $S \neq \{0\}$ and $S \neq V$. Such a subspace is invariant under the action of a group representation if the representations of all of the group elements map all vectors from the subspace back into the subspace, *i.e.*, $\Pi(g)s \in S$ for all $s \in S$ and $g \in G$.

2.5.1 The Relation of $SU(2)$ to $SO(3)$

The group $SU(2)$ is the group of all unitary matrices with determinant one of dimension two. The group $SO(3)$ is the set of all rotations about some three-space vector $\hat{n}$, through an angle $\theta \in [0, 2\pi)$.

There is a homomorphism that takes the elements of $SU(2)$ to $SO(3)$ that is two-to-one. This can be seen by observing that two elements of $SO(3)$ are the same if they are $R_O(\hat{n}, \theta)$ and $R_O(-\hat{n}, 2\pi - \theta)$. However, $SU(2)$ can also be represented as rotations $R_U(\hat{n}, \theta)$, however, rotations about an vector $\hat{n}$ through an angle 2θ. So for $SU(2)$, $R_U(\hat{n}, \theta) \neq R_U(-\hat{n}, 2\pi - \theta)$.

More explicitly, a basis for a three dimensional real vector space, $\mathbf{i}, \mathbf{j}, \mathbf{k}$ is a set of three orthogonal unit vectors. We could chose to represent $\mathbf{i} = X, \mathbf{j} = Y$

[3] The imaginary unit i is included in this definition according to the convention used by physicists.

and $\mathbf{k} = Z$ (X,Y,Z are the Pauli matrices). Then $\mathbf{v} = v_i\mathbf{i} + v_j\mathbf{j} + v_k\mathbf{k}$ becomes $V = v_iX + v_jY + v_kZ$.

Following Hall [Hal06], operations that are representations of the orthogonal group will will preserve the inner product of two vectors when the operation acts on both vectors, that is, $\langle w| O^T O |v\rangle = \langle w|v\rangle$. In the new basis for $\mathbb{R}^3$, $\langle w|v\rangle = \frac{1}{2}\mathrm{Tr}[W^\dagger V]$. So an orthogonal map on this representation, Φ_U is one that

$$\frac{1}{2}\mathrm{Tr}[\Phi_U(W^\dagger)\Phi_U(V)] = \frac{1}{2}\mathrm{Tr}[W^\dagger V].$$

Such maps are unitary maps on the matrix representations:

$$\Phi_U(V) = UVU^\dagger.$$

So there is a correspondence between the set of orthogonal operators on $\mathbb{R}^3$ (the elements of $SO(3)$) and the unitary matrices on two dimensions $SU(2)$. Notice that the map $\Phi_U = \Phi_{-U}$, but $U \neq -U$. This implies the mapping from $SU(2)$ to $SO(3)$ is two-to-one.

It is also possible to relate the Lie algebras of $SU(2)$ and $SO(3)$, $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$. It has been already shown that for unitary matrices the elements of the Lie algebra should be Hermitian. Further, for a matrix with determinant one, $\det(U) = 1 \Rightarrow \det(e^{iAt}) = 1$. Hermitian matrices are diagonalizable, so the trace of A is the sum of its eigenvalues $\sum_i a_i$ and its determinant of e^{iA} is the product of the exponentiated eigenvalues $\prod_i e^{a_i}$. Since $e^{\sum_i a_i} = \prod_i e^{a_i}$, therefore $\det(e^{iAt}) = e^{i\mathrm{Tr}[A]t}$. This will equal one for all real t if and only if $\mathrm{Tr}[A] = 0$. So $\mathfrak{su}(2)$ is the space of all 2×2 complex traceless Hermitian matrices.

Such matrices have the form:

$$A = a_1X + a_2Y + a_3Z,$$

where $|a_1|^2 + |a_2|^2 + |a_3|^2 = 1$ and $a_1, a_2, a_3 \in \mathbb{R}$.

For $\mathfrak{so}(3)$,

$$O^{-1} = O^{\mathrm{T}} \Rightarrow e^{-iBt} = (e^{iBt})^T. \tag{2.88}$$

Again, from the Taylor series of the exponential function and additivity of taking the transpose, $(e^{iBt})^\dagger = e^{iB^{\mathrm{T}}t}$, so

$$-B = B^{\mathrm{T}}, \tag{2.89}$$

and therefore, these are also a traceless matrices with elements

$$B = \begin{pmatrix} 0 & b_{12} & b_{13} \\ b_{21} & 0 & b_{23} \\ b_{31} & b_{32} & 0 \end{pmatrix}.$$

Then $b_{12} = -b_{21}, b_{13} = -b_{31}$, and $b_{23} = -b_{32}$ where $|b_{12}|^2 + |b_{13}|^2 + |b_{23}|^2 = 1$ so that the inner product of a real 3-vector is preserved and all of the matrix elements are real.

There are the same free parameters to define the elements of both Lie algebras. This demonstrates that there can be a one-to-one and onto mapping from $\mathfrak{su}(2)$ to $\mathfrak{so}(3)$ that takes $a_1 \mapsto b_{12}, a_2 \mapsto b_{13}$, and $a_3 \mapsto b_{23}$. Therefore, $\mathfrak{su}(2)$ is isomorphic to $\mathfrak{so}(3)$.

2.5.2 Group Theory Applications to Reference Frames and Angular Momentum

The symmetries that occur in the angular momentum states as spins are tensored together are conveniently described in the framework of group theory.

A Hilbert space $\mathcal{H}$, can be decomposed as a direct sum of spaces corresponding to inequivalent representations of Π_q of a Lie group, G,

$$\mathcal{H} = \bigoplus_q \mathcal{H}_q. \tag{2.90}$$

These spaces can each be written as the tensor product of a space that hold irreducible representations of G, Π_q, and a multiplicity space that hold trivial representations of G, $\mathcal{M}_q$, $\mathcal{H}_q = \mathcal{H}_{q,\text{irr}} \otimes \mathcal{M}_q$. Trivial representations are the representation equal to the identity for all group elements, $\Pi_{\text{trivial}}(g) = \mathbb{I}$ for all g. So, the multiplicity spaces have dimension equal to the number of copies of the irreducible representation held by $\mathcal{H}_q$. The irreducible representation space and the multiplicity space behave differently under the action of a map that destroys reference information.

As an example, when two spin-$\frac{1}{2}$ systems are combined under the tensor product, they can form a spin-1 system or a spin-0 system ($J = 0$ or $J = 1$, see equation 2.82). This can be expressed in terms of the Hilbert spaces:

$$\mathcal{H}_{\frac{1}{2}} \otimes \mathcal{H}_{\frac{1}{2}} = \mathcal{H}_1 \oplus \mathcal{H}_0, \tag{2.91}$$

and in terms of matrix representations

$$\begin{bmatrix} \boxed{\begin{matrix} & & \\ & J=1 & \\ & & \end{matrix}} & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 0 \end{matrix} & \boxed{J=0} \end{bmatrix}, \tag{2.92}$$

where along the diagonal there is a irreducible representation of $SU(2)$ on three dimensions (all possible transformations on a spin-1 system) and an irreducible representation of the spin-0 transformations which are just phases (*i.e.*, in $U(1)$). These both have multiplicity one.

Tensoring with another spin-$\frac{1}{2}$ system gives

$$\mathcal{H}_{\frac{1}{2}} \otimes \mathcal{H}_{\frac{1}{2}} \otimes \mathcal{H}_{\frac{1}{2}} = \mathcal{H}_{\frac{3}{2}} \oplus \mathcal{H}_{\frac{1}{2}} \oplus \mathcal{H}_{\frac{1}{2}}, \tag{2.93}$$

and

$$\begin{bmatrix} \boxed{\begin{matrix} & & \\ & J=\frac{3}{2} & \\ & & \end{matrix}} & \begin{matrix} 0 \\ \\ \end{matrix} & \begin{matrix} 0 \\ \\ \end{matrix} \\ \begin{matrix} 0 \\ \\ \end{matrix} & \boxed{J=\frac{1}{2}} & \begin{matrix} 0 \\ \\ \end{matrix} \\ \begin{matrix} 0 \\ \\ \end{matrix} & \begin{matrix} 0 \\ \\ \end{matrix} & \boxed{J=\frac{1}{2}} \end{bmatrix}. \tag{2.94}$$

Now the space is composed of an irreducible representation of $SU(2)$ on four dimensions (spin-$\frac{3}{2}$ transformations) and two copies of the irreducible representation of $SU(2)$ on two dimensions. The multiplicity of the $J=\frac{1}{2}$ space is therefore two. $\mathcal{H}_{\frac{1}{2}} = \mathcal{H}_{\frac{1}{2},\text{irr}} \otimes \mathbb{I}_2$.

The lack of a reference frame can be described as a *twirling* operation on a state ρ_{RF} over the elements of the group [BRS07] associated to the symmetry broken by the reference:

$$\rho_{\text{noRF}} = \mathcal{E}(\rho_{\text{RF}}) = \int_G T(g)\rho_{\text{RF}}T(g)^\dagger, \tag{2.95}$$

where $g \in G$ and $T(g)$ is an irreducible representation of G.

Lemma 2.1 (of Schur)**.** *[Sch05, BRS07] 1. Any operation, X, with the property that $\Pi(g)X\Pi(g)^\dagger = X$ for all $g \in G$ where $\Pi(g)$ is an irreducible representation of G, is a multiple of the identity. $X = c\mathbb{I}$.*

2. Any operation X, with the property that $\Pi_q(g)X\Pi_{q'}(g)^\dagger = X$ for all $g \in G$ where Π_q is an irreducible representation and $q \neq q'$ so that Π_q and $\Pi_{q'}$ are inequivalent representations of G, then $X = 0$.

This gives a restriction on how the map $\mathcal{E}$ can act. It must completely mix states in the spaces of the irreducible representations, but will leave states on the the multiplicity spaces intact. It can be expressed [BRS07] as:

$$\mathcal{E} = \sum_q (\mathcal{D}_{\mathcal{H}_{q,\text{irr}}} \otimes \mathbb{I}_{\mathcal{M}_q}) \circ P_q \tag{2.96}$$

where $\mathcal{D}(\rho) = \frac{\mathbb{I}}{d}$ if ρ is any state with dimension d and P_q is a projection onto the space $\mathcal{H}_q$. Then the operations that are invariant under the action of the group G are

$$A = \bigoplus_q \mathbb{I}_{\mathcal{H}_{q,\text{irr}}} \otimes B_{\mathcal{M}_q} \tag{2.97}$$

where B is any operator on the space $\mathcal{M}_q$.

Therefore, an operation which is invariant under the action of $SU(2)$ can only act non-trivially on the multiplicity spaces. In the example of the three spins tensored together, there is a space which is invariant under $SU(2)$. It is the multiplicity two space of the $\mathcal{H}_{\frac{1}{2}}$ space. Since it is two dimensional, it can store a logical qubit, even without a Cartesian frame ($SU(2)$ invariance). That is, imagine a three-qubit quantum state is prepared by Alice and sent to another Bob. If the Alice and Bob do not agree on a Cartesian frame, this is equivalent to there being a channel between them that applies a random rotation from $SU(2)$. Even so, Alice can still send one qubit of information to Bob. This will be referred to again in Chapter 4 Section 4.4.1 as an "energy conserving" subspace.

Also note that two spin-J systems tensored together have a Hilbert space that can be decomposed as

$$\mathcal{H}_J \otimes \mathcal{H}_J = \mathcal{H}_{2J} \oplus \mathcal{H}_{2J-1} \oplus ... \oplus \mathcal{H}_0, \tag{2.98}$$

where each $\mathcal{H}_j$ is an irreducible representation that appears with multiplicity one. Any operator on this Hilbert space that is invariant with respect to $SU(2)$ rotations, by Schur's Lemma must take the form $\rho = \bigoplus_{j=0}^{2J} c_j \mathbb{I}_j$, where the $c_j \in \mathbb{C}$ are

constants. This provides intuition for the proof of Theorem 3.1 in from Chapter 3, presented in Appendix A.

2.6 Reference frames

2.6.1 Introduction

Decoherence is the loss of quantum coherence, that is, the potential to observe quantum interference effects, and has a destructive effect on attempts to implement a quantum computation because it reduces the strength of quantum interference effects which are required to obtain the apparent speed-up over classical computation.

Decoherence can have many sources in an apparatus designed to exploit quantum effects. It can come as a result of technical imperfections in the devices that perform the operations on the qubits, or that perform the measurements, or from background noise which adds unwanted and unknown Hamiltonians to all or some qubits. In practice all of these effects will play a role, however, there are also fundamental sources of decoherence, which are the result of quantum theory itself and is thus unavoidable no matter how well-made the physical apparatus is or what implementation is selected.

One type of fundamental decoherence occurs when operations are completed by interacting an external system with the quantum system performing the computation. Any system used to implement such an operation, when treated quantumly may become entangled with the computation system. A *quantum reference frame* is a system that is interacted with the quantum system in order to provide a convention for measurements and operations. Normally, such reference systems are treated classically because they tend to be large and in a state which has thermalized. Together, these facts make such an approximation very good. However, if we believe that quantum mechanics is fundamental, then strictly speaking, this reference system is a quantum system and it is most accurate to treat it as such. The shortcomings of assuming it to be classical become apparent when conservation laws are considered and it seems that quantum mechanics leads to a paradox. However, as we shall come to later, the paradox is instantly resolved when the operating system is allowed to have quantum operators for its relevant degrees of freedom.

Treating the operating system classically prevents the consideration of any "back-action" of the computation system on the reference system. It is this back-action that causes the loss of coherence in the quantum system and for this reason, this source of decoherence is often overlooked. The classical model also allows each operation performed on the computation system to be modeled as a Hamiltonian that is applied to the system. Whenever we describe an operation in this way, we are implicitly assuming that there is a large classical, external system that is coupling to the quantum system to effectively apply this Hamiltonian to it and drive the new transitions in the system being treated quantumly. In this way, every Hamiltonian is an approximation, albeit, generally a very accurate one.

A qubit is a two-state system that displays quantum coherence. For a qubit to be useful for quantum information purposes it must be possible to perform a measurement that distinguishes these two states. Ideally, they should be orthogonal and thus perfectly distinguishable under the measurement associated to the operator of the degree of freedom that defines the two levels. This also requires access to a reference frame for that degree of freedom to perform the measurement, as is explained in the following.

In classical mechanics, a reference frame is a convention for describing the state of a system mathematically. For example, it can provide an origin for a coordinate system. In quantum mechanics, a classical reference frame is assumed when writing a state: this provides the basis or coordinate system in which the state is expressed. A case in point is that of a spin-$\frac{1}{2}$ system that is written in z basis. Without some definition of the z axis (often assumed to be provided by the classical measuring device) it is impossible to write a description of what the likelihoods of the outcomes corresponding to each "z projection" will be in a "z axis" measurement. We can relate one local reference frame to another reference frame by means of a transformation between the two basis conventions. For example two offset Cartesian frames can be transformed, one to the other via the appropriate $SU(2)$ rotation. This is a unitary basis change.

Suppose we have a *directional reference frame*, that is, a unit vector in three-space. Such a frame defines an axis in three-space. Let us call this axis the z axis, and assign values to projections of angular momentum m_J of spin states onto this axis. We can write the state of a spin-J system in terms of the basis defined by the system's possible values of m_J. However, without an additional notion of either an x or y axis, we cannot express a spin-J state as a superposition over different values of m_J. Our ignorance of this convention is described mathematically by integrating over all the possible directions the x (or y) axis might point in — that is, the whole

plane perpendicular to z:

$$\int_0^{2\pi} R_z(\phi)\rho_J R_z(\phi)^\dagger d\phi, \tag{2.99}$$

where $R_z(\phi)$ is a rotation about the z-axis through an angle ϕ. In effect we have a superselection rule that prevents us from creating coherent superpositions over different spin angular momentum z-direction values m_J. In addition, if we also lack a reference frame to define z, then it is not possible to create coherent superpositions of different values of total angular momentum, J. We say in this case that the angular momentum of this individual quantum spin is conserved.

Now suppose this axis is defined by a real physical object, rather than a conceptual ideal, against which we will make measurements in order that we can describe the relative state of other systems to this reference object. If the reference object is classical, then we can measure its state without altering it. However, if the reference object is a quantum system, for example a spin with a large value of J, then measurements relating its state to that of smaller spin systems will have some back-action on the reference state. Suppose we have both a classical and a quantum reference which are initially aligned. If we use the quantum state as a measurement reference repeatedly, eventually such measurements will cause a degradation of the quantum reference system as an indicator of the direction of the classical system.

2.7 Superselection Rules and Restrictions on Measurement

We can relate this to superselection rules imposed by conserved quantities. A degree of freedom is said to be conserved if its associated operator A commutes with the fixed Hamiltonian of the system. That is,

$$[A, H] = 0. \tag{2.100}$$

In nonrelativistic quantum mechanics, conservation laws appear to restrict measurements that can be performed on a quantum system and also the superpositions that can be created over particular variables. Initially it was suggested that charge conservation would impose a charge superselection rule [WWW52], implying that no coherent superpositions of states with different charge could be created. Likewise, component-wise angular momentum conservation counter-intuitively seems to

prevent a measurement of m_j for any particle. We shall see why this is not a paradox. Consider: if it is possible for the x-component of the total angular momentum to commute with a j_z measurement on some subsystem, then both can have well defined values at the same time and the measurement can be completed even under the momentum conservation restriction.

Let us isolate a particular pair of systems so that we can conserve momentum jointly within them. Note that by saying systems are isolated this means no other external system can couple to them. If J is the angular momentum operator of our reference system and j is the angular momentum of a spin system which we wish to perform measurements on, then let $\mathcal{J}$ be the total angular momentum of both systems together. For angular momentum to be conserved we have

$$[\mathcal{J}_x, H] = [\mathcal{J}_y, H] = [\mathcal{J}_z, H] = 0$$

and for the conservation of total momentum we have

$$[\mathcal{J}^2, H] = 0,$$

which in fact is a direct consequence of the first set of commutation relations. For example, let us conserve the total momentum of both systems (which can interact with each other, but are otherwise totally isolated) along the x-axis, $\mathcal{J}_x$:

$$\hat{\mathcal{J}}_x = \hat{J}_x + \hat{j}_x, \tag{2.101}$$

but this will not commute with j_z:

$$[\mathcal{J}_x, j_z] = [J_x + j_x, j_z] = -i\hbar j_y, \tag{2.102}$$

therefore we might conclude that this projection measurement cannot be done, because the total angular momentum should be conserved for all components of J (J_x, J_y, and J_z) in any closed physical system. However, as our system is completely isolated, we will only ever be able to relate the quantum spin's direction to the large reference spin's direction. The measurement that it is possible to perform is $\mathcal{J}^2$, which is a measurement of the total angular momentum of both systems together.

This *does* commute with the total angular momentum along the x-direction:

$$\begin{aligned}[\mathcal{J}_x, \mathcal{J}^2] &= [\mathcal{J}_x, (\mathcal{J}_x^2 + \mathcal{J}_y^2 + \mathcal{J}_z^2)] \\ &= i\hbar(\mathcal{J}_y\mathcal{J}_z + \mathcal{J}_z\mathcal{J}_y - \mathcal{J}_z\mathcal{J}_y - \mathcal{J}_y\mathcal{J}_z) \\ &= 0. \end{aligned} \tag{2.103}$$

If the two systems are completely isolated, it is not possible to speak about their relation to an ideal z-axis, but the large reference system J we will use to define for the smaller spin system a new, local z'-axis. The new total angular momentum $\mathcal{J}$ will indicate whether j points in the same direction as J, in which case $\mathcal{J}^2 > J^2$ or in the opposite direction ($\mathcal{J}^2 < J^2$). In fact it will give us a measurement of $m_{j(z')}$, the projection of the spin along the direction of the axis z', for the small spin system by virtue of the fact that $\mathcal{J} = \mathbf{J}+\mathbf{j}$. This measurement can be completed. If we wish to consider everything quantum mechanically, then we will always be encountering this sort of restriction. We cannot define a z-direction except by means of the orientation of a (possibly very large and approximately classical) quantum system.

When the reference system, which is the system employed to perform a operation, is quantized, then it is clear that the measurement being made is a *relative* measurement, which means firstly that the measurement is imperfect, since the limited size of the reference can only determine a measurement convention to a finite accuracy, and secondly that the act of measurement degrades the reference frame.

Chapter 3

The Degradation of a Quantum Direction Reference

Contents

3.1 Overview

In this chapter, a quantum spin is considered as a directional reference resource for performing measurements and operations on other quantum spin systems. These operations can be described by directionally covariant maps and a general form for covariant maps is found. This allows the analysis of the evolution of the moments of the angular momentum operator of the quantum reference system and from this a theorem governing the longevity of a quantum directional reference frame is discovered. Many examples then demonstrate how these results can be put to use to analyze measurements of and rotations on a variety of quantum systems. The work in this chapter has appeared in [BSLB08].

3.2 Introduction

In order to conserve angular momentum in a quantum system while performing operations or measurements on it, access to a quantum system that serves as a

direction reference is needed. A spin can be used in this way to indicate a direction in three dimensional space. This quantum reference allows other quantum systems to be rotated or measured projectively while the conservation law is upheld. However, as the reference spin is used in this capacity its usefulness for indicating a direction is reduced. That is, the state of the spin will become increasingly mixed and it contains less and less information about direction. In this sense the quantum direction reference is a resource that is consumed when it is used as a measuring apparatus or means of performing quantum operations on other systems.

The quantum operation describing the unitary rotation is then an operation *conditional* on the state of the quantum reference frame. As a result of quantizing the reference frame, there will be an inherent uncertainty in its direction and so the conditional operation will not be identical to a classical conditional operation. Further, as a result of the operation the state of the quantum reference will undergo an unknown disturbance.

Consider the case that we wish to implement a unitary rotation operation on a qubit. In this work, the qubit is assumed to be spin-1/2 system, but reference frame restrictions will also apply to other kinds of qubit, where the basis states are another type of degree of freedom that also obeys a conservation law. In the usual implementations the apparatus employed to perform such a transformation will use a system which can be described as a classical directional reference frame in order to define an axis for the rotation. However, there could be constraints on the system that make this model infeasible. For example, the apparatus performing the measurement may need to be miniaturized in order to fit on a chip. Alternatively, the device may need to be small in order to couple to a particular spin and not the surrounding spins in a solid state architecture. If this occurs, then it is not accurate to describe the system providing the reference direction classically, and instead it should be treated quantumly, using a Hilbert space of finite size. The quantum operation describing the unitary rotation is then an operation *conditional* on the state of the quantum reference frame. However, there will be inherent uncertainty in the direction of the quantum reference frame, so this conditional operation cannot be perfect. Also, the state of the quantum reference frame will experience back action from this operation, and experience decoherence as a result of tracing over entanglement formed during the interaction between the reference frame and system.

It should be noted that the constraints that lead to the requirement that the reference system be treated quantumly are ones that are expected to arise in a quantum computing architecture. For example, measuring devices for such a sys-

tem must couple strongly to specific registers (subsystems), and this implies a very small device so that coupling can be controlled and unwanted couplings to neighbouring systems prevented. Ideally, a measuring apparatus should be well-modeled classically, to reduce noise, however, that may not be compatible with the previous requirement, since a very small device is better modeled quantumly. Therefore, it is important to consider how a moderately sized system decoheres when being used in this way. This work is directly applicable to the magnetic resonance force microscopy (MRFM) proposal of [RBMC04, SGB$^+$95]. In that scheme a magnet on the scale of tens of nanometers in length is placed on tip of a nanomechanical resonator, which at very low temperatures may allow single-spin measurements in a solid state quantum computer. In this regime the assumption of classical behavior of the device leads to a poor approximation. Similar problems would occur in situations in which Hamiltonians must be applied to specific subsystems without disturbing neighboring ones.

This type of scenario was considered previously by Bartlett, Rudolph, Spekkens, and Turner [BRST06]. They considered the effect on the reference frame of its repeated use measuring the direction of spin-1/2 particles. For a measure of the quality of the reference frame, they defined the *longevity* of a reference frame as the number of uses that can be made of it before the fidelity of the measurement being made on a spin-1/2 system with the ideal, classical measurement falls below a chosen threshold. This allowed them to study how the usefulness of the quantum reference deteriorated with its use. They also made the assumption that all of the spin-1/2 systems started in the completely mixed state. This is equivalent to saying the systems were drawn from a reservoir composed of spin-1/2 systems in identical states and overall the reservoir had no prior correlation to the quantum reference system. By no prior correlation, what is meant is that no relative direction could be defined between the reservoir and the reference initially. Under these circumstances, the longevity of the reference scales with the square of its total angular momentum, $O(j)^2$.

Poulin and Yard [PY07] also considered this situation, except without the constraint that the spin-1/2 particles should be in a mixed state. Since all of the particles are assumed to be interchangable, it this gives the reservoir an overall polarization. In this case, for any non-zero polarization the longevity scales instead as $O(j)$.

3.3 Mathematical Description and Physical Intuition

This type of scenario can be described mathematically as follows. For the system to be used as a directional reference, let us use a spin-j system, which has dimension $d = 2j+1$. This system will begin in some initial state $\rho^{(0)}$, which will be chosen to indicate a classical direction given by a real three-vector $\hat{n}$. Therefore, to serve as a directional reference, $\rho^{(0)}$ should be correlated with $\hat{n}$. Following the ideas of [PY07], the initial state in this analysis is permitted to be arbitrary, except that it is only allowed to depend on the direction $\hat{n}$: $\rho_j^{(0)}(\hat{n}) = \int_0^{2\pi} R_j(\phi)\rho_j^{(0)}(\hat{n})R_j(\phi)\mathrm{d}\phi$ for the spin-j quantum directional reference frame, where the integration is over rotations about the axis $\hat{n}$. This restricts the initial density matrix to be diagonal in the particular basis that corresponds to $\hat{n}$. That is, $\rho_j^{(0)}(\hat{n})$ commutes with the angular momentum operator $J_{\hat{n}}$ in the $\hat{n}$-direction. The quantum reference frame transforms under rotations according to the spin-j representation of the group $SU(2)$, $R_j(\Omega)$, where $\Omega \in SU(2)$ denotes the various possible rotations, for example by indexing a rotation axis, a rotation angle, and an inconsequential global phase.

Let $\sigma^{\otimes n}$ be the state of a reservoir composed of n ordered subsystems on which sequential operations will be performed using the same quantum reference frame. A quantum operation χ is applied on the joint system of the reference and the first reservoir subsystem $\rho^{(0)} \otimes \sigma_1$. The map χ is required to be *rotationally invariant*, meaning that it is independent of any specific direction in space. Formally, a map χ is rotationally invariant if and only if $\chi\big(R(\Omega)(\cdot)R(\Omega)^\dagger\big) = R(\Omega)\chi(\cdot)R(\Omega)^\dagger$ for any rotation $\Omega \in SO(3)$ of the joint system (or more generally $SU(2)$ which is the double covering group of $SO(3)$, that is, there is a homomorphism with kernel 2 that takes $SU(2)$ to $SO(3)$, so their Lie algebras are isomorphic; see Chapter 2, Section 2.5.1). Here, R is the unitary representation of the rotation group $SO(3)$ on the joint system. This implies that all directions are treated equally by the map so that rotating the state and then performing the map is equivalent to performing the map on the state and then rotating its output. An example of a rotationally invariant operation on two qubits would be the application of a Hamiltonian of the form $H = XX + YY + ZZ$. This can be seen to be true by noting that an arbitrary unitary U can be expressed as $U = e^{i\alpha}\big(\cos(\beta/2)\mathbb{I} + i\sin(\beta/2)(n_xX + n_yY + n_zZ)\big)$ and that $U \otimes U$ commutes with H.

This restriction is made because if the map were not to be rotationally invariant, this would imply the use of a classical reference in the action of the map and thus

it would not be reasonable to consider the quantum reference as the only reference resource available. Subsequent to the operation, the reduced state of quantum reference frame is $\rho^{(1)} = \mathrm{Tr}_{\sigma_1}[\chi(\rho^{(0)} \otimes \sigma_1)]$, where Tr_{σ_i} denotes the partial trace over the i^{th} subsystem of the reservoir. The first subsystem of the reservoir is then discarded, and the same operation is performed on the second subsystem, but using the updated quantum reference frame $\rho^{(1)}$. These steps are repeated with the following subsystems of the reservoir, always using the updated quantum reference frame. The state of the quantum reference frame after the i^{th} step is $\rho^{(i+1)} = \mathrm{Tr}_{\sigma_i}[\chi(\rho^{(i)} \otimes \sigma_i)]$. By discarding the previous subsystems of the reservoir at every step, it follows that the reservoir cannot be used to increase our knowledge about the quantum reference frame.

As in [BRST06, PY07], a quantum directional reference frame could be used to measure whether a series of spin-$\frac{1}{2}$ particles are parallel or anti-parallel to some classical arrow. (See Figure 3.1.) However, there are a myriad of other possible cases, including measuring the angular momentum of a series of spin-j' systems with $j' > \frac{1}{2}$, or using the quantum reference spin to indicate an axis about which a unitary rotation operation is performed.

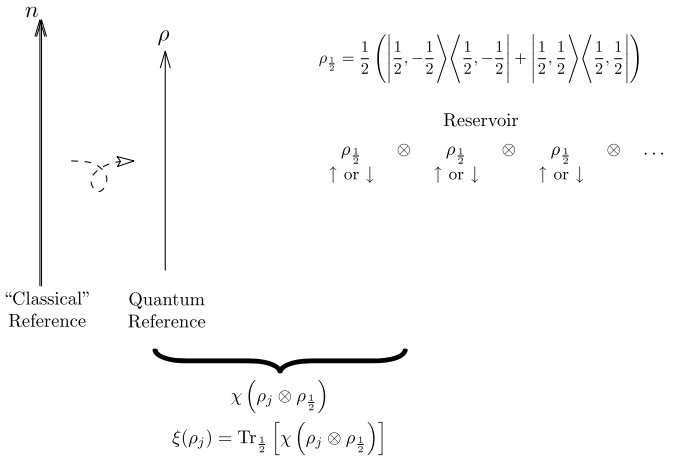


Figure 3.1: The scenario studied by Bartlett *et al.* [BRST06]. The quantum reference spin is used to measure the direction of a series of spin-$\frac{1}{2}$ particles in the completely mixed state by means of a projection onto the $J - \frac{1}{2}$ or $J + \frac{1}{2}$ subspaces.

3.3.1 Aside: Intuition from the Stern-Gerlach Arrangement

A physical scenario that this can be related to for physical intuition is a modified version of the Stern-Gerlach experiment. A diagram of the basic experimental setup is shown in Figure 3.2. A beam of silver atoms (which are spin-$\frac{1}{2}$ systems for the electron in the lowest energy S state) is sent through a magnet polarized in the z-direction. In other words, the South pole of a horseshoe cross-section magnet is directly above the North pole, such that there is a gap between them through which the beam passes. The shape of the magnet is designed to make the field inhomogeneous: the field lines have a higher density at the North pole than at the South pole. The source of the silver atoms is an oven, which heated the atoms to a particular temperature, imparting them with a corresponding kinetic energy and velocity. After emerging from a small opening in the oven wall, the atoms pass through collimating slits, in order to ensure that the beam is tightly focused, but is also incoherent in the sense that the silver atoms' spin directions are all randomly oriented over the state space. A fluorescent screen placed behind the magnet to detects where the beam strikes the screen. The magnet directs different spin orientations in different directions according to the equation

$$F_z = -\frac{e}{m_e c} s_z \frac{\partial B_z}{\partial z} \tag{3.1}$$

from electromagnetic theory, where $-e$ is the charge of an electron, m_e is the mass of the electron, c is the speed of light, s_z is the z-component of the atom's spin vector $\mathbf{s}$ (think of this as indicating the direction that the atom's own magnetic field is oriented in), and $\frac{\partial B_z}{\partial z}$ is the gradient of the inhomogeneous magnetic field, $\mathbf{B}$, along the z direction. This alludes to the fact that the magnetic moment of a silver atom is mostly due to the spin of the single least-strongly-bound electron. The Hamiltonian experienced by the atom will be

$$H = \frac{p^2}{2m} + \mu \boldsymbol{\sigma} \cdot \boldsymbol{B} \tag{3.2}$$

where p is the linear momentum of the silver atom, $\boldsymbol{\mu} = \mu \boldsymbol{\sigma}$ is the electron's magnetic moment, and m is its mass.

The magnet has a sufficiently strong field to act as a classical measuring device which couples strongly to the quantum systems, that is, the silver atoms. After leaving the magnet, only two trajectories are followed by the electrons. These

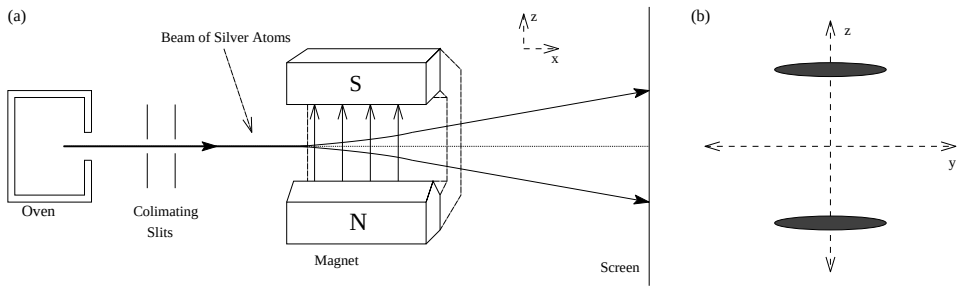


Figure 3.2: (a) The apparatus for a Stern-Gerlach experiment and (b) the resulting pattern on the screen.

correspond to the two discrete states of spin, $s = +\frac{1}{2}$ and $s = -\frac{1}{2}$. So the coupling to the magnetic field projects the spins into being either aligned with the magnetic field (and the z-axis) or anti-aligned.

Now imagine modifying this scheme. For convenience, consider an approximation that the magnetic field that the silver atom passes through is a homogeneous magnetic field $\mathbf{B} = B_z\hat{z}$ (even though for such a field the Maxwell equation $\nabla\cdot\mathbf{B} = 0$ is violated, in this case the projection is of a von Neumann type [SLB87]) that leaves the spin directed either along the field or against the field, in this case a simple projection onto a state of well-defined z axis angular momentum:

$$\Pi_\uparrow = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \Pi_\downarrow = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.3}$$

As the source of such a field, replace the magnet with a large quantum system that has a spin j (for example, a nucleus with a very large spin or a collection of spins with total angular momentum j) and therefore an intrinsic magnetic moment $\mu_j = g\frac{q}{2m}j$ where g is a constant, q is the charge of the system, and m is its mass. The coupling of this field with the magnetic moment of the spin-$\frac{1}{2}$ particle, the silver atom, will project the silver atom's spin to be either aligned with the moment of the spin-j system or anti-aligned, by projecting the joint system of both spins together onto a state of well-defined total angular momentum:

$$\Pi_{j+\frac{1}{2}}, \qquad \Pi_{j-\frac{1}{2}}. \tag{3.4}$$

However, since the magnetic field is now caused by a quantum system instead of a classical bar magnet, the state of the quantum system, ρ_j will be altered by this coupling to the state of the spin-$\frac{1}{2}$ system, σ. It is this process of altering with

repeated interactions that we aim to describe.

Each time the system couples to a silver atom's spin, the system is given a "kick" in an unknown direction. If the apparatus that performs the measurement is considered to be in a closed box, that only reveals the outcomes "up" and "down", then the results of the measurement will become less and less useful with repeated interactions with each silver atom, since eventually the direction of the quantum system's magnetic moment will be uncorrelated with the z-axis. By describing the degradation of the quantum system's state as an indicator of the direction z, we also quantify its usefulness for subsequent measurements or operations.

3.3.2 Organization

In the remainder of this chapter the results of [BRST06, PY07] are generalized in three different ways. First, performing operations on a reservoir of *qudits* (quantum systems with d-dimensional Hilbert spaces) is considered instead of restricting to the case of qubits ($d = 2$) as in [BRST06, PY07]. Second, all possible rotationally-invariant quantum operations on the reference frame and the system are considered, while the analyses in [BRST06, PY07] were restricted to very particular kind of interaction (specifically, interactions describing measurements). Note that this joint quantum operation could be a measurement, a conditional unitary, or any other completely-positive map, provided that it is rotationally invariant. Third, the concept of the quality function is generalized, and general conditions are placed on its form and evolution. In addition to using the *longevity* of a reference as a measure of usefulness of the reference, in this work a broader notion of a measure of the quality of the quantum reference frame is defined and a consideration is made of how it decreases with the number of times the reference frame is used. Defining the quality of a quantum reference can be quite arbitrary and depends on how it is intended to be used. In [BRST06, PY07], only a particular kind of quality function was considered, one which was based on the average probability of a correct measurement given a known spin direction, while in this work any possible quality function is considered. Here, the term *quality function* is used for any function F that is meant to quantify the ability of the reference frame to perform a particular task. General conditions are set on such functions, and how such functions evolve with the number of uses of the quantum reference frame is analyzed.

This chapter is organized in the following way. In Section 3.4, a set of parameters, called *moments*, are introduced for describing a quantum reference frame, and it is demonstrated that any quality function must depend only on these moments.

Then the central results are presented, Theorem 3.1, which provides a general form of all rotationally-invariant maps and the evolution of the moments (Theorem 3.2) is found to be described by a pair of recursive relations in the same section. In Section 3.5, the longevity of a quantum reference frame for a general quality function under the repetitive application of a rotationally-invariant map is explored, by analyzing the dependance of the evolution of its moments. Section 3.6 is dedicated to a pair of examples intended to illustrate the power of the techniques developed in the earlier sections of the paper, in particular the specific cases of measuring spin-1 particles relative to a quantum reference frame, as well as the use of a reference frame to perform Pauli operations on qubits.

3.4 Moments and Fidelity Functions

Now consideration is given to how the state of the quantum reference frame is updated through its use in repeated rotationally-invariant operations χ which act on both the frame and the reservoir. It is assumed as in [BRST06] that the reservoir is initially in a state that is invariant under rotations. (The completely mixed state is one example of such a state, but there are other non-trivial states, for example each σ could be a pair of qubits in the antisymmetric state.) The assumptions that the initial state of the reservoir and the joint quantum operation are spatially invariant imply that the map ξ used to update the state of the quantum directional reference frame, $\rho_j^{(i+1)} = \xi(\rho_j^{(i)}) = \mathrm{Tr}_{r_i}[\chi(\rho_j^{(i)} \otimes \sigma_i)]$, is also rotationally invariant:

$$R^{(j)}(\Omega)\xi(\rho_j)R^{(j)}(\Omega)^{-1} = \xi(R^{(j)}(\theta)\rho_j R^{(j)}(\Omega)^{-1}). \tag{3.5}$$

The restricted map is denoted ξ on the quantum directional reference and is here called the *disturbance* map. It describes the back-action on the reference which occurs as a result of the interaction.

Let $R_{\hat{n}}(\theta)$ be the rotation by an angle θ around the vector $\hat{n}$. Suppose that a rotationally-invariant map ξ is applied to density matrix ρ_j that is diagonal in a basis given by the eigenvectors of $J_{\hat{n}}$. Recalling that the reference state is assumed to define only one direction in space (along the vector $\hat{n}$) and is thus invariant with respect to rotations about the $\hat{n}$ axis such that

$$R_{\hat{n}}(\theta)\xi(\rho_j)R_{\hat{n}}^{-1}(\theta) = \xi(R_{\hat{n}}(\theta)\rho_j R_{\hat{n}}^{-1}(\theta)) = \xi(\rho_j), \tag{3.6}$$

then $\xi(\rho_j)$ is also diagonal in the basis given by the eigenvectors of $J_{\hat{n}}$. As a result,

the evolution of the quantum reference frame under the repeated application of the rotationally-invariant map ξ can be described by $2j+1$ equations, one for each of the eigenvalues of ρ_j. A different set of parameters that are equivalent to the eigenvalues is the set of *moments* given by

$$\{\mathrm{Tr}[\rho_j J_{\hat{n}}^{\ell}] \mid 1 \leqslant \ell \leqslant 2j\}, \tag{3.7}$$

where $J_{\hat{n}}^{\ell}$ is the ℓth power of the operator $J_{\hat{n}}$. Note that only $2j$ moments are necessary because the sum of the eigenvalues must be one. The use of moments instead of eigenvalues is in order to simplify the analysis of the evolution of the quantum reference frame.

The main motivation for studying the moments of the quantum reference frame is that any quality function F will depend only on these moments. Consequently, the behavior of the different moments will determine the behavior of the different quality functions. To see why F depends only on the moments given by equation 3.7, first note that any reasonable form of F depends only on the state of the quantum reference, and that it must respect the relation $F(\rho_j) = F(R^{(j)}(\Omega)\rho_j R^{(j)}(\Omega)^{-1})$ for all rotations $\Omega \in SU(2)$ and the state ρ_j is diagonal in the basis composed of the eigenvectors of $J_{\hat{n}}$. In other words, the quality measure should not be biased such that it favors a quantum reference frame that is pointed in any particular direction relative to some external frame. All directions must be equally valid. Therefore, F does not depend on the direction of $\hat{n}$, but only on the eigenvalues or the moments of ρ_j. Note that the set of moments with respect to the direction $\hat{n}$ can be written as a function of the moments with respect to any other direction — this is simply a change of basis.

Since the eigenvalues of the reference system convey the same information as the moments, the quality function could also be written in terms of the eigenvalues. However, in this problem the evolution of the quality function is more readily expressed in terms of the moments. The idea behind this work is to restrict the number of parameters needed to describe the behaviour of the reference. Working in terms of eigenvalues, it would depend on the initial state of the reference. Using moments as the parameters in the equations describing the evolution of the reference frame allows the restriction of dependencies based only on the map in question, ξ. In other words, the calculations can be simplified in the same way for any rotationally symmetric initial reference state; what matters is the process for which the reference is being used. For this reason, attention is directed to how the moments update as the map ξ is applied repeatedly to the reference state, ρ_j.

Recall that $\rho_j^{(k)}$ is the state of the quantum reference frame after the k^{th} joint operation with a subsystem of the reservoir, and $\rho_j^{(0)}$ as the initial state of the quantum reference frame. For a map ξ, that is, for a given state $\sigma^{\otimes n}$ of the reservoir and particular joint quantum operation χ, the evolution of a frame under ξ must be described by $2j+1$ explicit linear functions that take the moments of $\rho_j^{(k-1)}$ to the moments of the updated state $\rho_j^{(k)}$. These results lead to the general recursion relation,

$$\mathrm{Tr}[\rho_j^{(k)} J_{\hat{n}}^{\ell}] = \sum_{i=0}^{2j} A_i^{(\ell)} \mathrm{Tr}[\rho_j^{(k-1)} J_{\hat{n}}^{i}], \tag{3.8}$$

where $A_i^{(\ell)}$ are real coefficients.

It is possible to express and classify different rotationally-invariant maps in terms of the generators of the Lie algebra, J_x, J_y, and J_z and such a form is now presented. For this purpose, an expression which is invariant with respect to 3−space rotations is required. Such rotations are described by matrices belonging to the matrix representation of the group $SO(3)$. This group is isomorphic to the group $SU(2)$ up to a complex multiplicative phase factor. Employing the definition of the exponential map that takes the matrix Lie algebra, $\mathfrak{g}$, to its associated matrix Lie group, G as $e^{iM} : M \in \mathfrak{g}$, the generators of the group $SU(2)$ are the angular momentum operators J_x, J_y, and J_z. In this work the focus is on the case for $SU(2)$, however study has been made of how to obtain noiseless subsystems for quantum channels constructed from generators of a Lie algebra, first in [LCW98], and proprieties of Lie algebra channels were analyzed in [Rit05]. Therefore, let J_k for $k = x, y, z$ be the angular momentum operators for some arbitrary Cartesian frame. On a spin-j system, define the map

$$\zeta(\rho_j) = \frac{1}{j(j+1)} \sum_{k\in\{x,y,z\}} J_k \rho_j J_k, \tag{3.9}$$

for ρ_j a density matrix of the spin-j system. Let $\zeta^{\circ n}$ denote the composition of ζ with itself n times for $n > 0$, for example $\zeta^{\circ 2}(\rho) = \zeta(\zeta(\rho))$, and define $\zeta^{\circ 0}(\rho_j) = \rho_j$.

Theorem 3.1. *Any map ξ which is invariant with respect to the spin-j irreducible representation of $SU(2)$ has the form*

$$\xi(\rho_j) = \sum_{n=0}^{2j} q_n \zeta^{\circ n}(\rho_j), \tag{3.10}$$

where $\{q_n\}$ are real coefficients.

This series expansion is a valid expression for any rotationally invariant map on a reference state. The parameters q_n can be found by considering the action of the process described by the map on a particular pure state, *e.g.* $|j,j\rangle\langle j,j|$. This decomposition will allow the number of variables that are needed to describe the evolution of the reference under the map ξ to be limited by using Theorem 3.2

The proof is provided in Appendix A.

Theorem 3.1 allows an immediate simplification by restricting the number of coefficients $A_i^{(\ell)}$ of equation (3.8) that are required to calculate the evolution of the reference state. The following theorem limits the number of coefficients $(A_i^{(\ell)})$ which determine the evolution of the $J_{\hat{n}}$-moments of a spin-j under the action of a general invariant channel, and can be very useful to study the behavior of moments with low degree:

Theorem 3.2. *If ℓ is even, then*

$$\mathrm{Tr}[\rho_j^{(k)} J_{\hat{n}}^{\ell}] = \sum_{i=0}^{\ell/2} A_{2i}^{(\ell)} \mathrm{Tr}[\rho_j^{(k-1)} J_{\hat{n}}^{2i}] \tag{3.11}$$

and if ℓ is odd, then

$$\mathrm{Tr}[\rho_j^{(k)} J_{\hat{n}}^{\ell}] = \sum_{i=1}^{(\ell+1)/2} A_{2i-1}^{(\ell)} \mathrm{Tr}[\rho_j^{(k-1)} J_{\hat{n}}^{2i-1}]. \tag{3.12}$$

There are never more than $2j+1$ moments required to describe any evolution of a spin-j ($2j+1$ dimensional) reference state. This theorem shows that low degree moments after k applications of the map do not depend on moments of higher degree from the $(k-1)$th application of the map. Also, they only depend on half of the lower degree moments - those with the same parity. An explicit example is that of a third order moment, $\ell = 3$. In that case, one would have $\mathrm{Tr}[\rho_j^{(k)} J_{\hat{n}}^3] = A_1^{(3)} \mathrm{Tr}[\rho_j^{(k-1)} J_{\hat{n}}] + A_3^{(3)} \mathrm{Tr}[\rho_j^{(k-1)} J_{\hat{n}}^3]$. That is, the third degree moment only depends on the previous first and third degree moments, and not all $2j+1$ moments. This can simplify the analysis considerably, as will be shown in Section 3.6.

Proof. For notational convenience, define the z-axis so that it is parallel to the vector $\hat{n}$ which describes the classical directional reference frame. Consider any rotationally-invariant map ξ. By Theorem 3.1, it is possible to write $\xi(\rho_j) =$

$\sum_{k=0}^{2j} q_k \zeta^{\circ k}(\rho_j)$ where the coefficients q_k are real numbers and ζ is given by equation (3.9). Therefore,

$$\mathrm{Tr}[\xi(\rho_j)J_z^\ell] = \mathrm{Tr}[\rho_j \xi(J_z^\ell)] = \sum_{n=0}^{2j} q_n \mathrm{Tr}[\rho_j \zeta^{\circ n}(J_z^\ell)]. \tag{3.13}$$

To prove this theorem, it is sufficient to show that $\zeta^{\circ n}(J_z^\ell)$ is a polynomial in J_z of degree ℓ, and where all the powers have the same parity as ℓ. Define $\lambda = j(j+1)$ and define the function

$$G(\ell) \equiv \frac{i}{\lambda}(J_x J_z^{\ell-1} J_y - J_y J_z^{\ell-1} J_x). \tag{3.14}$$

Using the commutation relations $[J_a, J_b] = i\epsilon_{a,b,c}J_c$, we can show that

$$\zeta(J_z^\ell) = \zeta(J_z^{\ell-1})J_z + G(\ell), \tag{3.15}$$

and

$$G(\ell) = G(\ell-1)J_z + \zeta(J_z^{\ell-2}) - \frac{J_z^\ell}{\lambda}. \tag{3.16}$$

By induction, it is easy to prove using those relations that $\zeta(J_z^\ell)$ is a polynomial in J_z of degree ℓ where all the powers have the same parity as ℓ.

First assume that $\zeta(J_z^\ell)$ for even ℓ can be expressed as $\zeta(J_z^\ell) = \sum_s^{\ell/2} b_s J_z^{2s}$ and $\zeta(J_z^{\ell-1}) = \sum_s^{\ell/2} b'_s J_z^{2s-1}$, and that $G(\ell) = \sum_t^{\ell/2} g_t J_z^{2t}$ and $G(\ell-1) = \sum_t^{\ell/2} g_t J_z^{2t-1}$, which is to say that they are polynomials of degree ℓ with terms in only even or odd powers of J_z. From equations 3.15 and 3.16 it is clear that

$$\zeta(J_z^{\ell+1}) = \zeta(J_z^\ell)J_z + G(\ell)J_z + \zeta(J_z^{\ell-1}) - \frac{J_z^{\ell+1}}{\lambda}, \tag{3.17}$$

so, all the terms on the right hand side of the equation have odd powers of J_z:

$$\zeta(J_z^{\ell+1}) = \sum_s^{\ell/2} b_s J_z^{2s+1} + \sum_t^{\ell/2} g_t J_z^{2t+1} + \sum_s^{\ell/2} b'_s J_z^{2s-1} - \frac{J_z^{\ell+1}}{\lambda}. \tag{3.18}$$

Also,

$$\zeta(J_z^{\ell+2}) = \zeta(J_z^{\ell+1})J_z + G(\ell+1)J_z + \zeta(J_z^\ell) - \frac{J_z^{\ell+2}}{\lambda}, \tag{3.19}$$

so all the terms on the right hand side of the equation have even powers of J_z:

$$\zeta(J_z^{\ell+2}) = \sum_s^{\ell/2} b_s J_z^{2s+2} + \sum_t^{\ell/2} g_t J_z^{2t+2} + \sum_s^{\ell/2} b'_s J_z^{2s} - \frac{J_z^{\ell+2}}{\lambda}. \tag{3.20}$$

Therefore, this oscillation between even and odd powers of J_z in the polynomial will occur if for $\ell = 1$ and $\ell = 2$ the map and function G have only terms in powers of J_z which are odd or even respectively. This is indeed the case. From equation 3.14:

$$G(1) = \frac{i}{\lambda}[J_x J_y] = -\frac{J_z}{\lambda}\,, \tag{3.21}$$

which has only an odd power, so, using equation 3.15 and noting that $\zeta(0) = 1$:

$$\zeta(J_z^1) = \left(1 - \frac{1}{\lambda}\right) J_z\,. \tag{3.22}$$

Then, from equation 3.16

$$G(2) = 1 - \frac{2}{\lambda} J_z^2\,, \tag{3.23}$$

and finally:

$$\zeta(J_z^2) = \left(1 - \frac{3}{\lambda}\right) J_z^2 + 1\,, \tag{3.24}$$

both of which have only even powers. Since it is true for these first cases, it is true for all integer values of ℓ.

Now if $\zeta^{\circ n}(J_z^\ell) = \sum_{s=0}^{\ell/2} b_s J_z^{2s}$ for even ℓ, then $\zeta^{\circ n+1}(J_z^\ell) = \zeta(\sum_s^{\ell/2} b'_s J_z^{2s}) = \sum_s b_s \zeta(J_z^s)$ since:

$$\zeta^{\circ n+1}(J_z^\ell) = \sum_{s=0}^{\ell/2} b_s \zeta(J_z^{2s}) = \sum_{s=0}^{\ell/2} b_s \sum_t^s c_t J_z^{2t}\,. \tag{3.25}$$

Therefore additional applications of the map ζ do not change the parity of the powers of J_z. This same is true if ℓ is odd. Since for $n = 1$ the series has only even or odd powers of J_z depending on ℓ, then by induction $\zeta^{\circ n}(J_z^\ell)$ is a polynomial in J_z of degree ℓ where all the powers have same the parity as ℓ. □

3.5 Longevity of a Quantum Reference Frame

There are situations in which it is useful to be able to calculate how rapidly a quantum reference state is decaying. For example, suppose there is a microscopic

device for performing an operation or measurement on a quantum spin using a quantum reference frame. After many uses, the quantum reference frame will need to be realigned with the classical arrow $\hat{n}$. However, it is desirable to make as many uses of it as possible before performing this reinitialization, without allowing the accuracy to fall below some preselected threshold. To define this accuracy some quality function will be chosen that meets the particular needs of the protocol. It has already been shown that any rotationally-covariant quality function F depends only on the moments of the $J_{\hat{n}}$ operator.

It is of interest to determine the scaling of how many times a quantum reference frame can be used before the value of its quality function falls below a certain threshold with respect to Hilbert space dimension, d, or equivalently with the total angular momentum of the reference state ρ_j ($d = 2j + 1$), as was done in [BRST06, PY07]. This property is the *longevity* of a quantum reference frame. There are three important features of this analysis to highlight. First, whereas the work of [BRST06, PY07] restricts attention to one particular quality function, here the behavior of arbitrary functions satisfying the invariance relation described in Sec. 3.4 is investigated. Because any such function is considered, the longevity of the reference frame can be arbitrary. However, some general statements can be proven about the decay of the moments, and relationships amongst them, and these results can then be used to infer the behavior of any particular quality function. Second, rather than considering only a particular state of the reference frame (in [BRST06, PY07], they considered the state $\rho_j^{(0)} = |j,j\rangle_{\hat{n}} \langle j,j|$ because it was optimal for the task at hand), in this work an arbitrary state is considered. As such, the initial state $\rho_j^{(0)}$ of the quantum reference frame, as specified by its moments, can depend on j in a quite arbitrary way. Finally, arbitrary rotationally-invariant joint quantum operations χ are considered. In [BRST06, PY07], the joint operation was chosen to describe a particular measurement that was optimal for measuring the direction of spin-1/2 particles. Recall that in [BRST06], it was shown that the number of times a reference frame can be used before its fidelity (which in their case depends only of the first moment) falls below a certain threshold value scales as $O(j^2)$.

In the following, a general theorem is proven about the longevity of quantum reference frames, using Theorem 3.1 as a starting point. Recall Theorem 3.1 states that any rotationally-invariant disturbance map and concluding that in a very wide range of cases the longevity goes as $O(j^2)$. However, if the range and values of the coefficients q_n can depend on j arbitrarily, it is impossible to make any general statements about the longevity. On the other hand, making some well-motivated

assumptions allows for this general statement about the longevity.

The first assumption is that the number of non-zero coefficients q_n is bounded independently of j. In the general case, the number of parameters q_n describing a map can increase with j. The form of the map allows $2j+1$ parameters to determine the evolution, and in the case of the completely depolarizing map $\chi(\rho_j \otimes \sigma) = \frac{1}{2j+1}\mathbb{I}_{2j+1} \otimes \frac{1}{2}\mathbb{I}_{1/2}$ for example, all $q_n \neq 0$ for $0 \leq n \leq 2j$. However, consider a map χ on the systems $\rho_j \otimes \sigma$ for $\dim(\sigma) < 2j+1$ that conserves angular momentum in the $\hat{n}$ direction. Following an argument in [PY07], this implies that $\chi(J_{\hat{n}}) = J_{\hat{n}}$ where $J_{\hat{n}}$ is the *total* angular momentum operator in the $\hat{n}$ direction. The change of angular momentum of the quantum reference frame caused by such a disturbance map ξ cannot be higher than the change in angular momentum of the subsystem σ, which is in turn bounded by the subsystem's dimension, $\dim(\sigma)$. Given that the map $\zeta(\cdot) = \frac{1}{\lambda}\sum_{k=x,y,z} J_k(\cdot)J_k$ cannot lower or increase the angular momentum in any direction by more then one unit, the coefficients of the map ξ therefore satisfy $q_n = 0$ for all $n > \dim(\sigma)$. Thus, if a rotationally-invariant map conserves angular momentum, the coefficients $q_n = 0$ for $n > \dim(\sigma)$.

The second assumption restricts the form of the dependence of the coefficients on j. If the parameters q_n can depend on j arbitrarily, then we would not be able to conclude anything about the longevity of the quantum reference frame. To understand this, consider the general expression for the map on the reference state

$$\xi(\rho_j) = \sum_{n=0}^{2j} q_n \zeta^{\circ n}(\rho_j)\,. \tag{3.26}$$

If the q_n depend on j in a way that the q_n will increase as j increases this will determine directly the rate of the decay of the moments as a function of j, rather than this decay being solely set by the factors that determine the evolution of the moments $A_i^{(\ell)}$. If the dependence of q_n on j can be arbitrary, then the rate of the decay of the moments can also be an arbitrary function of j. For the following theorem, it is assumed that each coefficient q_n converges to a constant when $j \to \infty$. An assumption is also needed about the rate of convergence: suppose that each q_n can be written as a quotient of two polynomials in j, such that the degree of the denominator is at least the degree of the numerator (i.e., $q_n \leqslant O(1)$). This assumption ensures the appropriate behaviour of the measurement or rotation interaction in the classical limit. As an example of what can go wrong, if some q_n is a linear function of j then as the reference increased in "size", j, the decay with respect to measuring or operating on small systems would be more pronounced, rather than

less.

The third and final assumption concerns the initial state of the reference frame. It is assumed that for $\rho_j^{(0)}$, $\mathrm{Tr}[\rho_j^{(0)} J_{\hat{n}}^\ell] = \Omega(j^\ell)$. A sufficient condition for this to be true is that there exists a $\beta > 0$ independent of j such that ${}_{\hat{n}}\langle j,j|\, \rho_j^{(0)}\, |j,j\rangle_{\hat{n}} > \beta$. This ensures that the quantum reference state is a reasonable indicator of the direction $\hat{n}$. For example, if it begins in the direction opposite to the desired axis, it will still be rotationally invariant about that axis, but all the wrong conclusions will be drawn from measurements using it.

Theorem 3.3 (Longevity)**.** *Consider a quantum reference frame, realised as a spin-j system with initial state $\rho_j^{(0)}$, which is used for performing a rotationally-invariant joint operation χ. If this operation induces a disturbance map $\xi = \sum_{n=0}^{2j} q_n(j)\zeta^{\circ n}$ that satisfies the following assumptions:*

1. *there exists some n_{max} such that $q_n = 0$ for all $n \geqslant n_{max}$,*
2. *$q_n \leqslant O(1)$,*
3. *$Tr[\rho_j^{(0)} J_{\hat{n}}^\ell] = O(j^\ell)$,*

then the number of times, t_c, that such a quantum reference frame can be used before its ℓ^{th} moment falls below a certain threshold value scales as j^2. That is, $t_c = O(j^2)$.

Proof. The J_k operators are Hermitian and so, using the cyclic property of the trace, the map ζ has the property that

$$\mathrm{Tr}[\zeta(\rho) J_{\hat{n}}^\ell] = \mathrm{Tr}[\rho\, \zeta(J_{\hat{n}}^\ell)]. \tag{3.27}$$

Therefore the moments of angular momentum in the $\hat{n}$ direction after the map ξ_j has been applied to the reference state $\rho_j^{(k)}$ can be expressed as

$$\mathrm{Tr}[\rho_j^{(k+1)} J_{\hat{n}}^\ell] = \sum_{n=0}^{2j} q_n \mathrm{Tr}[\rho_j^{(k)}\ \zeta^{\circ n}(J_{\hat{n}}^\ell)] \tag{3.28}$$

Using the commutation relations, $\zeta^{\circ n}(J_{\hat{n}}^\ell)$ can be expanded as a polynomial in $J_{\hat{n}}$ of degree ℓ, as shown in the proof of Theorem 3.2. Again, let $\lambda = j(j+1)$. Using a proof by induction on equation (3.15) and observing that $\sum_{k\in\{x,y,z\}} J_k J_{\hat{n}} J_k = (\lambda - 1) J_{\hat{n}}$, it is easy to show that the coefficient of the leading term will be $A_\ell^{(\ell)} =$

$\sum_{n=0}^{n_{max}}[q_n(1-O(\frac{1}{\lambda}))] = 1-O\left(\frac{1}{\lambda}\right)$ where we used the normalization condition $\sum q_n = 1$, the assumption that each q_n is $O(1)$ and the fact that $q_n = 0$ for $n > 2j$. The coefficients of the non-vanishing lower terms will be $A_i^{(\ell)} = O(1)$ for $i < \ell$.

Assume $\zeta(J_z^\ell) = (1 - O(\frac{1}{\lambda}))J_z^\ell + \sum_i^{\ell-1} b_i J_z^i$ and $\zeta(J_z^{\ell-1}) = (1 - O(\frac{1}{\lambda}))J_z^{\ell-1} + \sum_i^{\ell-2} b_i J_z^i$ and $G(\ell) = O\left(\frac{J_z^\ell}{\lambda}\right) + \sum_t g_t J_z^{t-1}$. Then:

$$\zeta(J_z^{\ell+1}) = (1 - O(\frac{1}{\lambda}))J_z^{\ell+1} + \sum_i^{\ell} b_i J_z^{i+1} + O\left(\frac{J_z^{\ell+1}}{\lambda}\right) + \sum_t^{\ell} g_t J_z^t + \sum_i^{\ell} b_i' J_z^{i-1} - \frac{J_z^{\ell+1}}{\lambda} \tag{3.29}$$

which implies that

$$\zeta(J_z^{\ell+1}) = (1 - O(\frac{1}{\lambda}))J_z^{\ell+1} + \sum_i^{\ell} c_i J_z^{i+1}\,. \tag{3.30}$$

Also note that if $G(\ell) = O\left(\frac{J_z^\ell}{\lambda}\right) + \sum_t g_t J_z^{t-1}$ then $G(\ell+1) = O\left(\frac{J_z^{\ell+1}}{\lambda}\right) + \sum_t^\ell g_t J_z^t$. So if these forms of the expressions hold for $G(1)$, $\zeta(J_z^1)$, and $\zeta(J_z^2)$, then they hold for all ℓ. But this has already been found to be the case in the proof of Theorem 3.2. Furthermore, this holds for all n since the leading term of $\zeta^{\circ n}(J_z^\ell)$ is of the form $(1 - O(\frac{1}{\lambda}))^n J_z^{\ell+1}$, then the leading term of $\zeta^{\circ n+1}(J_z^\ell)$ is of the form $(1 - O(\frac{1}{\lambda}))^{n+1} J_z^{\ell+1} = (1 - O(\frac{1}{\lambda}))J_z^{\ell+1}$.

So, $A_\ell^{(\ell)} = \sum_{n=0}^{n_{max}}[q_n(1 - O(\frac{1}{\lambda}))] = 1 - O\left(\frac{1}{\lambda}\right)$, since q_n is assumed to be of order less that on equal to a constant with respect to j.

This reasoning about the constants $A_k^{(i)}$ is sufficient to characterize the rate of change of the moments with repeated application of the map. To demonstrate this, let $\ell = 1$. Now the minimum value of t is found such that $\mathrm{Tr}[\rho_j^{(t)} J_{\hat{n}}] < c$, for some constant c. Using equation (3.12), solve

$$c = \mathrm{Tr}[\rho_j^{(0)} J_{\hat{n}}] \left(1 - O\left(\frac{1}{\lambda}\right)\right)^{t_c}, \tag{3.31}$$

for t_c. Therefore,

$$t_c = O\left(j^2\right). \tag{3.32}$$

To generalize to higher moments, observe that $\mathrm{Tr}[\rho_j^{(0)} J_{\hat{n}}^i] \leqslant O(j^i)$ for $i < \ell$. Recall that we assumed that $\mathrm{Tr}[\rho_j^{(0)} J_{\hat{n}}^\ell] = O(j^\ell)$. Using equation (3.12) (the facts that $A_\ell^{(\ell)} = 1 - O\left(\frac{1}{\lambda}\right)$ and $A_i^{(\ell)} = O(1)$ for $i < \ell$), we can extend this result to higher odd moments by using strong induction. Then the minimum value of t such that $\mathrm{Tr}[\rho_j^{(t)} J_{\hat{n}}^\ell] < c$ for a properly selected initial state $\rho_j^{(0)}$ is $O(j^2)$. For even moments, a

similar approach using equation (3.11) can be used to obtain an identical conclusion. □

Note that if the assumptions of Theorem 3.3 fail, yet there is a specific dependance on j of the map χ and the state $\rho_j^{(0)}$, equations (3.11) and (3.12) may still be useful for studying the longevity on quantum reference frame. In the following section the applicability of these assumptions is further explored in the context of several examples.

3.6 Examples

3.6.1 Measuring Spin-1/2 Systems

In this section, the work of [BRST06] is revisited in the notation used in this work. They considered the specific process of using a spin-j system, initially in an optimal state, as a reference for repeatedly measuring spin-1/2 systems that began in the completely mixed state to determine whether they are aligned with or against the direction of the reference, using the map for this task that corresponds to the optimal protocol. The resulting disturbance map that they considered is

$$\xi(\rho_j) = \left(\frac{1}{2} + \frac{1}{2(2j+1)^2}\right)\rho_j + \frac{2j(j+1)}{(2j+1)^2}\zeta(\rho_j)\,, \tag{3.33}$$

where $\zeta(\rho_j)$ is given by equation (3.9). This form of covariant map was analyzed in [Rit05]. This specific map corresponds to the process of projecting the joint system of the spin-1/2 and the spin-j particles onto an eigenstate of well-defined joint total angular momentum.

The fidelity is defined as the probability that the task is performed successfully. If the reference frame used is classical (in the limit $j \to \infty$) then the measurement always returns the correct result of the measurement: if the spin is projected from the mixed state to the direction against the reference, the result is "down" and if the spin is projected along the reference direction, the result is "up" and the reference is not altered by this process. When a quantum reference with a finite total angular momentum is used, this is not always the case. The fidelity *after* being used k times is found by considering the projection onto the correct joint

angular momentum state:

$$F_{\frac{1}{2}}^{(k)} = \mathrm{Tr}\left[\frac{1}{2}\sum_{\mu\in\{-\frac{1}{2},\frac{1}{2}\}}\Pi_{j+\mu}(\rho_j^{(k)}\otimes|\mu\rangle\langle\mu|)\right] \tag{3.34}$$

$$= \frac{1}{2}\sum_m\Big(\langle j,m|\left\langle\tfrac{1}{2}\right|\Pi_{j+\frac{1}{2}}(\rho_j^{(k)}\otimes\left|\tfrac{1}{2}\right\rangle\left\langle\tfrac{1}{2}\right|)\Pi_{j+\frac{1}{2}}|j,m\rangle\left|\tfrac{1}{2}\right\rangle + \langle j,m|\left\langle-\tfrac{1}{2}\right|\Pi_{j-\frac{1}{2}}(\rho_j^{(k)}\otimes\left|-\tfrac{1}{2}\right\rangle\left\langle-\tfrac{1}{2}\right|)\Pi_{j-\frac{1}{2}}|j,m\rangle\left|-\tfrac{1}{2}\right\rangle\Big)\,. \tag{3.35}$$

The terms in this equation are given by the Clebsch-Gordan coefficients [CS51]:

$$\langle j,m|\left\langle\tfrac{1}{2}\right|\Pi_{j+\frac{1}{2}}|j,m\rangle\left|\tfrac{1}{2}\right\rangle = \sqrt{\tfrac{j+m+1}{2j+1}} \tag{3.36}$$

$$\langle j,m|\left\langle-\tfrac{1}{2}\right|\Pi_{j-\frac{1}{2}}|j,m\rangle\left|-\tfrac{1}{2}\right\rangle = \sqrt{\tfrac{j+m}{2j+1}}\,. \tag{3.37}$$

Noting that the expectation value of the operator J_z is $\mathrm{Tr}[\rho J_z] = \langle m\rangle_\rho$, the fidelity function can be expressed in terms of the first moment of the reference frame as

$$F_{\frac{1}{2}}^{(k)} = \mathrm{Tr}\left[\frac{1}{2}\sum_{\mu\in\{-\frac{1}{2},\frac{1}{2}\}}\Pi_{j+\mu}(\rho_j^{(k)}\otimes|\mu\rangle\langle\mu|)\right] = \frac{1}{2}+\frac{1}{2}\frac{1}{2j+1}\mathrm{Tr}[\rho_j^{(k)}J_z]\,. \tag{3.38}$$

It is possible to express the fidelity in terms of the *original* value of the first moment, *i.e.* in terms of the initial state of the reference frame. This requires the use of Theorem 3.2 and a simple calculation to evaluate $A_1^{(1)}$, as defined in equation (3.12). First note that:

$$\mathrm{Tr}[\rho_j^{(k)}J_{\hat{n}}] = A_1^{(1)}\mathrm{Tr}[\rho_j^{(k-1)}J_{\hat{n}}] = \left(A_1^{(1)}\right)^k\mathrm{Tr}[\rho_j^{(0)}J_{\hat{n}}]\,. \tag{3.39}$$

An expression for $A_1^{(1)}$ can be found by considering the action of ξ (the map describing the measurement of the spin-1/2 systems) on the state $|j,m\rangle$. Notice,

$$\mathrm{Tr}[\xi(|j,m\rangle\langle j,m|)J_z] = A_1^{(1)}m \tag{3.40}$$

and that the effect of the map ξ after the trace is taken is given by a sum over the

different ways of combining and separating a pair of spins:

$$\begin{aligned}\mathrm{Tr}[\xi(|j,m\rangle\langle j,m|)J_z] \\ &= \sum_{\mu'',\mu',\mu\in\{-\frac{1}{2},\frac{1}{2}\}} \left|\langle j,m-1|\left\langle\tfrac{1}{2}\right|\Pi_{j+\mu'}|j,m\rangle\left|-\tfrac{1}{2}\right\rangle\right|^2(m-1) \\ &\quad + \left|\langle j,m|\langle\mu''|\Pi_{j+\mu'}|j,m\rangle|\mu\rangle\right|^2 m \\ &\quad + \left|\langle j,m+1|\left\langle-\tfrac{1}{2}\right|\Pi_{j+\mu'}|j,m\rangle\left|\tfrac{1}{2}\right\rangle\right|^2(m+1). \qquad (3.41)\end{aligned}$$

The terms in this equation are the products of Clebsch-Gordan coefficients, for example the first term is:

$$\begin{aligned}\left|\langle jm-1;\tfrac{1}{2}|j-\tfrac{1}{2}m-\tfrac{1}{2};-\tfrac{1}{2}\rangle\langle j-\tfrac{1}{2}m-\tfrac{1}{2};-\tfrac{1}{2}|jm;-\tfrac{1}{2}\rangle\right|^2(m-1) \\ = \left(\frac{j+m}{2j+1}\right)\left(\frac{j-m+1}{2j+1}\right)(m-1). \qquad (3.42)\end{aligned}$$

Substituting all the appropriate Clebsch-Gordon coefficients gives the expression:

$$\begin{aligned}\mathrm{Tr}[\xi(|j,m\rangle\langle j,m|)J_z] &= 2\left(\frac{j+m}{2j+1}\right)\left(\frac{j-m+1}{2j+1}\right)(m-1) \\ &\quad + \left(\left(\frac{j+m}{2j+1}\right)+\left(\frac{j-m+1}{2j+1}\right)\right. \\ &\quad\quad \left. +\left(\frac{j+m+1}{2j+1}\right)+\left(\frac{j-m}{2j+1}\right)\right)m \\ &\quad + 2\left(\frac{j+m+1}{2j+1}\right)\left(\frac{j-m}{2j+1}\right)(m+1) \qquad (3.43) \\ &= 2m\left(1-\frac{2}{(2j+1)^2}\right), \qquad (3.44)\end{aligned}$$

which gives $A_1^{(1)} = 1 - \frac{2}{(2j+1)^2}$. This is all the information needed to construct the expression for the evolution of the fidelity:

$$F_{\frac{1}{2}}^{(k)} = \frac{1}{2} + \frac{\mathrm{Tr}[\rho_j^{(0)}J_z]}{2j+1}\left(1-\frac{2}{(2j+1)^2}\right)^k. \qquad (3.45)$$

All of the subsequent calculations of the coefficients $A_i^{(\ell)}$ and the fidelities given in this chapter can be performed in the same way, though the details are omitted for brevity. Note that the fidelity function in this particular case depended only on the first moment.

Figure 3.3 shows the decay of the fidelity of the measurement using the reference

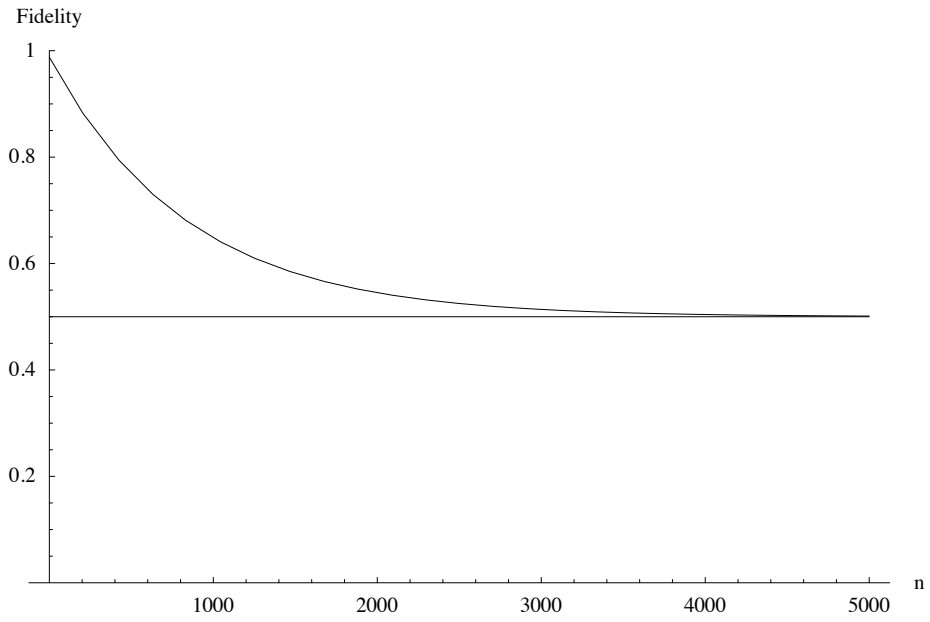


Figure 3.3: The fidelity of the measurement on spin-$\frac{1}{2}$ systems using the reference system of size $j = 20$ after n uses of it.

system with the repeated action of the map for the choice of $j = 20$ and an initial moment of $\mathrm{Tr}[\rho_j^{(0)} J_z] = j$. It will not fall below $\frac{1}{2}$, since, if it did, this would indicate that the state of the reference was becoming anti-correlated to the initial direction, and thus could become useful again. Instead the state of the reference becomes increasingly mixed until it does not provide a better measurement than a simple guess of which direction the spin-$\frac{1}{2}$ system is pointing, with a probability of $\frac{1}{2}$ of being correct.

3.6.2 Measuring Spin-1 Systems

This type of measurement problem can be extended to measuring systems that are not two dimensional. Now that the framework is developed, larger spins can be handled via this same technique. Suppose that the reservoir consists of spin-s systems which are initially in the completely mixed state. The quantum reference direction is a spin-j system with $j > s$, aligned with some classical direction $\hat{n}$. The goal is to measure each spin to determine its component μ along the vector $\hat{n}$ by performing a joint measurement on both the system from the reservoir and the reference system. The optimal rotationally-invariant joint operation for this task [BRS04] is a POVM given by the projectors $\{\Pi_{j+\mu} | \mu \in \{-s, \ldots, s\}\}$ where

$\Pi_{j+\mu}$ corresponds to a projection onto the subspace where the total angular momentum of the reference spin with the measured spin is $j+\mu$. Define the fidelity F_s as the probability that, when the reference is used to measure a system in a *known* state $|s,\mu\rangle_{\hat{n}}$ by means of the above POVM on the joint system the result it returns is correct, assuming that all values of μ are equiprobable.

Now consider the new question of measuring the angular momentum of a spin-1 particle along some direction using a quantum direction reference frame. First, an expression is determined for the fidelity in terms of the moments:

$$F_1^{(k)} = \mathrm{Tr}\left[\frac{1}{3}\sum_{\mu}\Pi_{j+\mu}(\rho_j^{(k)}\otimes|\mu\rangle\langle\mu|)\right], \tag{3.46}$$

where $\mu\in\{-1,0,1\}$. For this measurement, the disturbance map ξ is

$$\xi(\rho_j^{(k)}) = \frac{1}{3}\sum_{\mu'',\mu',\mu}\langle\mu''|\,\Pi_{j+\mu'}\,|\mu\rangle\;\;\rho_j^{(k)}\;\;\langle\mu|\,\Pi_{j+\mu'}\,|\mu''\rangle\,. \tag{3.47}$$

Explicit expressions for the reduced operators on the reference system $\langle\mu''|\,\Pi_{j+\mu'}\,|\mu\rangle$ can be found from the Clebsch-Gordan coefficients. The expression for the fidelity in terms of j and the moments is

$$F_1^{(k)} = \frac{1}{6} + \frac{[(2j+1)^2-2]}{6j(j+1)(2j+1)}\mathrm{Tr}[\rho_j^{(k)}J_z] + \frac{1}{2j(j+1)}\mathrm{Tr}[\rho_j^{(k)}J_z^2]. \tag{3.48}$$

Note that the fidelity in this case depends on the second moment as well as the first. Using Theorem 3.2:

$$\mathrm{Tr}[\xi(\rho_j^{(k)})J_z] = A_1^{(1)}\mathrm{Tr}[\rho_j^{(k)}J_z], \tag{3.49}$$

$$\mathrm{Tr}[\xi(\rho_j^{(k)})J_z^2] = A_0^{(2)} + A_2^{(2)}\mathrm{Tr}[\rho_j^{(k)}J_z^2]. \tag{3.50}$$

Substituting some particular values of ρ_j (i.e., $\rho_j = |j,m\rangle_{\hat{n}}\langle j,m|$ for $m=j,j-1$) and solving the system of equations:

$$A_1^{(1)} = \frac{3j^4+6j^3-j^2-4j+2}{3j^2(j+1)^2} = 1-\frac{4}{3j^2}+O\left(\frac{1}{j^3}\right), \tag{3.51}$$

$$A_0^{(2)} = \frac{2\,(8j^4+16j^3-8j-3)}{3j(j+1)(2j+1)^2} = \frac{4}{3}-\frac{5}{3j^2}+O\left(\frac{1}{j^3}\right), \tag{3.52}$$

$$A_2^{(2)} = \frac{4j^6+12j^5-3j^4-26j^3+j^2+16j+6}{j^2(j+1)^2(2j+1)^2} = 1-\frac{4}{j^2}+O\left(\frac{1}{j^3}\right). \tag{3.53}$$

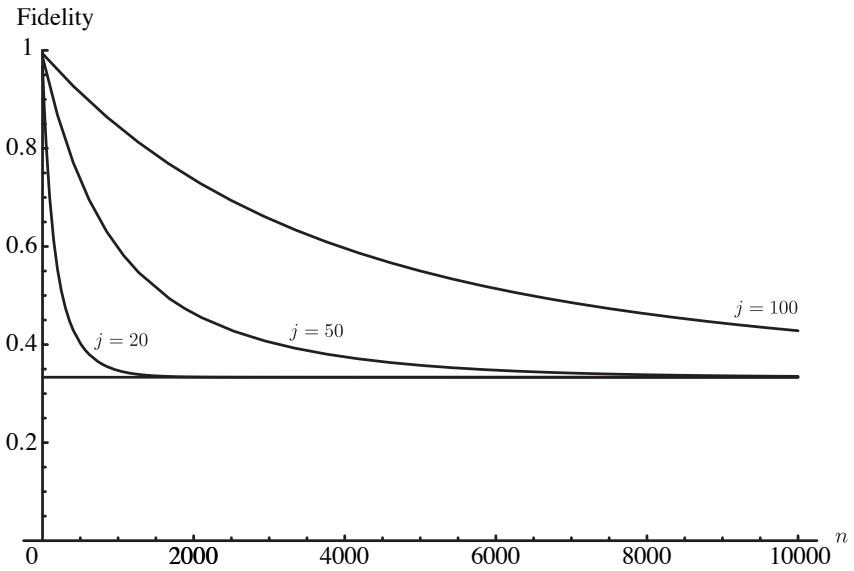


Figure 3.4: The fidelity of the measurement of the direction of spin-1 systems using reference systems of size $j = 20, 50$, and 100 after n uses of the reference.

(The series and bounds apply as $j \to \infty$.) In terms of these constants, the fidelity evolves with k, the number of applications of the map corresponding to the measurement of the z projection of the spin-1 system, as:

$$\begin{aligned} F_1^{(k)} &= \frac{1}{6} + \frac{[(2j+1)^2 - 2]}{6j(j+1)(2J+1)} \left(A_1^{(1)}\right)^k \mathrm{Tr}[\rho_j^{(0)} J_z] \\ &\quad + \frac{1}{2j(j+1)} \left(A_0^{(2)} \frac{1 - \left(A_2^{(2)}\right)^k}{1 - A_2^{(2)}} + \left(A_2^{(2)}\right)^k \mathrm{Tr}[\rho_j^{(0)} J_z^2] \right). \end{aligned} \tag{3.54}$$

The decay of this fidelity is illustrated in Figure 3.4, for three different values of j: $j = 20, 50$, and 100. Notice that in this case the fidelity drops as low as $\frac{1}{3}$, asymptotically, since a random guess amongst three equally likely outcomes is right $\frac{1}{3}$ of the time. This also demonstrates that larger reference systems (those with a higher value of j) are more robust to repeated use. That is, they should have higher longevities. This is shown in Figure 3.5, where in plot A the logarithm of the longevity is plotted against the logarithm of j.

Because the assumptions given in Section 3.5 are satisfied, the longevity of the quantum reference frame must scale as $O(j^2)$. This result is easily verified numerically. See Figure 3.5.

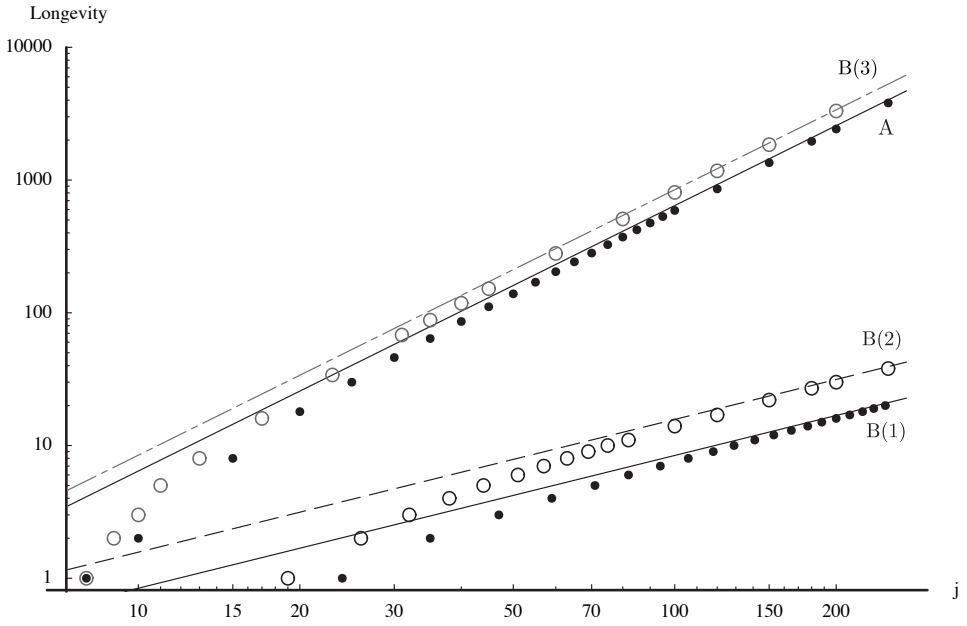


Figure 3.5: A plot of the longevity, as defined in Section 3.5, against the reference system of size j for, A, the scenario Section 3.6.2 measuring spin-1 systems (points converging to the solid line of gradient 2), and, B, the case of qubit rotation, Method 1, measure and rotate (points converging to the solid line of gradient 1), Method 2, filtering operation (circles converging to the dashed line of gradient 1), and Method 3, coupled spins (grey circles converging to grey dot-dashed line of gradient 2). The points are found numerically. The gradients of the lines in A and B(3) are 2, indicating longevity scaling as $O(j^2)$, and the gradients of the lines in B(1) and B(2) are 1, indicating a scaling of $O(j)$.

3.6.3 Implementing a Pauli Operator on a Qubit

Another case in which a quantum reference frame could be used is to define a particular unitary gate. Since the reference state considered here only defines one axis, it can only be used to implement the set of gates corresponding to rotations about this axis. Suppose, for example, the desire is to implement a Pauli Z operation on a qubit,

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.55}$$

using a quantum reference direction ρ_j in the form of a spin-j system to define the z-axis. To do this, a rotationally-invariant operation must be implemented on the combined reference and system. Unlike the measurement considered in previous examples, in this case the task is to implement a unitary operation on the system that is *conditioned* on the state of the quantum reference direction, or alternatively,

to perform a unitary rotation about a direction *programmed* by the state of the reference frame. However, it is not possible to construct an ideal conditional unitary because the size of the reference frame is bounded and the number of programmed operations (possible directions) is continuous, which follows from the impossibility of a "programmable quantum gate" [NC97]. It is possible to approximate it, as is now demonstrated in the examples.

Three different operations for achieving this gate are proposed and studied. The first is a measurement of the reference frame followed by the operation performed on the qubit conditioned on the measurement outcome. The next is a joint operation on the reference and the qubit in which the qubit is rotated coherently so that the z-axis matches the direction of the reference spin. This is different from the first case because the reference spin is never projected, so the range of angles about which the definition of "z-axis" for the qubit is distributed according to amplitudes, rather than a probability distribution over possible directions of the reference state as in the first case. The result is a filtering operation which is not trace preserving. The final one is also joint operation on the reference and the qubit, but one that is completed by coupling the different angular momentum components of the reference and the qubit. It is a trace preserving operation.

3.6.3.1 Method 1

One possible procedure for applying a rotation about the axis defined by the reference spin is to first measure the direction indicated by the reference and then perform the appropriate rotation to the qubit system, discarding the measurement result. The measurement result is discarded because the desire is to consider the reference information as being stored in a system of finite size, rather than measured once and then stored in an arbitrarily large classical system. Explicitly, consider the following operation, to be performed on the joint state of a quantum reference frame ρ_j and a single qubit σ:

$$\chi(\rho_j \otimes \sigma) = \int_\Omega d\mu_\Omega \sqrt{\Lambda(\Omega)} \otimes Z(\Omega)(\rho_j \otimes \sigma)\sqrt{\Lambda(\Omega)} \otimes Z(\Omega)^\dagger \,, \tag{3.56}$$

where the covariant measurement on the spin-j reference frame is described by the POVM

$$\{\Lambda(\Omega) = (2j+1)R(\Omega)|e\rangle\langle e|R(\Omega)^\dagger \,,\ \Omega \in SU(2)\} \,, \tag{3.57}$$

where $|e\rangle$ is a normalized state of a spin-j system. The measurement effects, $\Lambda(\Omega)$, of this POVM satisfy the normalization condition

$$\int_\Omega d\mu_\Omega \Lambda(\Omega) = \mathbb{I}_j \,, \tag{3.58}$$

where $\mathbb{I}_j$ is the identity on the spin-j Hilbert space and $d\mu_\Omega$ represents the Haar measure over $SU(2)$. This measurement is performed on the quantum reference frame state ρ_j, and then conditional upon obtaining the outcome Ω, the operation $Z(\Omega) = R(\Omega)ZR^{-1}(\Omega)$ is applied to the system. If subsequently the information about the measurement result Λ is discarded, then the effective joint map (3.56) is clearly invariant. The net operation on the system σ is given by the map

$$\tau(\sigma) = \int_\Omega d\mu_\Omega \mathrm{Tr}_j \left[\Lambda(\Omega)\rho_j\right] Z(\Omega)\sigma Z^{-1}(\Omega) \,. \tag{3.59}$$

Note that this operation τ is also rotationally invariant; although the measurement and subsequent unitary appear to require an external spatial reference frame, the net operation is invariant under changes of this frame. Also, because the operation is constructed explicitly from a POVM measurement and a unitary operation conditional on this classical result, followed by tracing out the state of the reference frame, it is necessarily completely positive and trace preserving (CPTP). (This fact is also clear by observation: the term $\mathrm{Tr}_j \left[\Lambda(\Omega)\rho_j\right]$ is a normalized probability distribution weighting possible unitaries $Z(\Omega)$, and thus the map τ is a valid unital CPTP map.)

This expression for τ explicitly gives a Kraus decomposition of this map (albeit with a continuum of Kraus operators), $\tau(\sigma) = \int_\Omega d\mu_\Omega E(\Omega)\sigma E(\Omega)^\dagger$, where

$$E(\Omega) = \sqrt{\mathrm{Tr}_j[\Lambda(\Omega)\rho_j]}Z(\Omega) \,, \tag{3.60}$$

are Kraus operators satisfying $\int_\Omega d\mu_\Omega E(\Omega)^\dagger E(\Omega) = \mathbb{I}$. The ability of this operation to approximate the Z operation on the system is defined using the *gate fidelity* [HHH99, BOS+02, Nie02, EAZ05] given by

$$F_{\mathrm{gate}}(Z,\tau) \equiv \frac{\int_\Omega d\mu_\Omega \left|\mathrm{Tr}[E(\Omega)^\dagger Z]\right|^2 + d}{d^2 + d} \tag{3.61}$$

where d (the dimension of the system on which the gate is applied) is 2 in this case.

Suppose that quantum directional reference frame ρ_j is aligned with the z-axis. Then ρ_j is a mixture of states $|j,m\rangle_z$ for $-j \leqslant m \leqslant j$. For simplicity,

let $\rho_j = |j,m\rangle_z \langle j,m|$. This is a pure state, but the analysis follows for convex combinations. The state $|e\rangle$ that defines the POVM could be any state, but for simplicity consider only the case where $|e\rangle = |j,j\rangle_z$. In this case, the rotationally-invariant operation on the combined reference and system is

$$\tau(\sigma,\rho_j) = (2j+1)\int_\Omega d\mu_\Omega \big|_z\langle j,m|R(\Omega)|j,j\rangle_z\big|^2 Z(\Omega)\sigma Z^{-1}(\Omega)\,. \tag{3.62}$$

The Kraus operators are then given by

$$E(\Omega) = \sqrt{(2j+1)}\big|_z\langle j,m|R(\Omega)|j,j\rangle_z\big|Z(\Omega)\,. \tag{3.63}$$

The irreducible representations of SU(2) can be parameterized

$$R(\Omega) = e^{i\alpha}\begin{pmatrix} e^{i\phi}\cos\theta & e^{i\psi}\sin\theta \\ -e^{-i\psi}\sin\theta & e^{-i\phi}\cos\theta \end{pmatrix} \tag{3.64}$$

where $0 \le \alpha,\psi,\phi \le 2\pi$ and $0 \le \theta \le \frac{\pi}{2}$. So, the rotation $R(\Omega)$ can be rewritten using Euler angles:

$$R(\Omega) = e^{i\alpha}R_z(-\phi-\psi)R_y(-2\theta)R_z(\psi-\phi), \tag{3.65}$$

where $R_k(\beta)$ is a clockwise rotation of angle β around the k-axis. It follows that

$$\mathrm{Tr}[R(\Omega)ZR^\dagger(\Omega)Z] = \mathrm{Tr}[R_y(-2\theta)ZR_y(2\theta)Z] = 2\cos 2\theta. \tag{3.66}$$

From [Ros57],

$${}_z\langle j,m|\,R(\Omega)\,|j,j\rangle_z = e^{-i\alpha-im(\phi+\psi)-ij(\phi-\psi)}\sqrt{\binom{2j}{j+m}}(\cos\theta)^{j+m}(-\sin\theta)^{j-m}. \tag{3.67}$$

Therefore,

$$\begin{aligned}|\mathrm{Tr}[E(\Omega)^\dagger Z]|^2 &= (2j+1)\binom{2j}{j+m}(\cos\theta)^{2j+2m}(\sin\theta)^{2j-2m}\big|\mathrm{Tr}[R_y(-2\theta)ZR_y(2\theta)Z]\big|^2 \\ &= 4(2j+1)\binom{2j}{j+m}(\cos\theta)^{2j+2m}(\sin\theta)^{2j-2m}(\cos 2\theta)^2,\end{aligned} \tag{3.68}$$

where, in the first step, the cyclic property of the trace was used. The Haar measure

in these coordinates is $d\Omega(U(2)) = (2\pi)^{-3} 2 \sin\theta \cos\theta d\theta d\alpha d\phi d\psi$.[1] It follows that

$$F_{gate}(Z,\tau) = \frac{1}{3} + \frac{2(j+1+2m^2)}{3(j+1)(2j+3)} \tag{3.69}$$

which is a function of the second moment of the z projection of the reference frame. The result easily generalizes to the case where ρ_j is a mixture of pure states of the form $|j,m\rangle_z$:

$$F_{gate}(Z,\tau) = \frac{1}{3} + \frac{2(j+1+2\mathrm{Tr}[\rho_j J_z^2])}{3(j+1)(2j+3)}. \tag{3.70}$$

This example is a case where the fidelity evolves with the expectation value of the second moment, despite the fact that the reservoir is composed of spin-$\frac{1}{2}$ particles. Notice that in Section 3.6.1, where measurements of spin-$\frac{1}{2}$ particles are considered, the fidelity only depends on the first moment. This indicates that rotation operations decohere the reference system differently than measurement operations. Specifically, in this particular case the fidelity does not depend on any odd moments because the operation depends only on the axis defined by $\hat{n}$, but not on the direction along this axis, since it is a rotation through an angle π. The rotations $R_{\hat{n}}(\pi)$ and $R_{-\hat{n}}(\pi)$ are equivalent transformations. Therefore, the sign of the J_z^ℓ moments cannot effect the fidelity, which will be the case if the fidelity only depends on the even moments.

Now consider how the quantum reference frame degrades with repeated use in performing this operation. Assume that the qubit systems (the reservoir) on which the approximate phase gate is applied are all in the completely mixed state $\frac{1}{2}\mathbb{I}_{\frac{1}{2}}$, so that the form of the map (3.10) for $\xi(\rho)$ applies. With each application, the reference frame evolves according to the invariant map $\xi(\rho_j) = \mathrm{Tr}_s \chi(\rho_j \otimes \frac{1}{2}\mathbb{I}_{\frac{1}{2}})$, where Tr_s is the partial trace of the qubit system on which the approximate phase gate is applied. To calculate how the second moment evolves as a function of the number of times the quantum reference frame has been used to perform the approximate phase gate, Theorem 3.2 can be used:

$$\mathrm{Tr}[\xi(\rho_j) J_z^2] = A_0^{(2)} + A_2^{(2)} \mathrm{Tr}[\rho_j J_z^2]. \tag{3.71}$$

To find the values of the coefficients $A_0^{(2)}$ and $A_2^{(2)}$, consider two possible initial states of the quantum reference frame. First, suppose that $\rho_j = \frac{1}{2j+1}\mathbb{I}_j$, which

[1] The integral is taken over the Haar measure for $U(2)$ instead of $SU(2)$ because this simplifies the notation, however, it should be noted that in this case the integral over α (the global phase) will not alter the result.

yields $\xi(\rho_j) = \frac{1}{2j+1}\mathbb{I}_j$. This evolution gives

$$\frac{j(j+1)}{3} = A_0^{(2)} + A_2^{(2)}\frac{j(j+1)}{3}. \tag{3.72}$$

Second, using $\rho_j = |j,j\rangle_z \langle j,j|$, a second equation can be obtained:

$$\frac{j(1+j(3+j+2j^2))}{(1+j)(3+2j)} = A_0^{(2)} + A_2^{(2)} j^2. \tag{3.73}$$

Solving those equations,

$$A_0^{(2)} = j - \frac{2j}{2j+3} = j - 1 + \frac{3}{2j} - \frac{9}{4j^2} + O\left(\frac{1}{j^3}\right), \tag{3.74}$$

and

$$A_2^{(2)} = 1 - \frac{3(2j+1)}{(2j+3)(j+1)} = 1 - \frac{3}{j} + O\left(\frac{1}{j^3}\right). \tag{3.75}$$

From the above equations, it can be deduced that the longevity of the quantum reference frame is $O(j)$ by noting that the fidelity after k repetitions of the gate is

$$F_{gate}^{(k)}(Z,\tau) = \frac{1}{3} + \frac{2}{3(j+1)(2j+3)}\left(j + 1 + 2A_0^{(2)}\frac{1-\left(A_2^{(2)}\right)^k}{1-A_2^{(2)}} + 2\left(A_2^{(2)}\right)^k \mathrm{Tr}[\rho_j^{(0)} J_z^2]\right). \tag{3.76}$$

The equations (3.74) and (3.75) imply that $A_0^{(2)}$ scales as $O(j)$ and $A_2^{(2)}$ goes as $O(1)$. Looking at equation (3.76), the fidelity must scale as $O(\frac{1}{j})$ and therefore longevity goes as $O(j)$. This is also seen in the numerical work shown in Figure 3.5. This result appears to be in contradiction with Theorem 3.3, but this is resolved by the fact that one of the assumptions of the theorem is not fulfilled. In this case, even though the map χ is rotationally invariant, it does not conserve the total angular momentum. In particular, note that $\langle j,m|_z\, \xi_j(|j,j\rangle_z \langle j,j|)\, |j,m\rangle > 0$ for all m in the range $-j,\ldots,j$. Because $\zeta^{\circ 2j}(|j,j\rangle_z \langle j,j|)$ is the only map of the form $\zeta^{\circ n}(|j,j\rangle_z \langle j,j|)$ for $0 < n < 2j+1$ that has support in the state $|j,-j\rangle_z \langle j,-j|$, then $q_{2j} > 0$. Therefore, there is no bound n_{max} independent of j such $q_n = 0$ for all $n > n_{max}$.

3.6.3.2 Method 2

There are methods for approximating a Pauli Z operation using a quantum reference frame other than the one just investigated. Another possibility is to do a joint operation on both the reference system and the qubit, such that the effective z axis for the qubit is rotated to match the direction of the reference spin coherently. This can be done using a *filtering operation* [Gis96], which is an operation related to a measurement that is not unitary or trace preserving in general. It is a map that distorts the state onto a subset of its Hilbert space. It could be represented, for example, by a POVM element. In particular, consider the filtering operation

$$\Gamma = (2j+1)\int_\Omega d\mu_\Omega R_j(\Omega)\otimes R_{1/2}(\Omega)\Big[|j,j\rangle_z\langle j,j|\otimes Z\Big]R_j(\Omega)^{-1}\otimes R_{1/2}(\Omega)^{-1}. \quad (3.77)$$

Observing that in spin notation

$$Z = |\tfrac{1}{2},\tfrac{1}{2}\rangle_z\langle\tfrac{1}{2},\tfrac{1}{2}| - |\tfrac{1}{2},-\tfrac{1}{2}\rangle_z\langle\tfrac{1}{2},-\tfrac{1}{2}|, \quad (3.78)$$

it is possible to replace the operator Z in equation (3.77) with this expression. Then, the tensor product written in the square brackets in equation (3.77) can be reexpressed using the Clebsch-Gordan coefficients as a joint system with total angular momentum having support on the $j-\frac{1}{2}$ and $j+\frac{1}{2}$ subspaces. The operation Γ is spatially-covariant and Schur's lemma can be used to show that it will be block-diagonal in the irreducible representations $j+\frac{1}{2}$ and $j-\frac{1}{2}$. Thus, it can be rewritten

$$\begin{aligned}\Gamma = (2j+1)\int_\Omega d\mu_\Omega R_{j+1/2}(\Omega)\Big[&|j+\tfrac{1}{2},j+\tfrac{1}{2}\rangle_z\langle j+\tfrac{1}{2},j+\tfrac{1}{2}| \\ &- \tfrac{1}{2j+1}|j+\tfrac{1}{2},j-\tfrac{1}{2}\rangle_z\langle j+\tfrac{1}{2},j-\tfrac{1}{2}|\Big]R_{j+1/2}(\Omega)^{-1} \\ - (2j+1)\int_\Omega d\mu_\Omega R_{j-1/2}(\Omega)&\Big[\tfrac{2j}{2j+1}|j-\tfrac{1}{2},j-\tfrac{1}{2}\rangle_z\langle j-\tfrac{1}{2},j-\tfrac{1}{2}|\Big]R_{j-1/2}(\Omega)^{-1} \\ = (2j+1)&\Big(\tfrac{1}{2j+2}(1-\tfrac{1}{2j+1})\mathbb{I}_{j+1/2} - \tfrac{1}{2j}\tfrac{2j}{2j+1}\mathbb{I}_{j-1/2}\Big), \end{aligned} \quad (3.79)$$

which simplifies to

$$\Gamma = (\frac{2j}{2j+2}\Pi_{j+1/2} - \Pi_{j-1/2}), \quad (3.80)$$

where Π_{j+k} is a projection into the subspace of total angular momentum $j+k$. Now define

$$\chi_j(\rho_j\otimes\sigma) = \Gamma(\rho_j\otimes\sigma)\Gamma^\dagger, \quad (3.81)$$

However, this map is not trace preserving, nor does it conserve angular momentum. As in method 1, let $\rho_j = |j,m\rangle_z \langle j,m|$. In that case, the Kraus operators are of the form

$$E_{m'} = {}_z\langle j,m'|\,\Gamma\,|j,m\rangle_z \tag{3.82}$$

and, using the appropriate Clebsch-Gordan coefficients, an expression is found for fidelity of the form

$$F_{gate} = \frac{1}{3} + \frac{2}{3}\left(\frac{1}{j+1}\right)^2 m^2, \tag{3.83}$$

which goes to 1 in the limit $j \to \infty$, for the case $m = j$. For a general state ρ_j (which must be diagonal in the basis $\{|j,m\rangle_z \mid -j \leqslant m \leqslant j\}$),

$$F_{gate} = \frac{1}{3} + \frac{2}{3}\left(\frac{1}{j+1}\right)^2 \mathrm{Tr}[\rho_j J_z^2]. \tag{3.84}$$

Even though the map is not trace preserving, by Theorem 3.2:

$$\mathrm{Tr}[\xi(\rho_j)J_z^2] = A_0^{(2)} + A_2^{(2)}\mathrm{Tr}[\rho_j J_z^2], \tag{3.85}$$

and, by the same method as in the previous subsection, for this map

$$A_0^{(2)} = \frac{j}{(j+1)} = 1 - \frac{1}{j} + \frac{1}{j^2} + O\left(\frac{1}{j^3}\right), \tag{3.86}$$

$$A_2^{(2)} = \frac{j}{(j+1)}\left(1 - \frac{3}{j(j+1)}\right) = 1 - \frac{1}{j} - \frac{2}{j^2} + O\left(\frac{1}{j^3}\right), \tag{3.87}$$

where the series expansions are appropriate in the limit $j \to \infty$.

Again it is found that the scaling of longevity with j to be $O(j)$. (See Figure 3.5.) However, the fidelity is higher than by using Method 1 for small numbers of repetitions of the map.

3.6.3.3 Method 3

Another idea to consider is performing an operation similar to the previous case, but in which trace is preserved on the map. Such a map could be realized through a coupling between the spins of the form seen in nuclear magnetic resonance:

$$H = w\left(J_x^{\left(\frac{1}{2}\right)}J_x^{(j)} + J_y^{\left(\frac{1}{2}\right)}J_y^{(j)} + J_z^{\left(\frac{1}{2}\right)}J_z^{(j)}\right) = \frac{w}{2}\mathcal{J}^2 - \frac{\omega}{2}\left(j(j+1) + \frac{3}{4}\right)\mathbb{I} \tag{3.88}$$

where w is a constant, the operator $J_x^{\left(\frac{1}{2}\right)} = X$ is the J_x operator on a spin-half system and the operator $\mathcal{J}^2$ is the total angular momentum operator on the joint system of the reference and the spin-half system. Recalling equation (2.68), $\mathcal{J}^2 = \left(j+\frac{1}{2}\right)\left(j+\frac{3}{2}\right)\Pi_{j+\frac{1}{2}} + \left(j-\frac{1}{2}\right)\left(j+\frac{1}{2}\right)\Pi_{j-\frac{1}{2}}$. This allows the unitary associated with this Hamiltonian $U = e^{i\frac{w}{2}(\mathcal{J}^2-(j^2+j+\frac{3}{4})\mathbb{I})t}$ to be defined. Choosing $t = \frac{2\pi}{w(2j+1)}$ gives the appropriate phase shift of the spin-half system relative to the spin-j reference system.

This can be readily related to the work of the previous subsection by now defining Γ as

$$\Gamma = U \cong (\mathbb{I}_{j+1/2} - \mathbb{I}_{j-1/2}), \tag{3.89}$$

up to a global phase. This differs from equation (3.80) by the factor on the $\mathbb{I}_{j+1/2}$ term.

In this case, the fidelity is given by

$$F_{gate} = \frac{1}{3} + \frac{2}{3}\left(\frac{2}{2j+1}\right)^2 \mathrm{Tr}[\rho_j J_z^2], \tag{3.90}$$

which, when $\mathrm{Tr}[\rho_j J_z^2] = j^2$, also tends to 1 in the limit $j \to \infty$.

For this map

$$A_0^{(2)} = 1 - \frac{1}{(2j+1)^2} = 1 - \frac{1}{4j^2} + O\left(\frac{1}{j^3}\right), \tag{3.91}$$

$$A_2^{(2)} = 1 - \frac{12}{(2j+1)^2} = 1 - \frac{3}{j^2} + O\left(\frac{1}{j^3}\right), \tag{3.92}$$

where, again, the series expansions are appropriate in the limit $j \to \infty$. Observe that in this case these parameters lack any terms in $\frac{1}{j}$.

The scaling of longevity with j are found to be $O(j^2)$, as depicted in Figure 3.5. Figure 3.6 compares the fidelities for the three methods for constant j. These last three examples demonstrate the importance of choosing the gate carefully to match the objectives of the given task.

3.7 Conclusion

This study has considered how a quantum directional reference frame state evolves under the action of maps invariant with respect to rotations in space. These maps

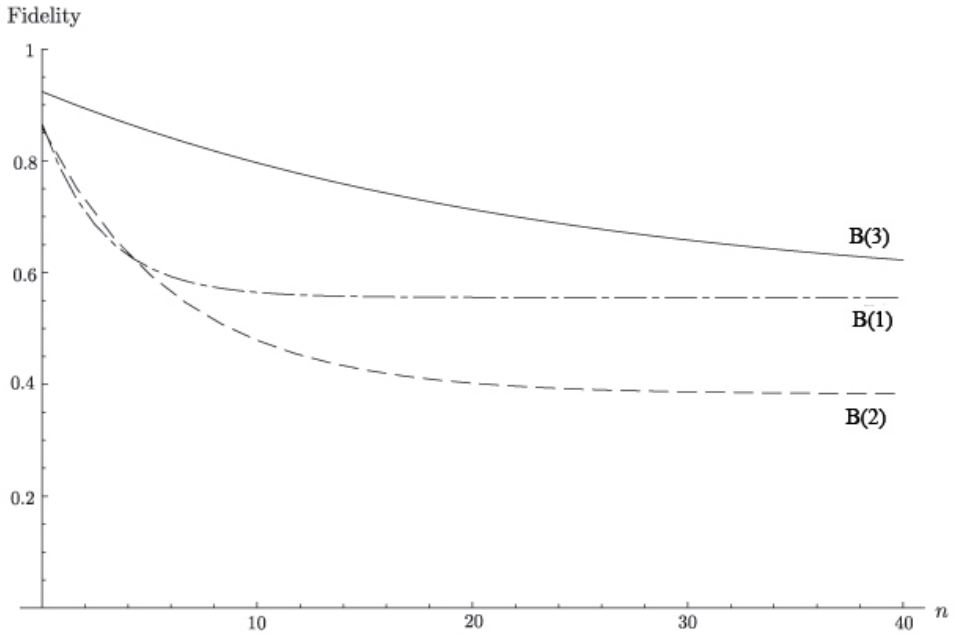


Figure 3.6: A plot of the fidelity with number of repetitions, n, for $j = 8$ for the three methods of qubit rotation, B(1) Section 3.6.3.1 measure and rotate (dot-dashed line), B(2) Section 3.6.3.2 filtering map (dashed line), and B(3) Section 3.6.3.3 coupled spins (solid line). This behavior of this value of j is representative.

are of physical interest because they describe the situation that occurs when an external, classical reference direction is lacking and there is rotational symmetry of the quantum system on which the quantum reference is to be used. The concept of quality of a quantum directional reference frame introduced by [BRST06] has been generalized. It has been demonstrated that the quality of a reference frame must be represented by a function that depends only on the eigenvalues of the quantum reference frame or an equivalent set of parameters called the moments. Equations (Theorem 3.2) are found that, used recursively, give expressions for how the moments evolve with the number of uses of the quantum reference frame. Some sufficient conditions (Theorem 3.3) for the longevity of a quantum reference frame to scale by a factor proportional to square the dimension of the quantum reference frame are given. Finally, the results are applied to different examples such as the use of a quantum directional reference frame to measure a spin-1 particle or to implement a Pauli operator on a qubit. The framework developed here can be used to compare different methods to perform some operation using a quantum reference frame and an example of such a case where this might be of interest is studied in

the final section.

Theorem 3.2 follows from Theorem 3.1, a theorem which implies that maps invariant with respect to $SU(2)$ can be written as a polynomial function of the Lie algebra generators of rotations. It would be interesting to investigate if a similar theorem could be applied to other Lie groups. One obvious case to consider is the group of $U(1)$ rotations, since invariance with respect to $U(1)$ describes the lack of a phase reference. However in this case, the group structure is not as rich as for $SU(2)$ since the representations of $U(1)$ are all one dimensional (*i.e.*, complex numbers z such that $|z| = 1$). This restricts the applicability of the methods used here to find constraints.

It is assumed in this analysis that the state of the reservoir is invariant under rotations. This is not the most general physical situation, since it is possible that the reservoir could be polarized, as is considered in [PY07]. It is an open question whether this theorem might be generalized in some way to the case where the rotational symmetry of the quantum system is broken.

Chapter 4

Quantum Reference Frames and Information Processing

Contents

4.1 Overview

In this chapter, results regarding requirements on reference frames in quantum computing, particularly for conserving energy, and the scaling of energy with error in a computation, are surveyed and combined. They are considered in conjunction with results from work on decoherence free subspaces, quantum cellular automata, and error correction to draw general conclusions about what restrictions the need for reference frames imposes and what are the resource requirements for quantum information processing.

4.2 Introduction

In the final section of the previous chapter, the concept that quantizing additional systems, such as measuring devices, allows for adherence to conservation laws was introduced. The same is true for performing operations on a quantum system. Every gate performed on a qubit or collection of qubits requires an interaction of those qubits with an external system. Since these external systems are required for quantum computation, it is important to consider them as a necessary resource in contemplating the practicality of quantum information processing.

For quantum computation, a system or collection of systems with physically distinguishable states is required. For example, a single qubit must have a state $|0\rangle$ and an orthogonal state $|1\rangle$. These states must be distinguished by some degree of freedom which differs in the two states. This degree of freedom could be (and typically is) energy. The physical difference between $|0\rangle$ and $|1\rangle$ allows measurements to distinguish the two states and operations to act in different ways on each state. This requirement implies there must be some Hamiltonian which breaks the symmetry between the two states by awarding one more energy than the other, causing the two states to have different Schrödinger evolution phases, *e.g.*, $H = E_0 |0\rangle\langle 0| + E_1 |1\rangle\langle 1|$ (note this is a rescaling of the Z operation). This allows for z-axis measurements and z-axis rotations, but this Hamiltonian will not mix the two states by transferring amplitude from one state to the other.

This degree of freedom will be associated to an operator that is considered a conserved quantity in classical physics. For example, in classical physics, energy is

strictly conserved, and likewise, the expected value of energy in a realistic model of quantum computation should be conserved so that a quantum computer together with its supporting apparatus is does not violate the first law of thermodynamics. However, the Hamiltonian associated with the degree of freedom does not give rise to all of the transformations that could be performed on a qubit. For example, we might wish to take the state $|0\rangle$ to $|1\rangle$ by means of an X operation. It is desirable to introduce a new Hamiltonian that can take the state $|0\rangle$ to the state $|1\rangle$, but that commutes with the Hamiltonian (*e.g.*, $H = E_0 |0\rangle \langle 0| + E_1 |1\rangle \langle 1|$) that breaks the symmetry.

This is possible if we introduce another system, which will be the reference frame. The reference frame can serve as a resource that allows for measurement or a near unitary state change to occur, as viewed by the qubit. This reference can exchange energy (or another degree of freedom) with the qubit and this will allow changes of amplitude between $|0\rangle$ and $|1\rangle$, while still maintaining constant energy overall. For example, consider the state of the reference to be $|\Phi_E\rangle$ where E denotes the expectation value for the energy of the state. Then consider the joint evolution of the system and the reference together:

$$|0\rangle \otimes |\Phi_E\rangle \mapsto |1\rangle \otimes |\Phi_{E-1}\rangle \tag{4.1}$$

where one unit of energy is transfered from the reference to the qubit. If the state of the reference is almost unchanged by gaining or losing and excitation (by virtue of having a large expected energy and an appropriate state, that is, one that is close to a classical state, not, for example, a large cat state), then the operation on the qubit can be approximately unitary because the coherence between states $|0\rangle$ and $|1\rangle$ can be preserved in the qubit when the reference is traced out. In other words, if we wish the interaction with the external quantum reference system to approximate well the case of an interaction with a classical system, the initial and final states of the reference should be close: $\langle \Phi_{E-1} | \Phi_E \rangle \approx 1$. In order to consider explicitly how this can work, we choose a particular physical system to explore, which is the qubits modeled as two-level atoms coupled to reference systems that are electromagnetic fields. The reason for this is that in most current implementations of quantum information processing the effects of the errors induced on the qubits by having a finite size reference system are negligible, because quantum computers currently have only a few coherent qubits run for several operations and the systems used to manipulate them are often composed of billions of atoms or excitations which are described exceptionally well classically. One practical scenario where it has been

suggested that this error might cause problems in the near future is the interaction of laser pulses with atomic qubits, where the pulses serve as a phase reference for the atomic qubits [BW99, KGBO09, GB02] due to finite laser coherence times, which are currently on the order of a few minutes, at the most.

The physical model of the world in which quantum systems can be employed for quantum computing is one which seemingly has contradictory requirements. It must be possible to model large systems classically, and yet, at the same time, there also must be no fundamental limit on how large a quantum system could be, so that they can be scaled up as needed. The first point is important because it allows the quantum computation system to undergo unitary evolution, as opposed to decohering when it is manipulated by external control systems. The second is desirable because the apparent speed up over classical processing in solving certain BQP problems becomes more pronounced as the input size of the problems increases. Superficially, these appear to be conflicting requirements. What is really necessary is that the systems which are modeled as classical are large compared to the quantum computation system. This leads to the question of how large do the classical systems need to be? If they must always be larger, but only by a linear or polynomial factor, this does not alter the problems which can be efficiently solved on a quantum processor. However, should this overhead grow superpolynomially in the input size to the problem, then finding the solutions would no longer be efficient in the total required resources.

It has been known for some time that conservation laws ought to impose restrictions on measurements and operations. Aharonov and Susskind [AS67] explicitly constructed a reference system for preparing a charge superposition state. For the case of performing a NOT gate on an atomic qubit using a Gaussian field state (coherent state), van Enk and Kimble [vEK01] found that such a gate could be performed up to an error of $O(\frac{1}{\bar{n}})$ where $\bar{n}$ is the expected number of photons in the Gaussian field. (See Section 4.6) Gea-Banacloche and Ozawa [GBO06] found that re-using such a field caused the fidelity to decrease with the square of the number of uses, k, so that the error on the kth gate is $O(\frac{k^2}{\bar{n}})$. Though they define $\bar{n}$ to be the total reference size across all k gates, therefore, their scaling agrees with van Enk and Kimble's result. Note that seems to imply that asymptotically the apparent speed up of algorithms such as Grover search would be negated, since a quadratic increase in the consumption of another resource is required.

It is possible to deduce fundamental restrictions on the fidelity of operations on quantum systems from commutation relations between the ideal operation and the operator associated with the quantity to be conserved. This has been done for gates

on one qubit [KGBO09], CNOT gates [Oza02], and multiplicative conservation laws [KMO08] as well as additive ones. Lower bounds for "gate infidelity" were found, and for arbitrary single qubit gates, the error must be greater than $\Omega\left(\frac{1}{N^2}\right)$ where N is the dimension of the ancilla system.

4.3 Energy Conservation

It is desirable for several reasons to have a model of quantum computing that conserves energy. For one, the expectation value of the energy should obey the correspondence principle so that classical thermodynamics is upheld, *i.e.*, a large quantum system, such as a computer should not be an engine that can generate energy from nothing. This is what appears to happen when a quantum system consisting of a string of qubits in the $|0\rangle$ ground state that has X gates applied to each qubit via unitary operations implemented by classical systems and all the qubits are mapped to excited states. Also, a quantum computer should not require an amount of energy to run that scales badly with the input size. Further, these interactions are no longer perfect unitary operations on the computation system, so it makes sense to consider how interactions with external systems alters the standard view of quantum computing. In particular, the error that occurs because the operations on the qubits will not be unitary is explored. First, however, we consider the definition of energy conservation.

This is important to consider because in order to find an appropriate scaling of this reference resource it is necessary to appreciate what systems must be quantized. The separation between quantum and classical must exist in all models of quantum computing, since the problem submitted to the computer is a classical question and the answer will also in the end be classical. This is simply the result of modeling the observer (who submits the input and obtains the output) classically. We would like to discover where is a good place to put this separation (the *Heisenberg cut*) so that necessary resources are not being neglected.

Hamiltonians cause states to evolve by altering the energy eigenstates, which causes different phases e^{iE_it} over a time t to be applied to the different eigenstates according to their different eigenvalues. The evolution is described by the Schrödinger equation:

$$H\sum_i \alpha_i |E_i\rangle = \sum \alpha_i E_i |E_i\rangle \tag{4.2}$$

where H is the applied Hamiltonian, $|E_i\rangle$ is the ith eigenstate of H, and E_i is its

eigenvalue. Any input state $|\psi\rangle = \sum_i \alpha_i |E_i\rangle$ can be expressed in this way as a superposition of eigenvectors of H. Under a constant Hamiltonian, the expected energy of the state $|\psi\rangle$, given by $\langle E\rangle_\psi = \mathrm{Tr}[E|\psi\rangle\langle\psi|] = \sum_i |\alpha_i|^2 E_i$, is fixed. However, changing the Hamiltonian on a state can cause a change in the expected energy of the state. An example of this is the X gate. Consider starting from the state $|0\rangle$ and moving to the state $|1\rangle$. Let the energy of the state $|0\rangle$ be zero and $|1\rangle$ be one by defining $H_Z = \frac{1}{2}(\mathbb{I} - Z)$. Clearly, the expectation value of energy has changed from the beginning to the end of this procedure since $\langle 0|H|0\rangle = 0$ and $\langle 1|H|1\rangle = 1$, but also there is an immediate change in the expected energy of the system as a result of applying the Hamiltonian which will implement the X gate, $H_X = \frac{1}{2}(\mathbb{I} - X)$. The eigenvectors of this Hamiltonian are $|+\rangle$ and $|-\rangle$ with eigenvalues 0 and 1 respectively. However, $|0\rangle$ can be written $|0\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)$. The expected energy of the state $|0\rangle$ under this Hamiltonian then is no longer 0, but $\frac{1}{2}$.

If an operator, A, commutes with the Hamiltonian then its expectation value $\langle A\rangle$ is preserved by the action of the Hamiltonian and the expectation value of a is said to be *conserved* during the time that this constant Hamiltonian is applied.

$$[A, H] = 0 \Rightarrow \quad \langle A\rangle \text{ is conserved.} \tag{4.3}$$

It is easy to see that this must be the case. Consider the system initially to be in a state $|\psi\rangle$. If a quantity is conserved over a time $0 \leq t \leq T$ then that means $a(0) = a(t)$ for all t in the specified range. The expectation value of the operator A is given by

$$\langle A\rangle_{\psi(0)} = \langle\psi(0)|A|\psi(0)\rangle$$

$$\langle A\rangle_{\psi(t)} = \langle\psi(t)|A|\psi(t)\rangle = \langle\psi(0)|U(t)^\dagger A U(t)|\psi(0)\rangle .$$

Using a Taylor expansion, the time evolution operator $U(t)$ is

$$U(t) = e^{-iHt} = \sum_{m=0}^{\infty} \frac{(-iHt)^m}{m!},$$

so that if an operator commutes with H it also commutes with U:

$$\begin{aligned}
[A,H]=0 \Rightarrow [A,U(t)] &= [A,\sum_m \frac{(-iHt)^m}{m!}] \\
&= \sum_m \sum_{k=0}^{m-1} \frac{(-it)^m}{m!} H^k [A,H] H^{m-k-1} \\
&= 0.
\end{aligned}$$

Then we can rearrange the expression for the expectation value as

$$\begin{aligned}
\langle A \rangle_{\psi(t)} &= \langle \psi(0)| U(t)^\dagger U(t) A |\psi(0)\rangle = \langle \psi(0)| A |\psi(0)\rangle \\
&= \langle A \rangle_{\psi(0)}. \qquad (4.4)
\end{aligned}$$

Therefore, the expectation value of that degree of freedom associated with any operator that commutes with the Hamiltonian is conserved under the action of that Hamiltonian.

It is possible to keep the expected energy in a system constant by employing a reference system to restrict to using operators that commute with H. The use of such a system will keep the number of excitations constant, as in Figure 4.1. Then the evolution on the joint system can be unitary. Consider the situation proposed in the previous section to use a qubit of reference together with a qubit of computation and let it start in the state $|01\rangle$.

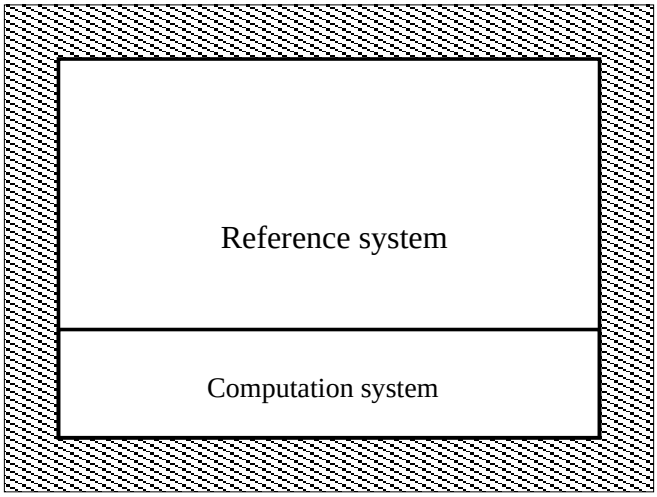


Figure 4.1: Fixing the total number of excitations in the joint system composed of the computation system and the reference system.

Then a X gate on the qubit will be accomplished by swapping the states of the computation qubit and reference qubit. The eigenstates of the swap operation, S, are $|\Psi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)$, but the expected energy of the state $|01\rangle$ does not change under this new Hamiltonian. Initially it has an energy 1 and under the swapping Hamiltonian $|01\rangle = \frac{1}{\sqrt{2}}\big(|\Psi^{+}\rangle + |\Psi^{-}\rangle\big)$ and the average energy of these terms is 1.

Notice that in the example of applying an X gate $[H_Z, H_X] = i \neq 0$. However, we can do a gate which serves as an effective X gate using the swap operation. In this case the two Hamiltonians do commute. Recall

$$S = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.5}$$

and

$$e^{iSt} = \cos(t)\mathbb{I} + i\sin(t)S. \tag{4.6}$$

So, let $H_S = \frac{1}{2}(\mathbb{I} - S)$. Then

$$[H_Z \otimes \mathbb{I} + \mathbb{I} \otimes H_Z, H_S] = 0. \tag{4.7}$$

If two measurement operators A and B commute with one another, then the two operators can have simultaneous well defined outcomes. That is, the system being measured can be in a state $|\psi\rangle$ such that $A|\psi\rangle = a|\psi\rangle$ and $B|\psi\rangle = b|\psi\rangle$ then $AB|\psi\rangle = BA|\psi\rangle = ab|\psi\rangle$ and the minimum possible joint uncertainty of these variables is zero, $\Delta a \Delta b \geq 0$. In addition, if two operators commute then they have joint eigenstates. In the case of the second example, the joint eigenstates are $|00\rangle, |11\rangle, |\Psi^{+}\rangle$, and $|\Psi^{-}\rangle$.

Then, any operations that commute with the total energy operator for a given system can be performed without disturbing the result of a measurement of total energy in the system. However, consider what happens within the $E = 1$ subspace that is spanned by the states $\frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ and $\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$. Under the Schrödinger evolution

$$e^{iH_Z t}\big(|\Psi^{+}\rangle + |\Psi^{-}\rangle\big) = e^{it}\big(|\Psi^{+}\rangle + |\Psi^{-}\rangle\big) \tag{4.8}$$

with the two states both acquiring the phase e^{it}. However, under the action of the

swap Hamiltonian this is not the case, rather

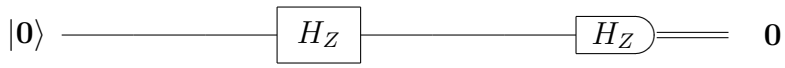
$$e^{iH_S t}(\left|\Psi^+\right\rangle + \left|\Psi^-\right\rangle) = e^{it}\left|\Psi^-\right\rangle + e^0\left|\Psi^+\right\rangle. \tag{4.9}$$

The states now have a relative phase.

This example demonstrates that even though the degree of freedom associated to the operator defined to be the "total energy operator", $H_{\text{total}} = H_Z \otimes \mathbb{I} + \mathbb{I} \otimes H_Z$, is conserved by the operation $e^{iH_S t}$, the act of applying H_S to the system must alter the relative energies of the states on that system. This is because H_S has different energy eigenstates than H_{total}. So the expected energy of the system $\langle H_S \rangle$ under H_S may differ from that of H_{total}. As an example of this consider the state $\alpha\left|\Psi^-\right\rangle + \beta\left|\Psi^+\right\rangle$. The expectation $\langle H_{\text{total}} \rangle = 1$, but the expectation $\langle H_S \rangle = 1|\alpha|^2 - 0|\beta|^2 \neq 1$ unless $|\alpha| = 1$. The operators H_{total} and H_S commute, but their expectation values will not be the same for all states.

Therefore the "energy" of the system's state as defined by the Hamiltonian operator dictating the Schrödinger equation changes with each new operation during a quantum computation. However, the "total energy" as defined by the consistent operator H_{total} can be conserved throughout.

Now we argue that a change of Hamiltonian is a requirement for quantum information processing. For quantum computation it does not suffice to prepare a state that is an eigenstate of the total energy, *e.g.*, the ground state, evolve it solely under this Hamiltonian, and measure it according to this same Hamiltonian operator. The state can acquire a global phase, but this cannot be detected physically, so it may as well be considered not to have evolved at all.

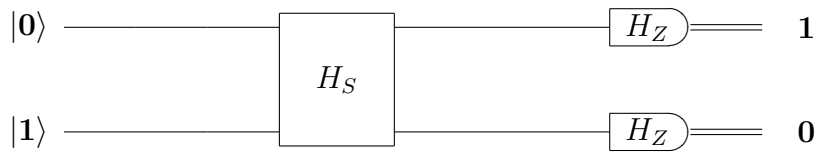


So, in order for there to be a non-trivial computation taking place, at the least a state must be prepared using one Hamiltonian, evolved according to a different one, and measured according to a third. In principle, the first and third Hamiltonians could be the same, but if the first and second or second and third are the same this is equivalent to preparing the state and measuring it immediately.

In this diagram the preparation is according to H_Z, the evolution is according to e^{iXt} and the measurement is again according to H_Z.

This implies that energy conservation in the strict sense of Noether's theorem, which arises from time-invariance does not lead to any detectable evolution quantum mechanically. This is logical, since if the system were evolving detectably, then this evolution should be able to serve as a clock which will break the time invariance.

Under the less strict definition that a particular operator which we shall call "total energy" must be conserved interesting evolutions can occur. This is illustrated in the diagram below.



Preparation and measurement are according to $H_Z \otimes H_Z$ which produces state that are eigenstates of H_{total}, and the evolution is according to H_S, but H_S preserves H_Z.

What we have just done when we replaced the X gate with the swap operation is move the Heisenberg cut, the conceptual line in the theory which separates systems considered classical and systems treated as quantum. In the case that a qubit is taken from $|0\rangle$ to $|1\rangle$ by the Hamiltonian e^{iXt}, some classical field is assumed to couple to the qubit. Being classical, it is treated mathematically as though it is unchanged by the coupling, even though it is clear that it has altered a property of the qubit: its expected energy. By adding an additional qubit the value $\langle H_{\text{total}} \rangle$ can be kept constant; the Heisenberg cut is moved so that the additional qubit is now included on the "quantum" side of the line. Note that the observer, and the observer's inputs, and the final measurement outcomes, are always on the "classical" side of the line.

What this demonstrates is that it is possible to replace a Hamiltonian in the quantum system by an interaction with an additional quantum system governed by a new Hamiltonian and further, that this can be used for the purpose of conserving the degree of freedom associated to an operator. This shift of the Heisenberg cut can clearly be repeated. The new Hamiltonian can itself be replaced by an interaction with yet another quantum system, governed by yet another Hamiltonian. Following the previous examples, this would allow the quantity $H_S \otimes \mathbb{I}_2 + \mathbb{I}_2 \otimes H_S$ to also be conserved (as well as $H_{\text{total}} = H_Z \otimes \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} + \mathbb{I} \otimes H_Z \otimes \mathbb{I} \otimes \mathbb{I} + \mathbb{I} \otimes \mathbb{I} \otimes H_Z \otimes \mathbb{I} + \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes H_Z$) by including two more qubits and making the effective X operation consist of a pair of swaps mapping $|0011\rangle \mapsto |1100\rangle$.

However, since the observer remains classical, there will never be completely quantum description of the evolution. It is better to place the Heisenberg cut somewhere convenient for the calculation. The intention in this work is to move the cut far enough to provide a rigorous accounting of what is required for a reference frame to be maintained so that quantum computation can occur. Therefore, we find the energy resources required to conserve H_{total} on the computation system and its reference state.

If a quantum computation were to require either overly large amounts of energy, even at a fundamental theoretical level, or else require a reference frame whose size scales superpolynomially, this would constitute a threat to the feasibility of quantum computation or a restriction on the class of problems a quantum computer could be expected to solve efficiently.

Fact 1. *For every new gate Hamiltonian applied to a quantum system an interaction between the qubit(s) that the gate acts on and an external system is required.*

Usually, the external system is modeled as classical, but here we will treat it as quantum mechanical, to properly assess its value as a resource. Gates can be separated into two kinds:

1. gates that do not commute with the total energy operator, H_{total}, and,
2. gates that do commute with H_{total}.

where now H_{total} is the total energy operator on the computation system only.

Both of these type of gates will require interaction with an externals system, which can be quantized. In the implementation setting of qubits being two-level atoms, both kinds of gate can be implemented using fields. However, the form of the interaction will be different in these two cases and the resource requirements for them are very different. For the first type of gate, non-commuting gates, the interaction exchanges excitations between the field (the reference) and the atom (the computation system). The external system in this case we consider to be a *reference frame*, since it will be necessary to keep coherence between the different energy states $|0\rangle$ and $|1\rangle$ of a qubit, as is explored in the next section. The second kind of gate, the commuting gates, require energy to be performed, but do not need to be kept coherent or correlated to the computation. They will be addressed in Section 4.9.

4.4 Coupling of the Computation System to the Reference

In the picture of energy conservation under which operations must commute with H_{total} this is equivalent to fixing the total number of excitations in the reference state and the qubits and then carrying the computation out according to the circuit model picture.

For example, gates such as Z, control-Z, phase gates, or fractional SWAPs, which commute with H_{total}, are assumed to be feasible without a phase reference system because they do not change the total number of excitations in the computation system, $\mathcal{C}$. The reference is required for any gate that can alter the expected number of excitations in computer system, such as an X, Y, or Hadamard, $\mathbf{H}$, gate, which will not commute with H_{total}, and need to be replaced with encoded operations on the joint reference-computation system.

The reference system, $\mathcal{R}$, will couple to the computation qubits and serve to preserve total energy while also providing a definition of phase, in the sense that it preserves coherences between states of different energy. Consider applying a Hadamard gate to a qubit in the $|0\rangle$ state:

$$\mathbf{H}\,|0\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big). \tag{4.10}$$

Now suppose that the Hadamard gate is performed via a mapping on the joint state of the reference and the computation systems, so that the number of excitations is conserved, where the reference is presumed to have n excitations:

$$|0\rangle\,|n\rangle \overset{\text{“}\mathbf{H}\text{”}}{\longmapsto} \frac{1}{\sqrt{2}}\big(|0\rangle\,|n\rangle + |1\rangle\,|n-1\rangle\big). \tag{4.11}$$

In the case of atomic qubits interacting with laser fields the Jaynes-Cummings Hamiltonian run for a time $t = \frac{2\sqrt{n}}{\pi g}$ will realize this interaction. If the reference state is traced out at this stage, the remaining state on the qubit is

$$\mathrm{Tr}_{\mathcal{R}}\big[\frac{1}{2}(|0\rangle\,|n\rangle + |1\rangle\,|n-1\rangle)(\langle 0|\,\langle n| + \langle 1|\,\langle n-1|)\big] = \frac{1}{2}\big(|0\rangle\,\langle 0| + |1\rangle\,\langle 1|\big) \tag{4.12}$$

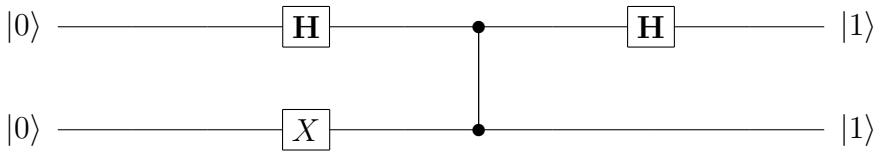
so all coherence is lost. It is possible to prepare a reference in a state that is close to the state of a classical field so that if can be traced out with limited decoherence effects. If the system $\mathcal{R}$ is in a coherent state $|\alpha\rangle$ containing a mean number of photons $\bar{n}$ and having a phase ϕ so that $\alpha = \sqrt{\bar{n}}e^{i\phi}$, then after an effective

Hadamard operation on the qubit the reference can be traced out to leave the qubit in the state $\frac{1}{\sqrt{2}}\big(|0\rangle + e^{i\phi}|1\rangle\big)$ up to an error of $O(\frac{1}{\bar{n}})$. This is considered in detail in Section 4.6. The reference can now be considered to be classically correlated to the qubit rather than entangled (with high fidelity). Notice that still, in this case, because the state of the qubit is correlated to that of the reference, if the reference system is discarded, even if it is classical, the state of the qubit will need to be represented as the completely mixed state.

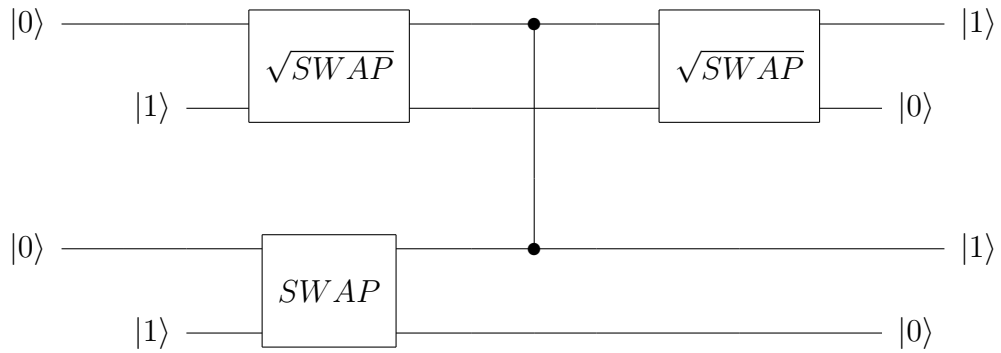
Fact 2. *If all two-qubit interactions within a computation system are ZZ couplings then the qubits do not require a common phase reference.*

In other words, a separate unentangled, uncorrelated reference system could be used in association with each qubit in the computation system so long as all two qubit interactions are for example, controlled-Z gates.

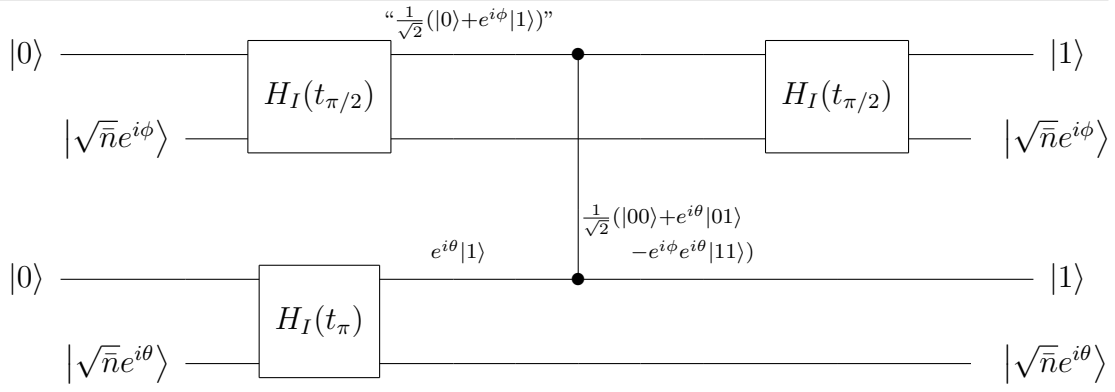
As an example, consider the circuit:



that takes input $|00\rangle$ to $|11\rangle$ by a phase kickback. An energy conserving implementation of this circuit using the encoding $|0_L\rangle = |01\rangle, |1_L\rangle = |10\rangle$ is:



where the ancilla qubits can be thought of as reference systems for the first and second qubit. Equivalently, independent coherent states that are not phase correlated could be used as references for each qubit:



Note that all single qubit gates with any two qubit entangling gate is universal for quantum computation. All single qubit gates that do not commute with Z can be accomplished by interactions between each reference system and each qubit. Also, each reference could begin in a state of fixed excitation number and since each interacts with only one qubit, the total number of excitations is fixed between that reference the that particular qubit. This means that there is no fundamental error in the application of the Jaynes-Cummings interaction, as is the case when a reference in a coherent state is used. In fact, the reference state could be reduced down to a system that contains a single excitation to be as economical with energy as possible. Now two physical qubits (one from the "computation system" and one from the "reference system") are required to simulate a single physical qubit and the computation takes place in the subspace of fixed excitation number. In this picture, if the computation system begins in the ground state, so that all qubits are initialized to $|0\rangle$ and the reference system begins in the state all qubits excited, $|1\rangle$, and the ideal final state on the computation qubits is a binary string, then this method should have no error at all due to the restriction that energy must be conserved.

A reason that large reference systems tend to be used in practice is that they are more robust. If the reference system is reduced to being a single excitation, then it is another qubit that an experimentalist must keep coherent. The field excitation may be trapped in a cavity, but there is some chance that it could leak from an imperfect cavity. If n excitations are used instead, and if one leaks from the cavity, this does not completely destroy the entanglement of the remaining excitations with the qubit:

$$\mathcal{N}a\frac{1}{\sqrt{2}}\big(|0\rangle\,|n\rangle+|1\rangle\,|n-1\rangle\big)=\frac{1}{\sqrt{n-1/2}}\big(\sqrt{n}\,|0\rangle\,|n-1\rangle+\sqrt{n-1}\,|1\rangle\,|n-2\rangle\big), \tag{4.13}$$

where $\mathcal{N}$ is some normalization, but this will result in an error of choosing the interaction time to achieve a particular rotation, *e.g.*, to implement $R_x(\theta/2)$ choose $t = \frac{\theta}{g\sqrt{n}}$, as well as shifting the relative amplitudes of $|0\rangle$ and $|1\rangle$. In general, the leakage will not be known to have occurred so the number of excitations in the cavity will not be n and the incorrect choice of t will be made. The result will be that the intended gate will not be precisely applied. However, if n is sufficiently large, due to the mis-timing error will be very small. This is considered in more detail in Section 4.6. In addition, the error caused by the shift of amplitude will be of order $\epsilon = O(\frac{1}{n^2})$.

Notice that even though different Hamiltonians must be switched off and on in this arrangement and that this will change the relative energy of the different sets of eigenvalues, this will not change the expected energy of the total system $\mathcal{R} \otimes \mathcal{C}$. If it begins in a state of well defined energy, *e.g.*, all computation qubits in the state $|0\rangle$ and all reference qubits in the state $|1\rangle$, then the joint system will also end in this well-defined-energy state. In this arrangement it is clear that the number of logical qubits is equal to the number of excitations required to complete the calculation, and further, that this energy input is effectively returned at the end of the calculation. That is the required energy scales linearly with the space requirement of the problem, not including the energy required to implement the gates that commute with the total energy operator.

There are then several approaches that might be taken to this problem of what to use as a reference system, which have different advantages and drawbacks.

1. Use an encoding for the qubits that keeps them in an energy-conserving subspace. This is considered in more detail in the next section. The advantage of this is that it is the most economical in terms of energy required in the reference state and in principle has no fundamental error due to reference concerns. The disadvantage is that it is not robust against photon loss.

2. Use an energy-conserving subspace, but allow the reference system to be a large photon number Fock state $|n\rangle$. These systems will be more robust, and will have faster interactions with the computation qubits, but will require more energy to construct.

3. Use coherent states as reference systems, one for each qubit. This has some error, but it is limited, and the coherent states need not be correlated, however, it requires maintaining as many cavities as qubits, each one containing a coherent reference state.

4. Use repeatedly the same coherent state for every non-commuting gate with every qubit. This is economical, only requiring one cavity, but the error scaling is worse.

5. Use as the reference a series of states each for one gate only that are mixtures of coherent states over phase, but phase correlated with each other. In other words the reference states are of the form

$$\int \left|\sqrt{\bar{n}}e^{i\phi}\right\rangle_1 \left\langle\sqrt{\bar{n}}e^{i\phi}\right| \otimes \left|\sqrt{\bar{n}}e^{i\phi}\right\rangle_2 \left\langle\sqrt{\bar{n}}e^{i\phi}\right| \otimes \ldots \otimes \left|\sqrt{\bar{n}}e^{i\phi}\right\rangle_G \left\langle\sqrt{\bar{n}}e^{i\phi}\right| d\phi$$

   where the subscript refers to which non-commutative gate each state will be used for $1, 2, ...G$. This is a model for what is usually currently done in practice. A single laser emits a series of Gaussian state pulses which are used to enact rotations on the qubits.

The first scheme is considered in the following sections, 4.4.1 and 4.5. It has already been contrasted with the second proposal in this section. The last three schemes are analyzed in Section 4.6 and compared in Section 4.6.1.

4.4.1 The Relation to Decoherence Free Subspaces

Note that the implementation which conserves the expectation of H_{total} uses an encoding that operates only in a subspace of the total Hilbert space $\mathcal{R} \otimes \mathcal{C}$, the subspace containing a constant number of excitations distributed among the qubits. Such a proposal was suggested by Kempe *et al.* [KBLW01]. They propose a scheme for combating "strong collective decoherence" in a logical qubit by encoding it in three physical qubits, where all logical qubit operations are performed by SWAP gates between the physical qubits. This conserves excitation number in the logical qubit. Also, extending a "conservative" encoding to many qubits becomes increasingly efficient. In the limit as the number of physical qubits $n \to \infty$, the ratio of the size of the largest decoherence-free subspace, k, which conserves energy, to n goes as:

$$\lim_{n\to\infty} \frac{k}{n} = 1 - \frac{3}{2}\frac{\log n}{n}. \tag{4.14}$$

Note the relation between this and Bartlett *et al.* [BRS07] for communication without a shared reference. The above formula also applies to the number of qubits of quantum information, k, that can be communicated between two parties that do not share a cartesian coordinate system using a transmission of n qubits. They also

find that the amount of classical information (k bits) that can be sent over such a quantum channel behaves as

$$\lim_{n\to\infty} \frac{k}{n} = 1 - \frac{1}{2}\frac{\log n}{n}, \tag{4.15}$$

which is the number of irreducible $SU(2)$ representations in the direct sum decomposition of the Hilbert space of n tensored qubits. However, in the case that the quantum channel also destroys ordering information on the transmitted qubits, that is, the receiver obtains all of the qubits but does not know the Cartesian frame of the sender or the order in which the sender transmitted the qubits, then communication is not possible. This is due to the fact that this restricts the parties to a symmetric subspace even before imposing the $SU(2)$ rotational invariance. This leads to the multiplicities of the spaces of different excitation number all being one and therefore there are no subspaces preserved under this channel as a consequence of Schur's first lemma.

Schur's first lemma (see Chapter 2, Section 2.5.2) states that, for matrix representations of a group G, closed under matrix multiplication and inversion and $R(g)$ are the group elements of an irreducible representation of G, then if a matrix A is commutes with $R(g)$ for all g then A is a scalar matrix, *i.e.* $A = c\mathbb{I}$ where c is a constant. [Sch05, Hal06]

This implies that the channel that destroys ordering information also cannot preserve any quantum information because destroying ordering information by twirling over all possible permutations

$$\mathcal{T}(\rho) = \sum_i \pi_i \rho \pi_i^\dagger \tag{4.16}$$

restricts the preserved states to the symmetric ones. That is, the states

$$\rho = \left(\sum_k a_k |k\rangle\right)\left(\sum_{k'} b'_k \langle k'|\right) \tag{4.17}$$

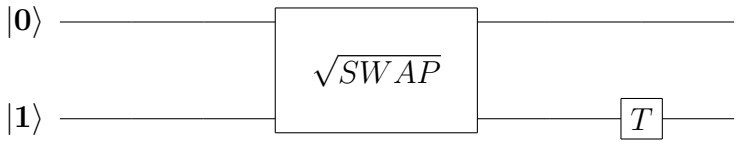
where $|k\rangle$ represents and equal superposition of all n-qubit strings with k qubits in the state $|1\rangle$. This is isomorphic to the states of each possible total angular momentum of n qubits. Then the irreducible representations of the group of all $SU(2)$ rotations is $\bigoplus_{k=0}^{n/2} \mathcal{H}_k$ if n is even or $\bigoplus_{k=1/2}^{n/2} \mathcal{H}_k$ if n is odd where $\mathcal{H}_k$ is the Hilbert space of dimension $2k+1$. Note that each space has multiplicity one. Schur's lemma implies that these spaces are not preserved and since there are no

multiplicity spaces, no quantum information can be transmitted by such a channel. Note however, that the number of classical bits that can be sent is just the logarithm of the number of different total angular momentum states for n qubits:

$$\left\lfloor \log\left(\left\lfloor \frac{n}{2}\right\rfloor + 1\right)\right\rfloor. \tag{4.18}$$

4.5 Quantum Computation with Only Commuting Gates

It is possible to do universal quantum computation using only SWAP and fractional-SWAP operations, as demonstrated by Kempe *et al.* [KBLW01]. In this picture, it is possible to calculate what the cost in reference frame will be, where here, the "reference frame" is any additional systems required to realize an energy conserving implementation beyond the number of required logical qubits. In other words the size of the reference is $n-k$, where n is the number of physical qubits and k is the number of logical qubits. In this case, $k = \log\binom{n}{n/2}$. For each pair of qubits to be either prepared initially or measured finally, a singlet state must be formed, which requires at least one applied Hamiltonian to be formed by the ground state of the input qubits $|00\rangle$. More conventionally, in the circuit model, the circuit:



will prepare the state $\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$, where the gate $T : \alpha|0\rangle + \beta|1\rangle \mapsto \alpha|0\rangle + e^{-i\pi/2}\beta|1\rangle$. Thus, the number of excitations can be conserved in this procedure. In terms of logical qubits, k, the number of physical qubits required as $n \to \infty$ is sub-linear, as can readily be seen using Stirling's Approximation:

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n + O\left(\frac{1}{n}\right). \tag{4.19}$$

This gives $\binom{n}{n/2} = \sqrt{\frac{2}{\pi n}}2^n$

$$\lim_{n\to\infty}\frac{k}{n} = 1 - O\left(\log n/n\right). \tag{4.20}$$

Therefore, as the number of logical qubits desired increases, the overhead per qubit decreases. Also, since one logical qubit can be encoded in three physical qubits under this scheme, there will never be worse than a factor of three increase, therefore,

Fact 3. *For a k qubit computation in a subspace that conserves energy, if suffices to have a reference system of $o(k)$ qubits.*

However, since the Hamiltonians for the square root of SWAP gates and the T gate will need to be switched on and off, which will be done in practice using an external field or similar interaction with a system (see Section 4.9.2), the requirements for an external field to construct the initial entangled pairs and perform the quantum computation can be accounted for. This requirement is philosophically different than a phase reference system requirement, however, and will be further considered in Section 4.9 and subsequent sections.

4.6 Decoherence due to Tracing Out the Reference System

We will now explore explicitly what is anticipated to occur when a coherent state is used to manipulate a two level atomic qubit. This has also been explored with different analysis by van Enk and Kimble [vEK01].

Consider the qubit initially in the state $|0\rangle$ and the reference initially in the state $|\alpha\rangle$, where $|\alpha\rangle$ is a coherent state. A coherent state is selected because it has the property that $a|\alpha\rangle = \alpha|\alpha\rangle$, *i.e.* the state is unchanged by the action of the annihilation operator, which means that losing an excitation should not alter it much.

$$|\psi_{ini}\rangle = |\alpha(0)\rangle \otimes |0\rangle \tag{4.21}$$

$$= \left(e^{-\bar{n}/2} \sum_{n=0}^{\infty} \sqrt{\frac{\bar{n}^n}{n!}} e^{i\phi n} |n\rangle \right) \otimes |0\rangle \tag{4.22}$$

We will entangle the two systems of this product state under the action of the Jaynes-Cummings Hamiltonian in the rotating wave approximation, which is the standard way of modelling atom-field interactions:

$$\hat{H} = \frac{1}{2}\hbar\omega_0(\mathbb{I} - Z) + \hbar\omega_\ell \left(a^\dagger a + \frac{1}{2} \right) + \hbar g \left(\sigma_+ a e^{-i\Delta t} + \sigma_- a^\dagger e^{i\Delta t} \right) \tag{4.23}$$

where ω_0 is the resonant frequency of the atomic transition, ω_ℓ is the cavity's frequency, g is a coupling constant, and $\sigma_\pm$ are the energy raising and lowering operators on the atom. Note that each term commutes with the operator $N_e = a^\dagger a + |1\rangle\langle 1|$, which measures the total number of excitations on both subsystems.

We now split this Hamiltonian into three parts: (1) H_I, the interaction, or "interesting" term, (2) H_{II}, a term which applies a different phase to terms with different total energies, and (3) H_{III}, a phase term which is global, so shall be ignored.

$$H = H_I + H_{II} + H_{III} \tag{4.24}$$

$$H_I = \hbar\Delta a^\dagger a + \hbar g\left(\sigma_+ a + \sigma_- a^\dagger\right) \tag{4.25}$$

$$H_{II} = \hbar\omega_0 \hat{N}_e \tag{4.26}$$

$$H_{III} = \frac{1}{2}\hbar\omega_\ell \mathbb{I} \tag{4.27}$$

where $\Delta \equiv \omega_\ell - \omega_0$. For an on-resonance, energy conserving interaction $\Delta = 0$.

As explained in the introduction, the degree of freedom that will distinguish $|0\rangle$ from $|1\rangle$ in the atom, which is the energy, should have an associated Hamiltonian that conserves the expectation value of this quantity. This Hamiltonian is H_{II}. Any other Hamiltonians that are used to control the system should commute with this Hamiltonian. This is the case. Firstly, $[H_{II}, \mathbb{I}] = 0$ follows trivially. Also, N_e commutes with H_{II}, $[H_{II}, N_e] = 0$, because each term in H independently commutes with N_e:

$$\begin{aligned}
[H_{II}, H_I] &= \left[\hbar\omega_0\hat{N}_e\ ,\ \hbar\Delta a^\dagger a + \hbar g\left(\sigma_+ a + \sigma_- a^\dagger\right)\right] \\
&= \hbar^2\omega_0\Delta\left[a^\dagger a + |1\rangle\langle 1|, a^\dagger a\right] + \hbar^2\omega_0 g\left[a^\dagger a + |1\rangle\langle 1|, \sigma_+ a + \sigma_- a^\dagger\right] \\
&= \hbar^2\omega_0\Delta(0) + \hbar^2\omega_0 g\left(\left[a^\dagger a\ ,\ \sigma_+ a + \sigma_- a^\dagger\right] + \left[|1\rangle\langle 1|\ ,\ \sigma_+ a + \sigma_- a^\dagger\right]\right) \\
&= \hbar^2\omega_0 g\big(\sigma_+[a^\dagger, a]a + a^\dagger\sigma_-[a, a^\dagger] + [|1\rangle\langle 1|, \sigma_+]a + [|1\rangle\langle 1|, \sigma_-]a^\dagger\big) \\
&= \hbar^2\omega_0 g(-\sigma_+ a + a^\dagger\sigma_- + \sigma_+ a - \sigma_- a^\dagger) \\
&= 0.
\end{aligned}$$

Now, consider the evolution that occurs on the reference state and the qubit under the action of the Hamiltonian $H_I + H_{II}$ (H_{III} is neglected because it only

contributes a global phase):

$$
\begin{aligned}
|\psi_{ini}\rangle \xrightarrow{H_I+H_{II}} & \left(e^{-\bar{n}/2}\sum_{n=0}^{\infty}\sqrt{\frac{\bar{n}^n}{n!}}e^{i\phi n}e^{i\omega_0 tn}\cos(g\sqrt{n}t)\,|n\rangle\right)|0\rangle \\
& + \left(e^{-\bar{n}/2}\sum_{n=1}^{\infty}\sqrt{\frac{\bar{n}^n}{n!}}e^{i\phi n}e^{i\omega_0 tn}i\sin(g\sqrt{n}t)\,|n-1\rangle\right)|1\rangle \\
= & \left(e^{-\bar{n}/2}\sum_{n=0}^{\infty}\sqrt{\frac{\bar{n}^n}{n!}}e^{i\phi n}e^{i\omega_0 tn}\cos(gt\sqrt{n})\,|n\rangle\right)|0\rangle \\
& + \left(e^{-\bar{n}/2}\sum_{n'+1=1}^{\infty}\sqrt{\frac{\bar{n}^{(n'+1)}}{(n'+1)!}}e^{i\phi}e^{i\omega_0 t}e^{i\phi n'}e^{i\omega_0 tn'}\sin(gt\sqrt{(n'+1)})\,|n'\rangle\right)|1\rangle \\
= & \left(e^{-\bar{n}/2}\sum_{n=0}^{\infty}\sqrt{\frac{\bar{n}^n}{n!}}e^{i\phi n}e^{i\omega_0 tn}\cos(gt\sqrt{n})\,|n\rangle\right)|0\rangle \\
& + \left(e^{-\bar{n}/2}\sum_{n'=0}^{\infty}\sqrt{\frac{\bar{n}^{n'}}{n'!}}\sqrt{\frac{\bar{n}}{n+1}}e^{i\phi}e^{i\omega_0 t}e^{i\phi n'}e^{i\omega_0 tn'}\sin(gt\sqrt{(n'+1)})\,|n'\rangle\right)|1\rangle \\
= & \, e^{-\bar{n}/2}\sum_{n=0}^{\infty}\sqrt{\frac{\bar{n}^n}{n!}}e^{i\phi n}e^{i\omega_0 tn}\,|n\rangle \\
& \otimes\left(\cos(gt\sqrt{n})\,|0\rangle + i\sqrt{\frac{\bar{n}}{n+1}}e^{i\phi}e^{i\omega_0 t}\sin(gt\sqrt{n+1})\,|1\rangle\right)
\end{aligned}
$$

Notice that this state is still slightly entangled, because of the $\sqrt{\frac{\bar{n}}{n+1}}$ factor on the $|1\rangle$ term and the dependency of the sine and cosine factors on n. This prevents us from extricating the sum over n from the qubit state and forming a tensor product state between the two systems.

In the case that $n \to \infty$, then we have $\frac{n+1}{n} \to 1$ and the only terms that contribute significantly to the sum are the ones for which $e^{-\bar{n}/2}\sqrt{\frac{\bar{n}^n}{n!}}$ is large, that is the states where $n \approx \bar{n}$. That will mean that the factor $\sqrt{\frac{\bar{n}}{n+1}} \approx 1$, and by choosing the interaction time $t = \frac{1}{g\sqrt{\bar{n}}}\theta$, we can remove the dependence on n from the qubit subsystem and only incur an error of size $O(\frac{1}{\bar{n}})$. In this case we can approximate the state as a tensor product:

$$
\alpha\left(\frac{\theta}{g\sqrt{\bar{n}}}\right)\otimes\left(\cos(\theta)\,|0\rangle + ie^{i\phi}e^{i\omega_0\theta/g\bar{n}}\sin(\theta)\,|1\rangle\right) \tag{4.28}
$$

For finite $\bar{n}$, let us consider the density matrix of the qubit state. For $t = \frac{1}{g\sqrt{\bar{n}}}\theta$:

$$\begin{aligned}\rho_q' &= e^{-\bar{n}}\sum_{n=0}^{\infty}\frac{\bar{n}^n}{n!}\Bigg(\cos^2(\sqrt{\frac{n}{\bar{n}}}\theta)\,|0\rangle\langle 0| \\ &\quad -ie^{-i\phi}e^{-i\omega_0 t}\sqrt{\frac{\bar{n}}{n+1}}\cos(\sqrt{\frac{n}{\bar{n}}}\theta)\sin(\sqrt{\frac{n+1}{\bar{n}}}\theta)\,|0\rangle\langle 1| \\ &\quad +ie^{i\phi}e^{i\omega_0 t}\sqrt{\frac{\bar{n}}{n+1}}\cos(\sqrt{\frac{n}{\bar{n}}}\theta)\sin(\sqrt{\frac{n+1}{\bar{n}}}\theta)\,|1\rangle\langle 0| \\ &\quad +\frac{\bar{n}}{n+1}\sin^2(\sqrt{\frac{n+1}{\bar{n}}}\theta)\,|1\rangle\langle 1|\Bigg) \end{aligned} \tag{4.29}$$

Consider the fidelity between this state and the ideal final state ρ_q:

$$F(\rho_q,\rho_q') = \mathrm{Tr}\left(\sqrt{\sqrt{\rho_q}\rho_q'\sqrt{\rho_q}}\right) \tag{4.30}$$

because ρ_q is pure, we can write this as:

$$F(\rho_q,\rho_q') = \sqrt{\langle\psi_q|\,\rho_q'\,|\psi_q\rangle} \tag{4.31}$$

Then, noting that $\cos(\theta(1-\delta)) = \cos(\theta)+O(\delta)$ and $\sin(\theta(1-\delta)) = \sin(\theta)+O(\delta)$ for $-1<\delta<1$, the fidelity of the actual density matrix of the qubit for finite $\bar{n}$ with the ideal pure final state is

$$F = \left[e^{-\bar{n}}\sum_{n=0}^{\infty}\frac{\bar{n}^n}{n!}\left(\cos^4(\theta)+2\cos^2(\theta)\sin^2(\theta)+\sin^4(\theta)+O\left(1-\frac{n+1}{\bar{n}}\right)\right)\right]^{1/2}. \tag{4.32}$$

Evaluating the sum

$$e^{-\bar{n}}\sum_{n=0}^{\infty}\frac{\bar{n}^n}{n!}O\left(1-\frac{n+1}{\bar{n}}\right) = O\left(\frac{1}{\bar{n}}\right) \tag{4.33}$$

recovers

$$F = \left[1-O\left(\frac{1}{\bar{n}}\right)\right]^{1/2} \in 1-O\left(\frac{1}{\bar{n}}\right), \tag{4.34}$$

using Bernoulli's inequality, and the channel that describes this transformation is a depolarizing channel.

In addition, note that the optimal phase estimation state, a phase squeezed state of the form [CW05] $|\psi\rangle = \frac{1}{\sqrt{(N+1)/2}}\sum_{n=0}^{N-1}\sin\left(\frac{(n+1)\pi}{N+1}\right)$ where the dimension of the

reference Hilbert space is not infinite, but is N, will only be a linear improvement on this, so the error will still be $\epsilon \in O\left(\frac{1}{\bar{n}}\right)$.

4.6.1 Error Scaling

Consider once more the three proposals for reference frame management that involved coherent states. The first of the three was to use coherent states as references, each paired with a qubit. In contrast to the second proposal, which was to use a Fock state $|n\rangle$, as seen in the last section, interaction with a coherent state causes decoherence on the qubit depending on the mean number of excitations, $\bar{n}$, in the coherent state as $O\left(\frac{1}{\bar{n}}\right)$. The advantage is that the coherent state is stable under photon loss: it is an eigenstate of the annihilation operator. Further, even though there is error due to the uncertain number of excitations, the mean number of total excitations does not change, since the coherent state exchanges only with its one qubit and all two qubit operations are assumed to be controlled-Z gates, so the error due to mis-timing the pulse ($t = \frac{\theta}{g\sqrt{\bar{n}}}$) does not get worse with repeated use of each reference.

For the scheme that use the same single coherent state repeatedly for all gates, Gea-Banacloche and Ozawa in [GBO06] suggest that the size of the reference will need to increase as the square of the number of times it will be used. It will be used each time an operation that does not commute with the total energy operator needs to be performed. This was demonstrated by the case of a coherent state reference being used repeatedly to create a large cat state $\frac{1}{\sqrt{2}}\left(|0\rangle^{\otimes \ell} + |1\rangle^{\otimes \ell}\right)$. In order to create this state from the state $|0\rangle^{\otimes \ell}$, the reference will need to enter a superposition with the qubits,

$$\frac{1}{\sqrt{2}}\left(|0\rangle^{\otimes \ell}\left|\sqrt{\bar{n}}e^{i\phi}\right\rangle + |1\rangle^{\otimes \ell}\left|\sqrt{\overline{(n-\ell)}}e^{i\phi}\right\rangle\right) \tag{4.35}$$

with each term having a different expected number of excitations. In this case, the mistiming errors will grow: for subsequent operations $t = \frac{\theta}{g\sqrt{\bar{n}-\ell/2}}$ could be chosen, but in the case of a large cat state, this will induce incorrect rotations, and tracing out the reference immediately also causes decoherence between the states. The error goes as $O\left(\frac{\ell^2}{\bar{n}}\right)$, where in the worst case, ℓ could be all of the qubits involved in the calculation.

This type of resource scaling would negate the benefits achieved by the square root improvement of many quantum algorithms, including Grover's search algo-

rithm [Gro96].

The last proposal was for the computation to be carried out directly as according to the circuit model, using a series of phase-correlated coherent states as the reference system. This reflects what is currently done in practice for implementations involving atoms manipulated by laser fields. A similar scaling to the proposed by Gea-Banacloche and Ozawa is expected naively. Diverging from their argument, if there is a coherent state containing a mean number of excitations $\bar{n}$ it can be divided by an appropriately chosen beamsplitter operation (which creates phase correlated states) into ℓ coherent states in separate modes, each containing a mean number of photons $\frac{\bar{n}}{\ell}$. Then the error on each gate will be $O(\frac{\ell}{\bar{n}})$. In the worst case, these errors can add, so then the total error $\epsilon_{total} = O(\frac{\ell^2}{\bar{n}})$. Fortunately, this is the scaling that is to be expected without active error correction. As discussed in detail in Section 4.7, active error correction can reduce this additive error below any constant threshold of $O(\epsilon)$, where ϵ is the expected error on a step of error correction, a gate, and another active error correction step together.

Therefore in this set up, the required reference size will scale as the number of gates in the circuit, G, times the required reference to do a single gate with accuracy ϵ, therefore $O(G \times \bar{n})$. For problems with polynomial sized circuits, G must not increase with input size faster than as a polynomial function of input size, k.

4.7 Error Correction

It has been suggested [KGBO09, GB02] that the finite laser coherence time could pose a problem for maintaining a phase reference for quantum computation in a system where the qubits are atoms manipulated by laser pulses.

Consider a quantum computer which is implemented on a system of atoms manipulated with a series of coherent states of the electromagnetic field. The phase of the coherent states will provide a reference frame for the computation. Over time the phase of the successive coherent states will drift with the net result that the definition of the reference frame drifts during the course of the computation. This implies that a Hadamard gate, *i.e.*, a $\frac{\pi}{2}$ gate, applied early in the computation will not be identical to one applied later. If one laser is used to control all of the atoms, which serve as the qubits, this is equivalent to a global drift of the relative phase on all qubits. Under these conditions, the deterioration of the reference, if sufficiently slow, can be addressed using error correction techniques.

There are two sources of noise in this situation. The first is a depolarizing effect due to the coherent state having a mean number of excitations which is finite. The second is the dephasing effect of the drifting reference frame.

The first effect occurs when the coherent state of the laser is treated as a quantum system, which is disturbed by its interaction with the qubit during an operation, as was analyzed in Section 4.6. Commonly, the laser is treated as a classical system, so that the operations it performs on the qubit will be unitary. Treating the laser quantum mechanically causes an operation on the qubit to be decohering, so it must be described by a more general map.

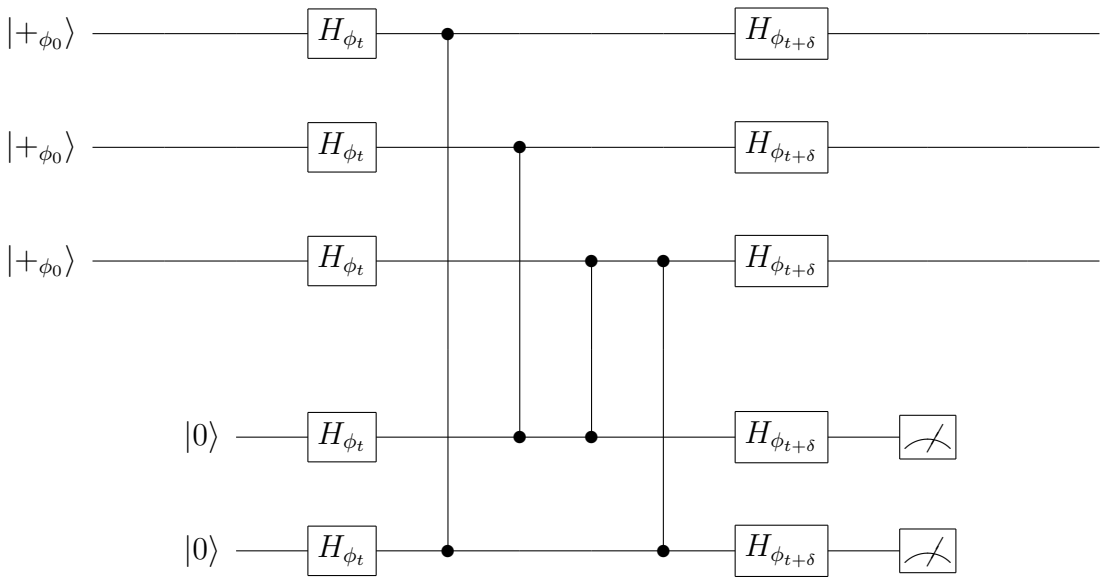
The second effect is caused by the imperfect phase correlation of the series of states emitted by the laser. Imagine that the state $|\alpha\rangle = \left|\bar{n}e^{i\phi}\right\rangle$ is emitted by the laser in one time step and then a state with a slightly different phase $\left|\bar{n}e^{i(\phi+\delta)}\right\rangle$. This will lead to the phase of the reference being "out of synch" with the phase of the qubit.

The dephasing due to the phase drift of the laser can be managed using a quantum error correcting code. It can be corrected for together with the depolarizing effect studied in the previous section using a code which corrects both bit and phase errors, such as Shor's nine-qubit code, the seven-qubit Steane code, or the five-qubit code, which can correct one X, Y, or Z error.

Alternatively, to allow for fewer error correction operations and higher thresholds, the depolarizing effect can be countered using the five-qubit error correcting code or the seven-qubit Steane code [Ste96], which also corrects one error and has a proven threshold of 2.73×10^{-5} [AGP06, Rei06], while the dephasing effect can be countered using the three-qubit phase error correcting code. Estimated and simulated error thresholds for the seven-qubit code can be much higher, ranging between $10^{-3} - 10^{-7}$ [Zal98, Pre98, KLZ98], with the higher estimates being more recent. A simulation gave the threshold for the three-qubit code as 1.2×10^{-3} for memory errors and 2.1×10^{-2} for gate errors [CS05]. Here, only noise due to the reference effect is considered so memory noise will be free of the depolarizing effect, but it will include the dephasing, since the phase of the laser is assumed to drift continuously, regardless of whether an operation is occurring on the qubit. The dephasing will also occur during every gate operation.

In most realistic arrangements, the average number of photons in a laser pulse will be large, so the depolarization effect will be small; smaller than the dephasing effect. However, it seems that the threshold for the three-qubit code is $10 - 100$ times larger for the three-qubit code that will correct the dephasing.

The following circuit could be used to correct the phase drift:



This makes use of controlled-Z gates, which commute with the constant Hamiltonian because they do not change the weight of the energy eigenstates only the phase. The Hadamard gates are performed using the coherent states, which are phase references. It is assumed that the state $|0\rangle$ can be prepared without the reference.

The encoding of the logical state chosen is $|\psi_L\rangle = \alpha\,|+_L\rangle + \beta\,|-_L\rangle = \alpha\,|+++\rangle + \beta\,|---\rangle$. For simplicity, assume that the state $|+++\rangle$ was prepared at time $t = 0$ and at some later time t the correction map will take place and that the drift that takes place during the correction is negligable. During this time, the phase of the coherent states will have drifted from ϕ_0 to ϕ_t. Let $\delta = \phi_t - \phi_0$ and let $H_{\phi_t}\,|0\rangle \to |+\rangle$ and $|+_t\rangle = |0\rangle + e^{-i\delta}\,|1\rangle$. Then $H_{\phi_t}\,|+_t\rangle \to \frac{1+e^{-i\delta}}{2}\,|0\rangle + \frac{1-e^{-i\delta}}{2}\,|1\rangle$, so after the H_{ϕ_t} operations the state in the circuit is:

$$\begin{aligned}
&\left(\frac{1+e^{-i\delta}}{2}\right)^3 |000\rangle\,|{+}{+}\rangle \\
&+ \left(\frac{1+e^{-i\delta}}{2}\right)^2 \left(\frac{1-e^{-i\delta}}{2}\right) (|001\rangle\,|{+}{-}\rangle + |010\rangle\,|{-}{-}\rangle + |100\rangle\,|{-}{+}\rangle) \\
&+ \left(\frac{1+e^{-i\delta}}{2}\right) \left(\frac{1-e^{-i\delta}}{2}\right)^2 (|011\rangle\,|{-}{+}\rangle + |101\rangle\,|{-}{-}\rangle + |110\rangle\,|{+}{-}\rangle) \\
&+ \left(\frac{1-e^{-i\delta}}{2}\right)^3 |111\rangle\,|{+}{+}\rangle\,. \qquad (4.36)
\end{aligned}$$

The probability that the state $|+_t\rangle$ is interpreted as $|-\rangle$ is the inner product between these states is:

$$\begin{aligned} p &= \left| \frac{1}{\sqrt{2}} \langle - | \left(|0\rangle + e^{-i\delta} |1\rangle \right) \right|^2 \\ &= \langle 0|0\rangle - e^{-i\delta} \langle 1|1\rangle \\ &= \left| \frac{1}{2}(1 - e^{-i\delta}) \right|^2 \\ &= \frac{1}{2} - \frac{1}{2}\cos(\delta). \end{aligned}$$

And the probability that the state is not corrected by the circuit is

$$\begin{aligned} p' &= 3 \left| \left(\frac{1 - e^{-i\delta}}{2} \right)^2 \left(\frac{1 + e^{-i\delta}}{2} \right) \right|^2 + \left| \left(\frac{1 - e^{-i\delta}}{2} \right)^3 \right|^2 \\ &= 3p^2(1-p) + p^3 \\ &= 3p^2 - 2p^3. \end{aligned}$$

So, as long as $3p^2 - 2p^3 < p$ this is a reduction in the error. In this case, in line with the threshold theorem [AGP06], this is ensured for $p < \frac{1}{2}$

Note also that logical qubits do not have to share a reference for this to work, the same reduction is true for independent phase references on each physical qubit, since again, the only gates that will be used to couple qubits are control-Z gates.

Note that if the drift in the laser is one in which a coherent state $|\alpha\rangle = \left|\bar{n}e^{i\phi}\right\rangle$ is drifting to a mixture over states of mean photon number $\bar{n}$ with different phases, this is still corrected for by the procedure described above. Suppose the state of the laser drifts from $\left|\bar{n}e^{i\phi}\right\rangle$ to the mixture $\frac{1}{2}\left|\bar{n}e^{i(\phi+\delta)}\right\rangle\left\langle\bar{n}e^{i(\phi+\delta)}\right| + \frac{1}{2}\left|\bar{n}e^{i(\phi-\delta)}\right\rangle\left\langle\bar{n}e^{i(\phi-\delta)}\right|$. Then after the correction the computation qubits will be in an entangled state with the reference. If the reference were to be traced out, then the qubits would move to the mixed state $\frac{1}{2}|+_\delta +_\delta +_\delta\rangle\langle +_\delta +_\delta +_\delta| + \frac{1}{2}|-_\delta -_\delta -_\delta\rangle\langle -_\delta -_\delta -_\delta|$. Crucially, the reference will not be discarded. It must continue to be used until at the end, the qubits are all returned to the computational basis.

In fact, it does not hurt the operation of such a quantum computer if the reference state starts in a mixed state over phase $\rho = \int_\phi P(\bar{n}) \left|\bar{n}e^{i\phi}\right\rangle\left\langle\bar{n}e^{i\phi}\right| d\phi$ so long as the same reference is used throughout the computation, from start to finish. The caveat is that the qubits must be initially in (or have very high fidelity to) a classical input bit string state for which each qubit is in one of the two energy

eigenstates $|0\rangle$ or $|1\rangle$ and must finally be measured in that same basis. It does not suffice to start with an input quantum state $|\psi\rangle$ prepared according to some other phase reference convention or to output some final quantum states $|\psi'\rangle$ which would be used later with another phase reference. These operations will decohere the states $|\psi\rangle$ and $|\psi'\rangle$ as the phase reference is lost. A reference of sufficient size $\bar{n}$ for the computation needs to be kept from the beginning of a computation, where the classical input specifies the instance of the problem to be solved, until the final measurement of the classical solution.

Error correction can also be completed in the other two bases by suitably adapting the circuit drawn above. This will similarly reduce the depolarizing error due to the uncertain number of excitations in the field. This means that like any other expected source of error in quantum information processing, reference errors can be suppressed by error correction.

The threshold theorem gives a bound on the required resources for implementing such a scheme. In the case a distance-three code, for example, [AGP06]

Theorem 4.1 (of Aliferis, Gottesman, and Preskill)**.** *Quantum accuracy threshold for independent stochastic noise. Suppose that fault-tolerant gadgets can be constructed such that all 1-exRecs obey the property exRec-Cor, and such that ℓ is the maximal number of locations in a 1-Rec, d is the maximal depth of a 1-Rec, and ϵ_{th}^{-1} is the maximal number of pairs of locations in a 1-exRec. Suppose that independent stochastic faults occur with probability $\epsilon < \epsilon_{th}$ at each location in a noisy quantum circuit. Then for any fixed δ, any ideal circuit with L locations and depth D can be simulated with error δ or better by a noisy circuit with L^* locations and depth D^*, where*

$$L^* = O(L(\log L)^{\log \ell}), \qquad D^* = O(D(\log L)^{\log d}). \tag{4.37}$$

Notice that the overhead in operations is only a logarithmic increase, raised to a power which will be a constant for a particular implementation, and for this increase in operations and depth, the error can be suppressed below a constant chosen by the user.

To allow for the possibility of coherent errors in such a system we can work in terms of error amplitudes $\sqrt{\epsilon}$. In terms of what can be achieved using concatenated codes, a number of layers of concatenated encoding, m, is required to reduce an error amplitude $\sqrt{\epsilon_g}$ where ϵ_g is the error on one gate down to $\frac{\sqrt{\epsilon}}{G}$ for a computation that will require G logical gates. For convenience let us chose $\epsilon = \epsilon_g$, though we can choose a much smaller threshold if we like, provided that the threshold condition

is met. This choice will keep the error constant through the computation. Then, if the threshold for the code being used is ϵ_{th}, m must be chosen so that:

$$\begin{aligned}
\frac{1}{\epsilon_{\text{th}}}\left(\frac{1}{\epsilon_{\text{th}}}\sqrt{\epsilon}\right)^{2^m} &< \frac{\sqrt{\epsilon}}{G} \\
2^m\left(\log\frac{1}{\epsilon_{\text{th}}}+\frac{1}{2}\log\epsilon\right)+\log\frac{1}{\epsilon_{\text{th}}} &< \frac{1}{2}\log\epsilon+\log\frac{1}{G} \\
2^m\left(\log\epsilon_{\text{th}}+\frac{1}{2}\log\frac{1}{\epsilon}\right) &> \frac{1}{2}\log\frac{1}{\epsilon}+\log G \\
m &> c'\log\log G + c
\end{aligned}$$

where in the final line comes from noticing that ϵ_{th} and ϵ are constants (ϵ_{th} is determined by the code and ϵ is chosen by the user to match the individual gate error in this analysis). Then requiring some number of qubits q per encoding, then in total for m-levels, the number of required gates is $Gq^m = G\text{poly }\log G$.

Now the required energy goes as $O\left(\frac{G\text{poly }(\log G)}{\epsilon}\right)$. This is an improvement over the previous scaling that went with G^2 in the asymptotic limit. In order to achieve this however, we had the restriction that the gates must have sufficiently low error that the error per gate $\epsilon < \epsilon_{\text{th}}^2$. This will mean that to ensure this scaling a reference of size $O\left(\frac{1}{\epsilon_{\text{th}}^2}\right)$ will need to be used for each gate. However for a given code, ϵ_{th} is a constant: once it is achieved the scaling as $O\left(\frac{G\text{poly }(\log G)}{\epsilon}\right)$ will take over. This converts a requirement that the error be lower per gate than some threshold, to an energy must exceed some threshold per gate $E > E_{\text{th}}$.

What this implies is that it is not necessary to increase the reference size with the square of the number of the operations in order to keep the error below some threshold. This should hold for all three of the schemes which show error due to reference effects, however the codes for which each is optimized will likely be very different. The critical point is that this scaling does not eliminate the apparent quantum speed up of algorithms that show a square root improvement in run time over classical algorithms in the asymptotic limit. Although, without error correction, this would be the behaviour of these effects.

4.8 Modeling Phase Reference Drift

There will be drift in a phase reference source based on a laser-type device that is treated quantum mechanically for information theoretic reasons, including the no

cloning theorem for quantum states. Consider the laser cavity to contain a very large quantum state. Quantum subsystems are emitted from the cavity in a series of time-bin modes. As the energy from the cavity is being emitted as a laser beam, the energy is being resupplied from an optical pump field.

Lasers operate by creating a population inversion in the gain medium. That means that within the laser cavity are atoms or molecules, most of which are in the excited state of the lasing transition. This state of affairs is maintained though constant pumping of energy from another optical source into the cavity. That source is not coherent with the laser. It serves only to excite the particles in the gain medium. The process by which the coherent laser beam is established is stimulated emission. A photon interacts with an atom in the excited state and causes it to emit a second photon with the same phase as it decays to the ground state. This reinforces the photon beam, creating more coherent excitations. However, spontaneous emission also occurs in the cavity. In this process the excited atom decays without interacting with a photon and will have a random phase.

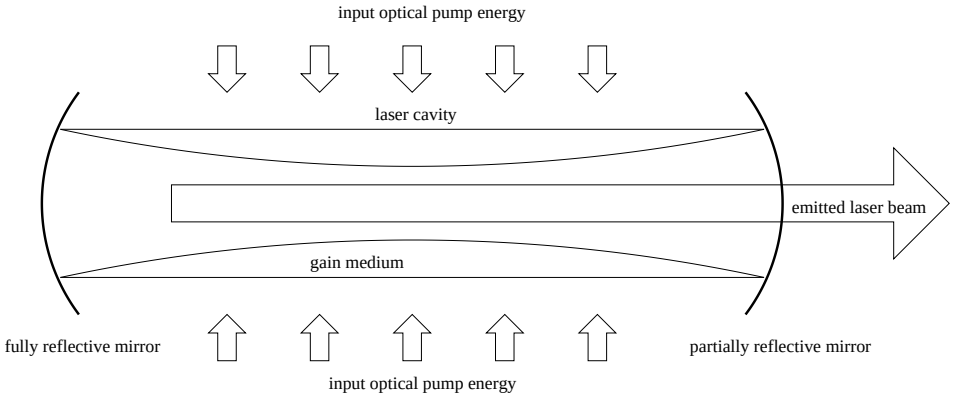


Figure 4.2: Schematic of a laser.

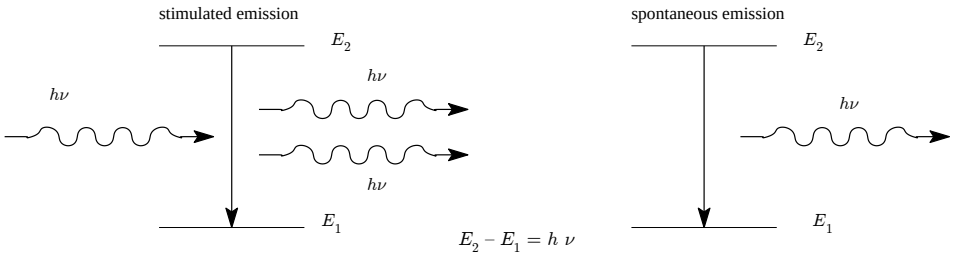


Figure 4.3: Stimulated and spontaneous emission from an excited two-level system.

However, this process does not and cannot perfectly preserve the state in the

cavity. We can consider, however, a model in which a number, N, of identical Gaussian field states are input to an optimal quantum cloning device, which produces $N+1$ identical states, each as close as possible to the input states. This type of situation was studied by Cerf and Iblisdir [CI00]. They found that the states after this procedure should be described by the expression

$$\rho_1^{(\ell)} = Tr_N(\rho_{N+1}) \tag{4.38}$$

$$= \frac{1}{2\pi\sigma^2_{N,N-1}} \int e^{-\frac{|\beta|^2}{\sigma^2_{N,N+1}}} D(\beta)\rho_1^{(\ell-1)}D(\beta)d^2\beta \tag{4.39}$$

where the integral is over all β in the complex plane and $D(\beta)$ is the displacement operator. The Gaussian noise this introduces leads to an increase in the variance of x, p, and n by $\sigma^2_{N,N+1}$. They find a lower bound for $\sigma^2_{N,M} = \frac{M-N}{MN}$ which in the case we are interested in gives $\sigma^2_{N,N+1} = \frac{1}{N(N+1)}$.

This increases the variance on the phase also, so that the error on the qubit operation is also increased from $O\left(\frac{1}{\bar{n}}\right)$ to $O\left(\frac{1}{\bar{n}}\left(1+\frac{\ell}{N(N+1)}\right)\right)$, where ℓ is the number of times the copying has occured. In the worst case the errors on each qubit can add, so in total $\epsilon = O\left(\frac{\ell}{\bar{n}}\left(1+\frac{\ell}{N(N+1)}\right)\right)$. For a fixed $\bar{n}$, to achieve an error ϵ on the ℓth gate, $N = O\left(\frac{\ell}{\sqrt{\epsilon}}\right)$.

4.9 Gates that Commute with the Total Energy

Even when the operation induced by the new Hamiltonian commutes with the total energy operator, switching on different Hamiltonians requires energy in the form of an applied field which will break the energy symmetry between two states. This is a separate issue from the need to maintain a reference frame. Phase references are required to keep coherence between different states in a superposition by providing a measurement convention (the way a coherent state's phase gives a definition of the x-axis), or by being the purification of the computation qubits. This is not required for gates that commute with the energy. However, energy must still be supplied to the qubits during the operation

For example, consider the controlled-Z gate. Symmetry under the Schrödinger equation must be broken between the state $|11\rangle$ and the states $|00\rangle, |01\rangle,$ and $|10\rangle$, so that there is a relative phase acquired by $|11\rangle$:

$$H = E_1 |11\rangle\langle 11| + E_0\big(|00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 10|\big) \tag{4.40}$$

Applying such a Hamiltonian for a time $t = \frac{\pi}{(E_1 - E_0)}$ will give a phase flip on the state $|11\rangle\langle 11|$ since the phase factor applied to $|11\rangle\langle 11|$ will be e^{iHt} and we require the relative phase to be $(E_1 - E_0)t = \pi$.

Following Gea-Banacloche [GB02], consider again a Jaynes-Cummings model of a two level system interacting with a field.

$$\hat{H} = \frac{1}{2}\hbar\omega_0(\mathbb{I} - c - Z) + \hbar\omega_\ell\left(a^\dagger a + \frac{1}{2}\right) + \hbar g\left(\sigma_+ a e^{-i\Delta t} + \sigma_- a^\dagger e^{i\Delta t}\right)(\mathbb{I} - c - Z) \quad (4.41)$$

Now $\omega_0 = 0$. The coupling Hamiltonian for a single mode of the form:

$$H_I = g(ae^{-i\omega t} + a^\dagger e^{i\omega t})\,|11\rangle\langle 11|\,. \quad (4.42)$$

If the field being coupled to is a coherent state, there will be uncertainty in the excitation number. For such a field we want:

$$\pi = \sqrt{\bar{n}}g\int_0^\tau (e^{-i\omega t + i\phi} + e^{i\omega t - i\phi})\mathrm{d}t, \quad (4.43)$$

but since there is a distribution over n, there will be fluctuations around this value of the exponent. They must be small for this procedure to work with high fidelity. Evaluating the integral

$$\pi = \sqrt{\bar{n}}g\frac{\sin(\omega\tau - \phi)}{\omega}, \quad (4.44)$$

we see that it is desirable for $g\frac{\sin(\omega\tau-\phi)}{\omega}$ to be small because we need

$$\left(\sqrt{\bar{n}} - \sqrt{\bar{n} - \sqrt{\bar{n}}}\right) g\frac{\sin(\omega\tau - \phi)}{\omega} \approx \frac{1}{2}g\frac{\sin(\omega\tau - \phi)}{\omega}, \quad (4.45)$$

to be small (where $\sqrt{\bar{n}}$ is the standard deviation of the Poisson distribution over n). So choose g and ω such that

$$\left|g\frac{\sin(\omega\tau - \phi)}{\omega}\right|^2 < \epsilon \quad (4.46)$$

Then, for equation (4.44) to be true, that means that $\bar{n} = \frac{\pi^2}{\epsilon}$. In general this may be done over a many modes of different frequencies, ω_k, in which case it is the total photon number that must be this large. This means that the photon number of the field must go as the inverse of the desired error, or, for some chosen error ϵ, the number of excitation is $O(\frac{1}{\epsilon})$.

This is a linear trade-off between energy needed and time taken to accomplish

a gate if frequency is fixed. The error in this process can also be maintained below a threshold of $O(\epsilon)$ using error correction.

Fact 4. *The energy required for a quantum computation consisting only of operations that commute with the total energy operator for a fixed mean frequency should scale linearly in the number of gates required to execute the computation, and should be scale with $\frac{1}{\epsilon}$, where ϵ is the error in one gate.*

The energy required will scale with the square of number of gates required and as one over the error in the operations, $E = O\left(\frac{G^2}{\epsilon}\right)$, if each time the field is generated it is discarded. However, in principle, there is no reason in principle why this energy could not be recycled from one gate to the next. The coherent state field is disturbed by the interaction with the qubits, but it does not need to be kept coherent. The energy from field in principle could be returned to a battery and used for a later gate with only minimal loss. In such a case, $E = O(\frac{G}{\epsilon})$ energy will be necessary. In practice, however, this could be quite difficult.

4.9.1 A Stricter Energy Conservation Restriction: Quantum Computation with a Single Time-Invariant Hamiltonian

There are strategies for implementing universal quantum computations that conserve energy automatically. In the strictest sense, Noether's Theorem implies that energy conservation within a system is a time invariance of the Hamiltonian governing the evolution of that system. It is possible to achieve a quantum computation under a fixed Hamiltonian, as is illustrated, for example in a proposal for programmable quantum cellular automata by Nagaj and Wocjan [NW08]. In their scheme a portion of the inputs encodes the program and a portion encodes the particular instance to be solved (the traditional input of a problem) and a constant Hamiltonian is applied. After an amount of time that depends on the problem, the output of the computation is produced with high probability and is flagged as completed by a single bit value that is measured. Of course in this case, in order to complete a non-trivial calculation, the input state cannot be an eigenstate of the constant Hamiltonian.

In particular, in this scheme a chain of ten-dimensional qudits is used, which is thought of as being decomposed into a tensor product of a program register on a five dimensional Hilbert space and and data register on a two dimensional space.

The initial state can be prepared in the computational basis (that is the initial state for each qudit can be one element along the diagonal of the 10×10 matrix), and the measurement of the flag qubit to check the completion of the program, which occurs with high probability after a polynomial amount of time in the size of the equivalent quantum circuit for the computation, can also be in this basis. This is the basis that we have been calling the basis of well-defined energy up to this point. However, for the duration of the quantum calculation, the definition of energy will be according to the eigenstates of the fixed Hamiltonian. The Hamiltonian suggested is a sum of translationally invariant terms,

$$H_{10} = \sum_j (R_{j,j+1} + R_{j,j+1}^\dagger) \tag{4.47}$$

where

$$R_{j,j+1} = \sum_{A \in \{W,S,\mathbb{I}\}} |A,\cdot\rangle\langle\cdot,A|_{p_j,p_{j+1}} \otimes \mathbb{I}_{d_j,d_{j+1}} + |A,\triangleright\rangle\langle\triangleright,A|_{p_j,p_{j+1}} \otimes A_{d_j,d_{j+1}} \tag{4.48}$$

where A is one of three two-qubit operations: controlled-Hadamard, SWAP, or the identity, and the program registers p_j, have five orthogonal states which denote the three operations, $\cdot$, which is just signaling "do nothing", and a pointer state $\triangleright$ which indicates that the operation in register p_{j+1} should be applied to the data qubits d_j and d_{j+i}.

This means the preparation will not be energy conserving, under the definition of energy supplied by this Hamiltonian. This is an example of a problem that has a quantum input and a quantum output, but that has no fundamental error due to the energy conservation during the computation itself. However, the preparation of the initial quantum input and the measurement of the flag qubit and (if required) of the final quantum output will suffer from this fundamental error due to the finite size of the preparation and measurement apparatus.

Finding a preparation for the initial state in this scheme that even commutes with the constant Hamiltonian used for the computation would be quite difficult, since the Hamiltonian can entangle all neighboring qubits.

A very similar scheme was proposed by Chase and Landahl [CL08] that uses a chain of eight-dimensional systems, but the Hamiltonian for that case is not translationally invariant. Nevertheless, it demonstrates that constant Hamiltonians can be universal on even smaller systems.

4.9.2 Hamiltonians and Time in a Quantum Computation

If a state is prepared as an eigenstate of a Hamiltonian, evolved under that Hamiltonian, and measured according to that Hamiltonian, no change will be observed in the state. Using a sequence of different Hamiltonians, as in the circuit model, even if they all commute with the total energy operator, breaks the symmetry of the static situation of the preceding sentence and give an arrow for time in the computation, in the sense that the quantum computation system becomes correlated to the classical observer's clock. Notice that in the scenario presented in Section 4.9.1 of a single Hamiltonian which is universal for computation, the arrow for time does not come from the Hamiltonian itself: if it were applied in discrete intervals it is clear that at each step it could undo the previous step, since it is composed of an operation and its Hermitian conjugate. In that case, the asymmetry for time is supplied by the initial input state which contains, after the programmed operations desired for the calculation, many identity operations, which correspond to a large amplitude after a certain time for the computation to be finished. In the standard circuit model the changing series of Hamiltonians supply the asymmetry with respect to time in order to shepherd the initial state to the final one.

4.10 Conclusion

Reference frames are a critical resource for accomplishing state transformations and measurements in quantum information and there has been concern that this resource has not properly been accounted for when assessing the potential of quantum computers.

This work demonstrates that even in the circuit model style implementation of quantum computation, where correlated phase references are employed and discarded, the overhead required in the reference frame will not need to be more than $O(\frac{G\text{poly}\log G}{\epsilon})$, where G is the number of gates required to be implemented, so G is polynomial in the input size to the problem for a problem, where ϵ is the error tolerated in the computation, assuming active error correction is employed. This is a resource requirement which is polynomial in the input size of the problem.

However, in principle, the reference requirement can be reduced further than this to be as small as $o(k)$ qubits, where k the number of logical qubits required for the computation, and requiring $\frac{k}{2}$ excitations distributed amongst all the qubits. This is the case if a subspace that conserves the excitation number is used and

active error correction is employed during preparations and measurements. Note that there is no fundamental error here due to the reference requirements. This is very efficient in principle, but could cause implementation difficulties since it requires many extra qubits be kept coherent when k is small. As k increases, this becomes a less taxing restriction. Also, there are certain to be errors from other sources that arise when implementing this scheme.

If a collection of atoms are paired with references that are Fock states $|n\rangle$ trapped in cavities, this is more robust against photon loss, but also requires additional excitations: for q physical qubits, q cavities are required and nq excitations are needed. There is no error due to reference frame requirements in this scheme either, but in practice maintaining many separate cavities, preparing perfect Fock states, and executing perfect timing will not be possible, so there will be many errors due to these other factors.

If the two-level atoms are paired with coherent state references that are trapped in cavities, this is even more robust against photon loss, since the coherent state is an eigenvector of the annihilation operator, but in this case, the timing of the interaction cannot be chosen to perform the desired operations free of error. The terms in the superposition of Fock states that make up the coherent state will each couple differently to the qubit and as a result, the time chosen for the interaction to run is $t = \frac{\theta}{g\sqrt{n}}$ for a rotation though θ where $\bar{n}$ is the coherent state's mean photon number, but this results in some decoherence which is small for large $\bar{n}$. Therefore, the number of cavities required is still q, but the number of excitations needed is $O(\frac{qD}{\epsilon})$, where D is the average number of gates needing to be performed on any one qubit. Still, this system is easy to error correct, as has been demonstrated, and this could help manage errors from other sources also.

Note that in all of these cases there will be errors related to the need to perform gates that commute with the total energy. This is separate from a reference requirement, but it still implies that a number of excitations $O(\frac{G}{\epsilon})$ is required to perform G gates that commute with the total energy.

Treating the reference frame as classical does not mask a restriction on quantum computation, in the sense that in theory quantum computers do not require a superpolynomial reference resource to implement polynomial sized circuits. That said, if one desires to complete a very large quantum computation and energy is at a premium, then using a method that operates in a DFS for excitation number is superior. However, these methods require more qubits to be kept coherent, and so are less robust in many circumstances and the analysis does not include

the requirements for the error correction of non-reference errors, including errors incurred from commuting gates.

Chapter 5

Quantum Walk with Entangled Particles

Contents

5.1 Overview

Here, a quantum walk algorithm is given in which a speed up over traditional quantum walks is achieved through the novel idea of using two entangled walkers. The idea of a quantum walk with two particles is developed and studied for the case

of a discrete time walk on a line [SPOB06, OPSB06]. Both separable and maximally entangled initial conditions are considered, and it is shown how the entanglement and the relative phase between the states describing the *coin* degree of freedom of each particle will influence the evolution of the quantum walk. In particular, these factors will have consequences on the distance between the particles and the probability to find them at a given point, yielding results that cannot be obtained from a separable initial state, be it pure or mixed. Finally, there is a brief review of proposals for implementations. The work in this chapter has appeared in [OPSB06] and [SPOB06].

5.2 Introduction

Quantum walks, first proposed in 1993 [ADZ93], are the quantum analogue of classical random walks. Quantum walks (for an overview, see [Kem03]) have generated much interest, in particular because they are proving to be a very useful technique for the construction of quantum algorithms, just as random walks are for classical algorithms. Several quantum algorithms based on quantum walks have been shown to be optimal [SKW03, Amb07], and an oracle algorithm that offers an exponential speed-up with respect to its classical counterpart [CCD$^+$03] is based on a (continuous time) quantum walk.

The crucial difference between quantum and random walks is that the former allow for quantum superpositions of the walker states and explore the interference of the terms in these superpositions. The resulting probability distributions after N steps are very different, as can be seen in Figure 5.1. Furthermore, a quantum walk exhibits a variance proportional to N, which represents a quadratic speed-up over the classical case, where the variance goes as $\sqrt{N}$.

There are several variants of quantum walks. Walks can evolve in discrete time or continuous time, and can take place on arbitrary graphs, with varying transition probabilities. Szegedy [Sze04] proposed a Markov chain walk model based on the diffusion operator, following a formalism presented by Watrous in [Wat01]. That model works even on directed graphs where each transition probability for each edge can be set individually and the walk evolution operator is a pair of reflections. The states of the walk are given by the tensor product of two copies of the vertex space of the graph $V \times V$:

$$|\psi_u\rangle = |u\rangle \otimes \sum_{v \in N(u)} a_{v \to u} |v\rangle \,, \tag{5.1}$$

where u and v denote vertices so that the first part of the product can denote a current position and the second the previous location, $a_{v\to u}$ is the amplitude corresponding to the square root of the probability of moving from vertex v to u, $\sqrt{p_{v\to u}}$, possibly with a phase, and $N(u)$ are all the vertices adjacent to u. (If there is a non-zero probability of remaining in the same location, *i.e.* $a_{u\to u} \neq 0$, then instead the closed neighbourhood of u should be used, which includes u, $N[u]$.) This large Hilbert space to describe the evolution of the walk allows the evolution to be very general, since there can be a probability of transition from any vertex of the walk graph to any other vertex. The two reflections that together, one applied after the other make up two steps in the evolution of the walk are then defined by

$$R_1 = 2\Pi - \mathbb{I} \qquad \text{and} \qquad R_2 = 2S\Pi S - \mathbb{I} \tag{5.2}$$

where S is the operator that swaps the two registers in the tensor product and $\Pi = \sum_{u\in V} |\psi_u\rangle \langle\psi_u|$. Then the operator that completes two steps of the walk is $W = R_2R_1$ or, one step of the walk is given by SR_1, that is, there is a reflection about the equal superposition of all states of well-defined current position $|\psi_u\rangle$ and then the two registers that hold the current and previous positions of the walkers are swapped. This is repeated for each step of the walk.

Note however, that the diffusion operator is chosen to be as far as possible from the identity [MR02], which in the case of a graph with each vertex having only two adjacent vertices will give the behaviour of a coin that is the X operator.

$$2 |+\rangle \langle+| - \mathbb{I} = X \tag{5.3}$$

This is not the approach taken here. We consider a walk on a line and wish to make a model analogous to the behaviour of a classical random walk on a line. Therefore, a Hadamard coin is used to ensure that there is a probability of one half for moving either left or right for a walker localized at a particular vertex. We next introduce the variations we propose to the standard quantum walk on a line.

5.3 The Multiparticle Walk

The new concept being introduced in this work is the idea of using multiple walkers — quantum particles — simultaneously in the same quantum walk. This idea has been used classically, where k walkers traversing the same graph are equivalent to k independent random walks. These additional resources can be used, for instance,

to improve the time to hit a certain node of the graph. In the quantum case though, the multiple-walker situation becomes even more interesting, because the particles can be entangled, which results in final probability distributions which cannot be described by multiple independent single-particle walks. It is demonstrated that the entanglement in a quantum walk on a line can be tuned to cover more space.

5.4 The Quantum Walk Formalism

First consider a single-particle discrete time quantum walk on a line. A single step of the walk comprises two operations: the coin flip C and the conditional shift S. The coin degree of freedom is encoded by a qubit, whose states will be denoted $|\uparrow\rangle$ and $|\downarrow\rangle$, and the coin flip is a unitary operator that sets a basis state into a superposition over that basis. The qubit system C acts on the coin Hilbert space $\mathcal{H}_C$. The most commonly used example is the Hadamard transformation, and this shall be used here throughout as the coin:

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{5.4}$$

This coin is said to be *unbiased* in the sense that an initial state $|\uparrow\rangle$ will be set into an equal superposition of "up" and "down", $\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$, after being acted upon by the coin. This corresponds to equal probabilities of the particle being found in the state $|\uparrow\rangle$ or $|\downarrow\rangle$ if measured after one operation. The second part of a single step of the walk is the shift operation S, acting on $\mathcal{H}_P$ — the Hilbert space encoding the possible positions of a particle on an infinite discrete line. The basis vectors of $\mathcal{H}_P$ will be denoted $|i\rangle$, where $i \in \mathbb{Z}$. Thus the total Hilbert space of the walk is given by:

$$\mathcal{H} = \mathcal{H}_C \otimes \mathcal{H}_P, \tag{5.5}$$

and the states of the quantum walk will be described by vectors in $\mathcal{H}$.

Consider the following shift operator:

$$S = \sum_i \left(|\uparrow\rangle\langle\uparrow| \otimes |i+1\rangle\langle i| + |\downarrow\rangle\langle\downarrow| \otimes |i-1\rangle\langle i| \right). \tag{5.6}$$

This has the effect of moving a particle in the "up" state one unit to the right and a particle in the "down" state one unit to the left. If the particle in question is in some superposition of "up" and "down", the terms evolve accordingly. C acts on

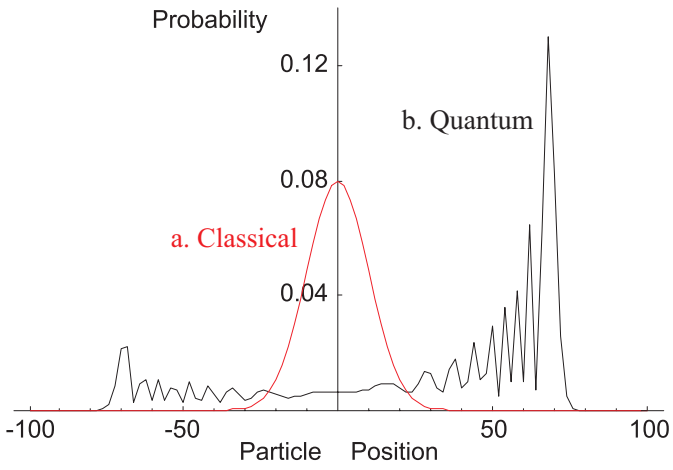


Figure 5.1: Probability distributions for discrete time random walks on a line after $N = 100$ steps: (a) the classical walk, (b) the quantum walk with a Hadamard coin and initial state $|0\rangle \otimes |\uparrow\rangle$. The asymmetry in the quantum case is the result of the initial state of the coin register. The shape of the interference pattern is symmetric when the initial state $|0\rangle \otimes \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$ is chosen.

the coin state space and S acts on both the position state space and coin space together. The total operator for each step is unitary and will have the form:

$$U = S(I_P \otimes C). \tag{5.7}$$

where I_P is the identity operator on the position space. So, for example, if the initial state is $|\uparrow\rangle \otimes |0\rangle$, the first step of the walk gives:

$$U(|\uparrow\rangle \otimes |0\rangle) = \frac{1}{\sqrt{2}}(|\uparrow\rangle \otimes |1\rangle + |\downarrow\rangle \otimes |-1\rangle). \tag{5.8}$$

If a measurement is performed at this point, then the walk agrees with its classical counterpart: there is a probability of one half for the particle being found at position $+1$ and a probability of one half for the particle being found at position -1. Yet, after the first two steps of the walk, the progression of the quantum and classical walks begin to diverge. This becomes particularly clear after a few dozen steps. In Figure 5.1 the probability distributions are presented for both the classical random walk and the quantum walk for $N = 100$.

Note that the quantum walk shows a relatively low probability associated with the walker being found close to the origin: rather, the peaks of the distribution correspond to the particle being found a considerable distance away. On the other

hand, in the classical case, the origin is precisely the point where the probability for the particle to be found is maximal. There is a $\sqrt{N}$ speed-up of the quantum walk over the classical in terms of expected deviation of the walker from the origin.

5.5 Quantum Walk with two Particles

Now generalize this definition of the discrete time quantum walk on a line to the case of two walkers. Consider two particles, 1 and 2, simultaneously completing quantum walks on the same line. Let the joint Hilbert space of the two-particles system be $\mathcal{H}_{12}$. Then,

$$\mathcal{H}_{12} = \mathcal{H}_1 \otimes \mathcal{H}_2, \tag{5.9}$$

where $\mathcal{H}_1$ and $\mathcal{H}_2$ are the Hilbert spaces of particles 1 and 2 respectively. Both $\mathcal{H}_1$ and $\mathcal{H}_2$ are isomorphic to $\mathcal{H}$, as defined in equation (5.5). Similarly, the new walk operator will be U_{12}, given by:

$$U_{12} = U \otimes U. \tag{5.10}$$

Suppose that two distinguishable particles are set into the walk in a pure separable initial state, for example:

$$|\Psi_0^{\text{sep}}\rangle = |0,\downarrow\rangle_1 |0,\uparrow\rangle_2 . \tag{5.11}$$

In this case, the two particles evolve throughout the walk simultaneously, but independently, so that after N steps:

$$|\Psi_N^{\text{sep}}\rangle = U_{12}^N |\Psi_0^{\text{sep}}\rangle = U^N |0,\downarrow\rangle_1 U^N |0,\uparrow\rangle_2 .$$

This is equivalent to two separate walks being completed on two different lines at the same time. The probability distributions are both identical (though one is inverted in the position coordinate relative to the other) and independent.

However, it is also possible for a pair of particles to be set into the walk in a joint state which is entangled. For example, consider two maximally entangled states:

$$|\psi_0^{\pm}\rangle_{12} = \frac{1}{\sqrt{2}}(|0,\downarrow\rangle_1|0,\uparrow\rangle_2 \pm |0,\uparrow\rangle_1|0,\downarrow\rangle_2), \tag{5.12}$$

Note that these also describe the cases of a pair of identical particles on the same vertex, either bosons (the "+" state) or fermions (the "−" state). Now the evolution of the walk cannot be described as two separate walks progressing independently.

The joint state of the particles' evolution after N steps is:

$$\left|\psi_N^{\pm}\right\rangle_{12} = U_{12}^N \left|\psi_0^{\pm}\right\rangle_{12} = \frac{1}{\sqrt{2}}\left(U^N|0,\downarrow\rangle_1 \, U^N|0,\uparrow\rangle_2 \pm U^N|0,\uparrow\rangle_1 \, U^N|0,\downarrow\rangle_2\right).$$

These two scenarios correspond to very different final joint probability distributions. Let $P_{12}(i,j,N)$ denote a joint probability over the two particles of finding particle 1 in position i and particle 2 in position j after N steps of the walk. In the case of the separable initial conditions, $|\Psi_0^{\text{sep}}\rangle = |0,\downarrow\rangle_1 |0,\uparrow\rangle_2$,

$$P_{12}^{\text{sep}}(i,j;N) = P_1^{\text{sep}}(i;N) \times P_2^{\text{sep}}(j;N), \tag{5.13}$$

which is just the product of two independent distributions for single-particle walks after N steps. This is in agreement with the observation that the two particles in the walks proceed independently, and without reference to each other, into the final state given in equation (5.12). However, this equation does not describe the case of entangled particles, which have walk evolutions which depend on the existence of the other particle in the walk. The joint probability distribution must be symmetric under exchange of the labels 1 and 2 for the "+" case and, for the "−" case, the distribution must remain unchanged when *both* the labels of the particles are exchanged and the position axes are reflected. These distributions are presented graphically in Figure 5.2 for the case $N = 60$.

In the case of the separable initial conditions, the plot in Figure 5.2(a) is just the product of two distributions like the one in Figure 5.1(b). The initial state is $|\Psi_0^{\text{sep}}\rangle = |0,\downarrow\rangle_1 |0,\uparrow\rangle_2$, so the distribution for particle 2 is biased towards the right (as shown in Figure 5.1(b)) and the distribution for particle 1 is biased towards the left — a mirror image of the distribution for particle 2. Figure 5.2(b) shows the distribution for P_{12}^{+} and Figure 5.2(c) shows the distribution for P_{12}^{-}.

In all three distributions the maxima occur around $i,j \simeq 42$, but the effect of the entanglement is dramatic. Entanglement can induce the walk to explore certain configurations with relatively high probability that for a walk beginning in a separable state would be very unlikely. For example, the higher maxima of the distribution for the symmetric entangled state ("+" state) correspond to the particles being found on the same side of the origin and both near $|i|,|j| \simeq 42$. Conversely, in the antisymmetric, "−", case, there are a whole set of configurations that have probability zero associated with their occurrence. These are the configurations along the line $i = j$, which implies that the two particles are never in the same position on the line. Note that a pair of identical fermions conducting a quantum

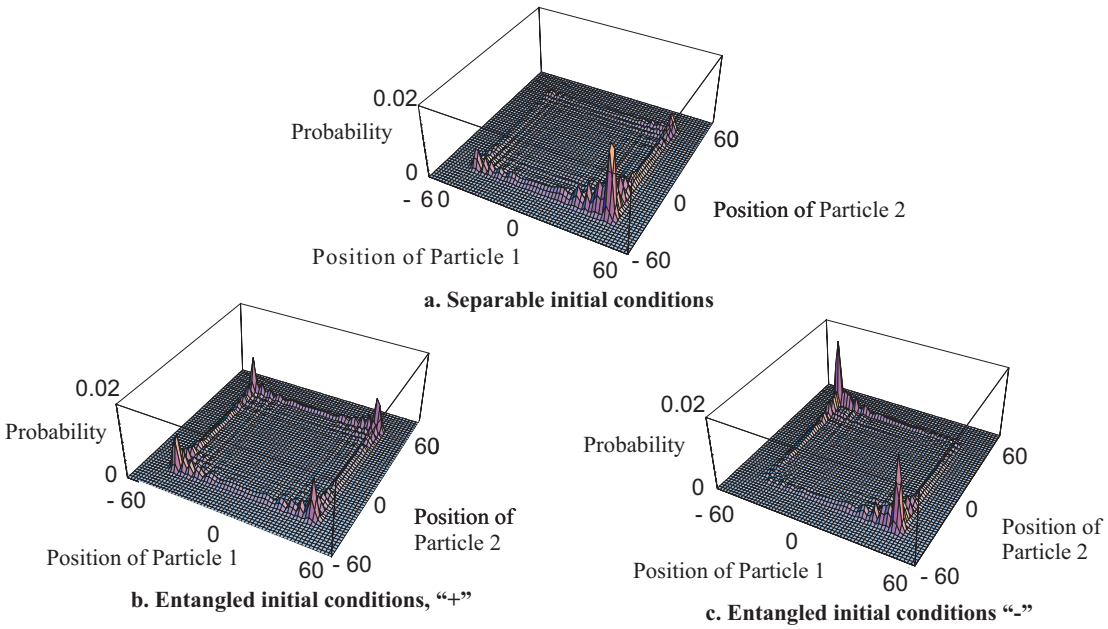


Figure 5.2: Two-particles probability distributions after $N = 60$ steps for different initial conditions: (a) separable state $|\psi_0^S\rangle_{12}$; (b) $|\psi_0^+\rangle_{12}$ state; and (c) $|\psi_0^-\rangle_{12}$ state. Note that the peak of the distribution in (a) has been cropped.

walk will follow this distribution (Figure 5.2(c)), in which the particles have high probability of being found at opposite ends of the line. A pair of identical bosons will follow the symmetric distribution shown in Figure 5.2(b).

The behaviors of the individual particles in the quantum walk for entangled initial conditions also exhibit interesting features. Tracing over the degrees of freedom of one of the particles leaves a reduced density matrix:

$$\rho_1 = \mathrm{Tr}_2\left(\left|\psi_N^{\pm}\right\rangle_{12}\left\langle\psi_N^{\pm}\right|_{12}\right) = \frac{1}{2}U^N\left|0,\downarrow\right\rangle\left\langle 0,\downarrow\right|U^{\dagger N} + \frac{1}{2}U^N\left|0,\uparrow\right\rangle\left\langle 0,\uparrow\right|U^{\dagger N}. \quad (5.14)$$

which is an equal mixture or the states $U^N\left|0,\downarrow\right\rangle$ and $U^N\left|0,\uparrow\right\rangle$. These are simply the evolution states after N steps for a walk starting from the state $\left|0,\downarrow\right\rangle$ and the state $\left|0,\uparrow\right\rangle$. Therefore the marginal probability distribution for finding one of the particles in position i after N steps is given by

$$P_1^{\pm}(i;N) = \frac{1}{2}[P_{\downarrow}(i;N) + P_{\uparrow}(i;N)] = P_2^{\pm}(i;N). \quad (5.15)$$

Note that $P_{\downarrow}(i;N)$ and $P_{\uparrow}(i;N)$ are the probability distributions of a single-particle walk initialized in the state $\left|0,\downarrow\right\rangle$ and $\left|0,\uparrow\right\rangle$. Again compare this situation to the case of separable states. The separable initial condition given in equation (5.11) gener-

ates the marginal probability distributions $P_1^{\rm sep}(i;N) = P_\downarrow(i;N)$ and $P_2^{\rm sep}(i;N) = P_\uparrow(i;N)$, giving a joint probability distribution which is just the product of these (equation (5.13)). This is not the case for the entangled state, where the joint probability distribution contains information about the correlations between the positions of the two particles, which will become apparent on measurement of the two particles' positions.

One can also demonstrate that no mixed separable state diagonal in the $\{|0,\downarrow\rangle_1|0,\uparrow\rangle_2, |0,\uparrow\rangle_1|0,\downarrow\rangle_2\}$ basis can exhibit the same properties between the two particles as entangled states. Consider the following mixed state, a weighted mixture of two alternatives described by the states $|0,\downarrow\rangle_1|0,\uparrow\rangle_2$ and $|0,\uparrow\rangle_1|0,\downarrow\rangle_2$, as a new initial state for the two particles quantum walk:

$$\rho_{12}^{\rm MS}(0) \;=\; a|0,\downarrow\rangle\langle 0,\downarrow|_1 \otimes |0,\uparrow\rangle\langle 0,\uparrow|_2 + b|0,\uparrow\rangle\langle 0,\uparrow|_1 \otimes |0,\downarrow\rangle\langle 0,\downarrow|_2, \quad (5.16)$$

where $a, b \in \mathbb{R}^+$ such that $a + b = 1$. After N steps of the quantum walk, the system will be in the state $\rho_{12}^{\rm MS}(N) = U^N \rho_{12}^{\rm MS}(0) U^{\dagger N}$, and the marginal probability distributions will be:

$$\begin{aligned} P_1^{\rm MS}(i;N) &= aP_\downarrow(i;N) + bP_\uparrow(i;N), \\ P_2^{\rm MS}(i;N) &= aP_\uparrow(i;N) + bP_\downarrow(i;N). \end{aligned} \quad (5.17)$$

For $a = b$, these marginal distributions are identical to the the marginal distributions in the case of entangled particles. In general for $(a \neq b)$ they are biased, and

$$\langle x_1\rangle = \sum_{i=-N}^{N} iP_1^{\rm MS}(i;N) = a\langle x_\downarrow\rangle + b\langle x_\uparrow\rangle = (a-b)\langle x_\downarrow\rangle = -\langle x_2\rangle. \quad (5.18)$$

Here, use has been made of the expression $\langle x_{\downarrow,\uparrow}\rangle \equiv \sum_{i=-N}^{N} iP_{\downarrow,\uparrow}(i;N)$, with $P_{\downarrow,\uparrow}(i;N)$, for the probability distributions after N steps for initial states $|0,\downarrow\rangle$ and $|0,\uparrow\rangle$, and the fact that $\langle x_\downarrow\rangle = -\langle x_\uparrow\rangle$.

However a difference comes in the joint probability distribution after N steps, as seen from equation (5.16), for the mixed state it is given by:

$$P_{12}^{\rm MS}(i,j;N) = aP_\downarrow(i;N)P_\uparrow(j;N) + bP_\uparrow(i;N)P_\downarrow(j;N), \quad (5.19)$$

which is distinct from the joint probability distributions for the entangled states, but also not the product of two one-particle marginal distributions $P_\downarrow(i;N)$ and

$P_\uparrow(j;N)$. It is the appropriately weighted sum of the distribution for separable initial conditions given in Figure 5.2(a) with the same distribution, but with the particles 1 and 2 reversed.

There are further joint properties of the two particles which are of interest. Define the distance $\Delta_{12}^{\text{sep},\pm}$ between the two particles' final (independently) measured positions, x_1 and x_2, after N steps:

$$\Delta_{12}^{\text{sep},\pm} \equiv |x_1 - x_2|, \tag{5.20}$$

where $x_1, x_2 \in [-N, \ldots, 0, \ldots, N]$. Table 5.1 presents the expectation value of this distance for the three different initial conditions, and for different N. From this data, it is evident that the particles are more likely to remain closer together for the "+"-entangled initial conditions than for the separable initial state, and more likely to remain farther apart in the case of the "−"-entangled initial state than for the separable state. In fact, an even stronger statement can be made; for a given N we always have:

$$\langle \Delta_{12}^{-} \rangle - \langle \Delta_{12}^{\text{sep}} \rangle = \langle \Delta_{12}^{\text{sep}} \rangle - \langle \Delta_{12}^{+} \rangle \tag{5.21}$$

Consider now the correlation function between the spatial distribution of each of the two particles:

$$C^{\text{sep,MS},\pm}(x_1, x_2) \equiv \langle x_1 x_2 \rangle - \langle x_1 \rangle \langle x_2 \rangle. \tag{5.22}$$

Clearly, in the case of the separable initial condition (5.11), this correlation is always zero. For the other cases, the values of $C^{\pm}(x_1, x_2)$ are presented in Table 5.2 for different N. Given the symmetry of the $P_{12}^{\pm}$ distributions, in those cases $\langle x_1 \rangle = \langle x_2 \rangle = 0$. Thus, the sign difference in the correlation function expresses the tendency for the two particles in the "−" case to end the quantum walk on different sides of the line (with respect to the origin, 0), and on the same side for the "+" case.

The correlation function for the mixed state $\rho_{12}^{\text{MS}}(N)$, $C^{\text{MS}}(x_1, x_2)$ is not zero, as can be seen using $\langle x_1 x_2 \rangle = -\langle x_\downarrow \rangle^2$:

$$C^{\text{MS}}(x_1, x_2) = -\left[1 - (a-b)^2\right] \langle x_\downarrow \rangle^2. \tag{5.23}$$

The values of $C^{\text{MS}}(x_1, x_2)$ for the case $a = b = \frac{1}{2}$ are presented in Table 5.2. Since the function $[1 - (a-b)^2] \in [0, 1]$, therefore $\rho_{12}^{\text{MS}}(N)$ is always less correlated than a singlet entangled state. Actually, $\rho_{12}^{\text{MS}}(N)$ is only a *classically correlated* state,

Expectation value $\langle\Delta_{12}^{\text{sep},\pm}\rangle$ after N steps						
No. of steps N	10	20	30	40	60	100
Init. cond. $\|\psi_0^-\rangle_{12}$	8.8	17.5	26.0	34.9	52.2	87.0
Init. cond. $\|\psi_0^{\text{sep}}\rangle_{12}$	7.1	14.7	21.9	29.5	44.3	73.9
Init. cond. $\|\psi_0^+\rangle_{12}$	5.5	11.9	17.8	24.1	36.3	60.8

Table 5.1: Average distance $\langle\Delta_{12}^{\text{sep},\pm}\rangle$ after N steps.

and when compared with entangled states $|\psi_N^{\pm}\rangle_{12}$, it exhibits the same features as the pure separable state $|\psi_N^{PS}\rangle_{12}$. Namely, the expectation value of the distance is the same for both pure and mixed separable states, $\langle\Delta_{12}^{\text{MS}}\rangle = \langle\Delta_{12}^{\text{sep}}\rangle$, as seen from the probability distribution (5.19).

Finally, for the different initial conditions, the probability of finding at least one particle in position i after N steps can be calculated: $\mathcal{P}^{\text{sep,MS},\pm}(i;N)$. This is a joint property as it depends on both one-particle outcomes:

$$\begin{aligned}\mathcal{P}^{\text{sep},\pm}(i;N) &= \sum_{j=-N}^{N}\left[P_{12}^{\text{sep},\pm}(i,j;N)+P_{12}^{\text{sep},\pm}(j,i;N)\right]-P_{12}^{\text{sep},\pm}(i,i;N)\\ &= [P_{\downarrow}(i;N)+P_{\uparrow}(i;N)]-P_{12}^{\text{sep},\pm}(i,i;N).\end{aligned}\tag{5.24}$$

Similarly, the probability of finding at least one particle in position i after N steps is the same for both pure separable and mixed separable case: $\mathcal{P}^{\text{sep}}(i;N)=\mathcal{P}^{\text{MS}}(i;N)$.

Given a one-particle probability distribution, say $P_{\downarrow}(i;N)$ (note that we have $P_{\uparrow}(i;N)=P_{\downarrow}(-i;N)$), the probability $\mathcal{P}^{\text{sep},\pm}(i;N)$ decreases with the joint probability $P_{12}^{\text{sep},\pm}(i,i;N)$ and is maximal in the "$-$" case, since $P_{12}^{-}(i,i;N)=0$. In fact, around the points $(40,40)$ and $(-40,-40)$ in Figure 5.2 it can be seen that:

$$\mathcal{P}^{-}(i;N)>\mathcal{P}^{\text{sep}}(i;N)>\mathcal{P}^{+}(i;N).\tag{5.25}$$

By introducing entanglement in the initial conditions of the two-particles quantum walk, the probability of finding at least one particle in a particular position on the line can be greater or less than in the case where the two particles are independent and note that in this case where both states considered are are maximally entangled the probability depends on the form of symmetry (symmetric or antisymmetric) in the state.

Correlation function $C^{\text{sep,MS},\pm}(x_1, x_2)$ after N steps

No. of steps N	10	20	30	40	60	100
Init. c. $\|\psi_0^-\rangle_{12}$	-16.8	-69.8	-153.5	-276.2	-619.7	-1718.3
Init. c. $\rho_{12}^{\text{MS}}(0)$	-6.0	-31.4	-70.6	-130.4	-299.3	-839.3
Init. c. $\|\psi_0^{\text{sep}}\rangle_{12}$	0	0	0	0	0	0
Init. c. $\|\psi_0^+\rangle_{12}$	4.8	7.3	13.7	15.1	23.1	39.1

Table 5.2: Correlation function $C^{\text{sep,MS},\pm}(x_1, x_2)$ after N steps.

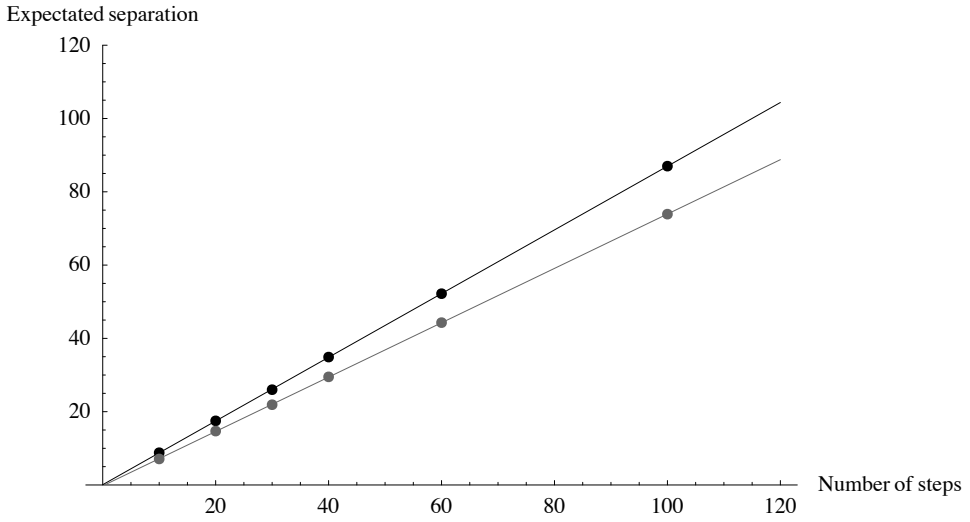


Figure 5.3: The data is taken from Table 5.1 and shows the expectation of the particles separation $\langle \Delta_{12}^{\text{sep},-} \rangle$ against the number of steps of the walk taken, N, for the separable initial condition (grey line and points) and the antisymmetric initial condition (black line and points). This shows that there is a linear speed up in the entangled case over separable initial condition.

5.6 Conclusions, Further Study, and Implementations

In this article the concept of a quantum walk with two particles was introduced and studied for the case of a discrete time walk on a line. Having more than one particle allowed the addition of a new feature to the walk: entanglement between the particles. In particular, initial states that are maximally entangled in the coin degrees of freedom and with opposite symmetries were considered, and compared to the case where the two particles were initially in a pure or mixed separable state (the particles being then independent or only classically correlated, respectively). It was found that the entanglement in the coin states introduced spatial correlations

between the particles, and that their average distance is larger in the "-" case than in the separable case, and is smaller in the "+" case. This could benefit algorithmic applications which require two marked sites to be reached which are known *a priori* to be on the opposite or on the same sides of a line. It was also found that the introduction of entanglement could increase or decrease the probability of finding at least one particle on a given point of the line. This increase could allow us to reach a marked site faster than with two unentangled quantum walkers. In particular, the speed up of the expected separation of the walkers for the "-" entangled case was a linear improvement over the separable case. The entanglement in the initial conditions thus appears as a resource that can be tuned according to the specific needs for enhancement of a given application (algorithmic or other) based on a quantum walk.

There are a number of natural extensions to this work on discrete time quantum walks on a line with two entangled particles, starting with the study of other initial entangled states and the introduction of more particles. It would also be interesting to consider other graphs for the walks, as well as less standard coins, such as unbalanced or even entangling ones [HBF03, BB04]. There have been a number of studies in recent years considering the effect of multiple coins and their entanglement with a single walker [PA07]. Another direction worth exploring would be a multiparticle quantum walk in continuous time, which does not use a coin, and which could be a relevant model to study the evolution of dilute quantum gases. These ideas may find use for particular applications or the design of quantum algorithms. Random walks have been employed in such tasks as estimating the volume of a convex body [DFK91] and the connectivity in P2P networks [LCC$^+$02]. Quantum algorithms based on walks could be useful in these problems. This proposal demonstrates that there are further quantum behaviours other than simply allowing superpositions of the walker that remain to be exploited in these types of algorithms.

There have been several proposals to implement a single-particle quantum walk, using cavity QED [SBTK03], optical lattices [DRKB03] and ion traps [TM02]. Another possibility is to send two photons through a tree of balanced beam splitters which implement both the coin flipping and the conditional shift, again generalizing a scheme proposed for a single particle [HBF03, JPK04]. A similar scheme was implemented by Bouwmeester *et al.* [BMK$^+$99]. Note that this could be implemented with other particles as well, e.g. electrons, using a device equivalent to a beam splitter [LOYT98]. Finally, observe that by considering indistinguishable bosons or fermions, the effects of quantum statistics can be used to prepare the ini-

tial entangled states, thus appearing again as a resource for quantum information processing [OPBV02, POBV02]. Indeed, since this work was first made public, a scheme for interfering two photons on a tree of polarizing beam splitters and phase shifting plates has been proposed and analyzed that also make use of the statistics of identical particles to generate entanglement [PA07].

Chapter 6

Approximating Fractional Time Quantum Evolution

Contents

6.1 Overview

In this chapter the relationship between continuous and discrete unitary evolution is studied by the construction of a new algorithm to simulate continuous evolution with discrete applications of a unitary oracle. Specifically, an algorithm is presented for approximating arbitrary powers of a black box unitary operation, $\mathcal{U}^t$, where t is a real number, and $\mathcal{U}$ is a black box implementing an unknown unitary. The complexity of this algorithm is calculated in terms of the number of calls to the black box and the errors in the approximation. For general $\mathcal{U}$ and large t, one should apply $\mathcal{U}$ a total of $\lfloor t \rfloor$ times followed by our procedure for approximating the fractional power $\mathcal{U}^{t-\lfloor t \rfloor}$. An example is also given where for large integers t this method is more efficient than direct application of t copies of $\mathcal{U}$. Further applications and related algorithms are also discussed. 5

6.2 Introduction

If a unitary operation is presented as a resource for running algorithms, but its description is not provided, this is a less general resource than if access is given to the unspecified Hamiltonian that generates this unitary in a particular amount of time. While it is clear how to implement a unitary operation from a Hamiltonian, exponentiating a unitary oracle is not so straightforward.

An n-qubit unitary $\mathcal{U}$ can be implemented by evolving (or simulating the evolution of) a *time-independent* Hamiltonian H for a period of time $\tau = 1$, that is, $U = e^{-iH}$. Then for any $t \in \mathbb{R}_+$, one can implement $\mathcal{U}^t$ by simply evolving the Hamiltonian for a period of time t. For example, if $t = \frac{1}{2}$, then a square root of $\mathcal{U}$, $e^{-i\frac{1}{2}H}$, could be implemented in this way, and in such model of computation the cost would be half of the cost of implementing $\mathcal{U}$.

In this work, the question what can be done if $\mathcal{U}$ is realized in some other way, such as a non-trivial sequence of time-dependent Hamiltonians, or a quantum circuit is explored. In other words, consider the situation when $\mathcal{U}$ is given in the form of a black box. The goal is to implement real valued powers t of this unitary operation by making use of the multiple copies of the black box implementing $\mathcal{U}$. The complexity of such a procedure is measured in terms of the total number of calls to the unitary.

It is possible to find the t^{th} power of an unknown unitary by first performing a sufficiently precise complete quantum process tomography of a $2^n \times 2^n$ dimensional

unitary $\mathcal{U}$, which uses $O(4^n)$ calls to the unitary with various input states [CN97, NC00] and measurements to achieve $\mathcal{U}$ with constant precision. The exponential scaling with n is necessary. In particular, a lower bound on the number of calls to the unitary can be derived from constructing an $\epsilon-$net over unitaries [1]. If the space of unitary operations is divided up into balls of radius ϵ then the total number of unitaries that can be specified up to this resolution is $\left(\frac{c}{\epsilon}\right)^{4^n}$, where c is a constant. To specify each of these requires $4^n \log(\frac{c}{\epsilon})$ bits. To discover information about $\mathcal{U}$ one might supply states to the unitary tensored with the identity operation of the same dimension (a larger dimension does not help) and perform measurements on the output states. The states will have dimension 4^n, and by Holevo's bound one could hope to obtain no more than $2n$ bits of information for each of these calls to $\mathcal{U}$ (in general it will be less than this). Therefore, it is not possible by tomography to use fewer than $\Omega\left(\frac{4^n}{2n}\log(\frac{c}{\epsilon})\right)$ calls to $\mathcal{U}$. Even allowing for error in some fraction δ of the basis states, this still cannot reduce the required number of calls to any function polynomial in n [Aar07].

In this study a more efficient approach to implementing powers is described. In particular, an algorithm is presented for approximating any constant power of an unknown unitary using only $O(\frac{1}{\epsilon}\log\frac{1}{\epsilon})$ calls to the unitary itself, with error ϵ calculated using the trace norm. Note that this complexity is independent of the number of qubits n that $\mathcal{U}$ acts on.

The relation between discrete and continuous oracles for various problems has been the subject of numerous previous studies. Farhi and Gutmann [FG98] introduced the concept of (continuous) Hamiltonian black box oracles for quantum computing. Ioannou [Ioa02], and Roland and Cerf [RC03] consider the problem of simulating a Hamiltonian for Grover search [Gro96] on a discrete computer. Mochon [Moc07] extends this to a more general setting where he is concerned with finding lower bounds for discrete oracle problems by considering them in the Hamiltonian setting and mapping them to the problem of finding geodesics in manifolds. In particular, he considers one-item Grover search and oracle interrogation, highlighting the case of computing the XOR function on a hidden bit string. In these cases, he is able to exploit symmetries in the oracle problems to solve what is otherwise a very difficult problem. The focus in this work is different. Firstly, an oracle is allowed to be applied only an integer number of times and not for a fractional amount of time. Secondly, the new algorithm presented here has a focus on (constructive) upper bounds. Thirdly, symmetry assumptions are not made on the oracles. Lastly, the eigenvalues of these unitaries are not restricted to ± 1 as

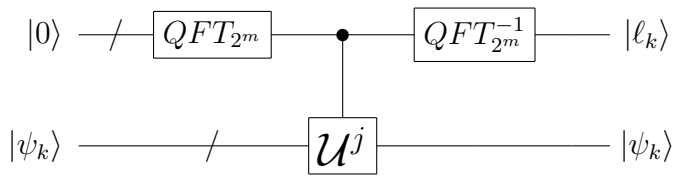
[1] We thank P. Hayden for this observation.

in the Grover search and XOR case, or the work in [CGM+08] which shows how to simulate a general continuous *Boolean* oracle evolving for a total of time T with $O(\frac{T \log T}{\log \log T})$ discrete oracle queries. Since this work was completed, Harrow, Hassidim, and Lloyd [HHL08] found an algorithm for finding the expectation value of a vector which is a solution to a system of linear equations which is based on eigenvalue estimation in a similar way to the one proposed here. That algorithm shows exponential speed up over the classical approach and is discussed further in the conclusion.

The remainder of this chapter is organized as follows. In Section 6.4 the basic construction is described for efficient computation of the powers of $\mathcal{U}$, the assumptions and constraints are discussed, the complexity of the computation is given, and the precision of the approximate solution is given. The full calculation of these results is included in Section 6.5. An example of a special case in which t is a large positive integer and the new algorithm is more efficient than direct application of the t copies of the unitary is given in Section 6.9. Related applications of this algorithm to computing fractional Fourier transform and noise filtering are explored in Section 6.10. Finally, the implications of this work are discussed in the concluding section.

6.3 A Brief Review of the Eigenvalue Estimation Algorithm

The new algorithm makes use of the Eigenvalue Estimation algorithm as a subroutine, so it is convenient to review it now. The algorithm in its traditional form accepts as inputs a unitary operator, $\mathcal{U}$, and one of its eigenvectors, $|\psi_k\rangle$, and determines the eigenvalue, λ_k, associated with this eigenvector. It does this by making use of the quantum fourier transform acting on m ancilla qubits to call in equal superposition each integer power of the controlled unitary operator $\mathcal{U}^j$ for $0 \leq j \leq 2^m - 1$, where m is the parameter that determines the accuracy of the estimate. The same superposition can be made using the Hadamard operator applied to each ancilla qubit. This causes each term in the superposition to acquire a phase term $e^{2\pi i \frac{j\ell_k}{2^m}}$.



where $\frac{\ell_k}{2^m} = \lambda_k \in [0,1)$ and $\mathcal{U}|\psi_k\rangle = e^{2\pi i\lambda_k}|\psi_k\rangle$. The algorithm proceeds in three steps:

1. The QFT_{2^m} takes $|0\rangle$ to $\frac{1}{\sqrt{2^m}}\sum_{j=0}^{2^m-1}|j\rangle$.
2. The $c-U_j$ gates takes $|j\rangle|\psi_k\rangle$ to $|j\rangle\mathcal{U}^j|\psi_k\rangle = e^{2\pi i\frac{j\ell_k}{2^m}}|j\rangle|\psi_k\rangle$ by the principle of phase kick-back.
3. The $QFT_{2^m}^{-1}$ takes the system to $\frac{1}{2^m}\sum_{j,j'=0}^{2^m-1}e^{2\pi ij\frac{(\ell_k-j')}{2^m}}|j'\rangle|\psi_k\rangle = |\ell_k\rangle|\psi_k\rangle$.

The input to the algorithm need not be an eigenstate, $|\psi_k\rangle$. This can be done in superposition over eigenstates indexed by k:

$$\sum_k \alpha_k|0\rangle|\psi_k\rangle \mapsto \sum_k \alpha_k|\ell_k\rangle|\psi_k\rangle,$$

and our goal is to perform the operation $\mathcal{U}^t$ on the state $|\Psi\rangle$, $\mathcal{U}^t|\Psi\rangle$:

$$\mathcal{U}^t|\Psi\rangle = \sum_k \alpha_k\mathcal{U}^t|\psi_k\rangle = \sum_k \alpha_k e^{2\pi i\lambda_k t}|\psi_k\rangle.$$

6.4 The Algorithm

In this section, the basic construction of the algorithm is described. For any unitary $\mathcal{U}$ on a finite dimensional state space, consider its spectral decomposition $\mathcal{U} = P\Lambda P^\dagger$, where P is a unitary matrix composed of the eigenvectors of $\mathcal{U}$, and Λ is a diagonal matrix containing the eigenvalues of $\mathcal{U}$. For such a decomposition, powers of $\mathcal{U}$ may be computed as follows:

$$\mathcal{U}^t = P\Lambda^t P^\dagger. \tag{6.1}$$

For a black box unitary oracle $\mathcal{U}$, the eigenvector matrix P will be unknown. However, an important observation is that one does not need to actually implement

the basis change operators P and $P^\dagger$ in order to exploit the above feature of the spectral decomposition – that the t^{th} power of a unitary is equivalent to finding the t^{th} power of each eigenvalue and multiplying each associated eigenspace by it.

For now, let t be real number between 0 and 1.

The algorithm is described in three stages. In the first stage, approximations to the eigenvalues of $\mathcal{U}$ are calculated using an eigenvalue estimation algorithm. Initially, assume that the eigenvalues are of the form $e^{2\pi i\lambda_k}$ with $\lambda_k = \frac{\ell_k}{2^m}$ for some integer $\ell_k \in \{0, 1, \ldots, 2^m - 1\}$ (so that the effect of precision errors can initially be ignored). In the second stage, phase shifts are applied to the eigenstates of $\mathcal{U}$. The third stage uncomputes the eigenvalue estimation. Refer to Figure 6.1 for details.

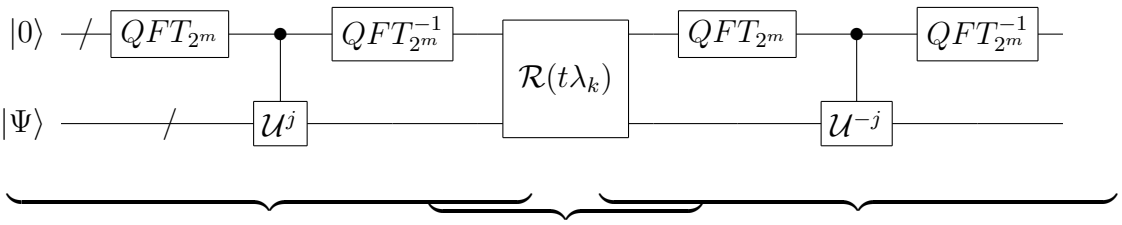


Figure 6.1: The quantum circuit diagram outlining the implementation of the algorithm that raises an unknown unitary $\mathcal{U}$ to some power t. Stage I performs eigenvalue estimation, where value l in the first register corresponds to the eigenvalue parameter estimate $\lambda = l/2^m$. Stage II then uses this eigenvalue estimate to induce a corresponding phase shift $e^{2\pi i l t/2^m}$. Stage III uncomputes the eigenvalue estimation step. In the case that the eigenvalue parameters λ_k are of the form $l_k/2^m$, the eigenvalue estimation is exact and the uncomputation step returns the control register back to the initial all-zeroes state. In general, when this assumption on the form of the eigenvalues is not valid, a slightly more complicated eigenvalue estimation algorithm is performed that gives an estimate with error at most $1/2^m$ with probability in $1 - O(1/2^m)$.

Suppose $|\Psi\rangle$ is a quantum state on which the transformation $\mathcal{U}^t$ is to be performed. Note that any state $|\Psi\rangle$ is a superposition of eigenvectors, $|\psi_k\rangle$, of $\mathcal{U}$, $|\Psi\rangle = \sum_k \alpha_k |\psi_k\rangle$. Prepare m ancilla qubits in the "all-zeros" state, where m is a user chosen parameter that characterizes the complexity of and errors in the algorithm; its role is discussed in more detail in Section 6.5. These ancilla are input

to the Quantum Fourier Transform (QFT) circuit [KLM07] (or equivalently, since applied to the all-zeros state, a circuit consisting of only Hadamard operations on each qubit) and later they are used as the controls for controlled-$\mathcal{U}$ operators, while the state $|\Psi\rangle$ is the target. The notation $c-\mathcal{U}^j$ denotes the operation $|j\rangle\,|\phi\rangle \mapsto |j\rangle\,\mathcal{U}^j\,|\phi\rangle$. Using the eigenstate expansion of $|\Psi\rangle$ in the target, it can be shown that such a controlled operation causes a phase kick-back to the control register, so that the total state becomes $\sum_k \sum_j \alpha_k e^{2\pi i \lambda_k}\,|j\rangle\,|\psi_k\rangle$. Next, an inverse QFT operation is performed on the ancilla register. The result left on the ancilla register are estimates of the eigenvalue parameter, λ_k, specifically $\ell_k = 2^m\lambda_k$, in superposition (as dictated by the input vector $|\Psi\rangle$). For a more detailed description of the eigenvalue estimation algorithm see [KLM07].

For stage two, the controlled operation $c-\mathcal{R}(2\pi\ell t/2^m)$ is applied to induce the phase shift $e^{2\pi i\ell t/2^m}$ when the eigenvalue parameter estimate ℓ is in the control register. This computation corresponds to the exponentiation of the diagonal matrix in the spectral decomposition formula. There are a few natural ways to implement this step, as shown in [CEMM98] or [Zal98].

Finally, the state of the control register (containing the ℓ_k values) is uncomputed back to the all-zeros (input ancilla) state. This completes the third and final stage of the computation.

This leaves the registers in the final state $|\mathbf{0}\rangle \otimes \sum_k \alpha_k e^{2\pi i\lambda_k t}\,|\psi_k\rangle$, which is simply the result of the application of $\mathcal{U}^t$ to the state $|\Psi\rangle$. When the assumption that the eigenvalue parameters λ_k are exact ratios of some integers ℓ_k divided by 2^m is dropped, in general precision errors will appear in the eigenvalue estimation. This imprecision in the resolution of λ_k means the phases applied to the state are not exact, which further implies that the uncomputation step does not return the ancilla register precisely to all zeroes. Still, by a proper choice of parameters, and a "gap" assumption discussed below, the errors can be managed—in particular, it is proven that they are exponentially small in the parameter m. This error reduction is done by choosing an eigenvalue estimation algorithm that outputs an estimate with error at most $\frac{1}{2^m}$ with probability in $1-O(\frac{1}{2^m})$. This can be done with $O(m)$ repetitions of the standard QFT-based eigenvalue estimation algorithm [KLM07] and applying the Chernoff bound. This uses $O(m2^m)$ calls to the black box. Relevant calculations may be found in Section 6.5.

The algorithm closely parallels a strategy employed by Klappenecker and Rötteler in [KR03], where they address the different problem of how to simulate a known unitary matrix A given other operators that generate an algebra containing A.

They conclude that their method is related to the Eigenvalue Estimation algorithm for a special case of their work. They, however, consider the case that $\mathcal{U}$ has a known efficient quantum circuit, and so, do not consider errors in their algorithm. Here, we wish to consider the case that $\mathcal{U}$ is unknown and we are attempting to simulate some function of it to a particular accuracy, with a certain number of calls. Errors will occur as a matter of course in our case since in general we cannot know *a priori* what the order of $\mathcal{U}$ is. In cases where we can, errors can be eliminated, and in particular, in the case of $\mathcal{U}$ being the QFT, considered in Section 6.10 the order is only four. This special case was also considered in [KR03] where they reach the same conclusions as presented here.

6.4.1 Underlying Assumptions

As formulated in the previous section, the algorithm requires in addition to the black box implementation of $\mathcal{U}$ itself, that access is also given to a black box implementing controlled-$\mathcal{U}$ and a black box implementing $\mathcal{U}^{-1}$. It is also necessary to assume, as discussed in more detail in the following section, that spectrum of matrix $\mathcal{U}$ has a small gap, in particular, that $\mathcal{U}$ has no eigenvalue $\lambda = e^{i\phi}$, where $\phi \in (2\pi(1 - g),\ 2\pi)$ for some value g. For simplicity, assume that $g \geq \frac{1}{2^m}$; otherwise, it would be necessary to replace some of the $O(2^m)$ terms with $O(\max\{2^m, \frac{1}{g}\})$ since the eigenvalue estimation must be done to within a higher precision than the size of the gap size with high probability. In this case, an approximation with error in $O(\frac{1}{2^m})$ is achieved. Note that there are no other assumptions about gaps elsewhere in the spectrum. Furthermore, an assumption of this form is essentially necessary in the case of worst-case complexity (as is explained in Section 6.4.2). Average-case performance is discussed more in Section 6.8.

Not all of the above restrictions are necessary, but they facilitate a clear and transparent description of the algorithm and its analysis. In some cases of interest, the individual restrictions may be dropped. In particular, in cases where only the black box implementation of $\mathcal{U}$ is given (and not the controlled-$\mathcal{U}$), one can exploit the technique for implementing a controlled-$e^{-i\phi}U$, where $e^{i\phi}$ is a random eigenvalue of $\mathcal{U}$, and use it in the algorithm to find arbitrary powers, as outlined in [Kit95] (see Section 6.6 for details). In some special cases, it is possible to remove the assumption of having a black box for $\mathcal{U}^{-1}$. For example, for non-integer values of $t \geq 2^m$, $\mathcal{U}^t$ can be approximated with error in $O(\frac{1}{t})$ using $O(t)$ calls to the controlled-$\mathcal{U}$, without any use of a black box for $\mathcal{U}^{-1}$ (see Section 6.7). Note that the above restriction on the form of the spectrum is naturally satisfied

in many practical unitaries, including the quantum Fourier transform, standard oracles $|j\rangle \mapsto (-1)^{f(j)}|j\rangle$ for computing Boolean functions $f(\cdot)$, and oracles such as those implicitly used in the adiabatic algorithms of Farhi *et al.* [FGG00] that reversibly compute a function with a finite discrete spectrum.

It is also worth noting that, for non-integer values of the power t, $\mathcal{U}^t$ is not unique; there are many operators that one might reasonably call the t^{th} power of $\mathcal{U}$. In other words, for any given $\mathcal{U}$, there are many Hamiltonians H such that $\mathcal{U} = e^{-iH}$, and thus one could naturally define $\mathcal{U}^t$ as e^{-iHt} for any one of these Hamiltonians. For instance, consider $t = \frac{1}{2}$ and the identity matrix $\mathbb{I} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. We can find all unitary matrices B such that $BB = \mathbb{I}$ where each such B is a possible square root of $\mathbb{I}$ and this is instructive for providing intuition for the general case. Firstly, matrices $\mathbb{I}$ and $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ are both square roots of $\mathbb{I}$. This happens because there are two possible square roots of the complex number 1, which is the eigenvalue of $\mathbb{I}$. In addition, each eigenspace of a unitary matrix may be broken into subspaces such that different square roots of the corresponding eigenvalue may be used to construct square roots of a given matrix. In our case, this helps to construct two more square roots of the form $\pm\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Furthermore, every matrix of the form $\pm\begin{pmatrix} \cos a & e^{b}\sin a \\ e^{-b}\sin a & -\cos a \end{pmatrix}$, and its transpose, where a and b are real valued parameters is a square root of the operator $\mathbb{I}$. For the choice of parameters $a = \frac{\pi}{2}$ and $b = 0$ this allows to construct root $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, also known as the Pauli-X matrix. Roots of the identity matrix in higher dimensions get more complicated. Thus, the problem of finding fractional powers is not completely straightforward to define. In this solution, fractional powers are constructed based on the spectral decomposition, choosing the primitive complex fractional power of every eigenvalue (though, in principle our method allows us to break the eigenspaces into subspaces and take different roots of the eigenvalue for each subspace). Thus, in a sense, the fractional power constructed is the most natural one, and also, it is the fractional power of a matrix most commonly defined in linear algebra textbooks. With the above example, our algorithm implements the $\mathbb{I}$ root of $\mathbb{I}$, though modifications are possible that would allow us to explore a wider range of square roots of $\mathbb{I}$.

6.4.2 Optimality and the Necessity of a Gap Assumption in the General Case

This algorithm is optimal in terms of the number of calls made to $\mathcal{U}$ in order to compute $\mathcal{U}^t$. It is not possible to compute $\mathcal{U}^{\frac{1}{2}}$ with constant error with fewer

than $\Omega(2^m)$ calls to $\mathcal{U}$, for a general black box, but with a spectral gap of size $g = \frac{1}{2^m}$. If $\mathcal{U}^{\frac{1}{2}}$ can be calculated with fewer calls then this method could be used to distinguish a pair of unitaries $\mathcal{U} = \mathbb{I}$ and $\mathcal{U} = \mathcal{U}_{1-1/2^m} = |0\rangle\langle 0| + e^{-2\pi i/2^m}|1\rangle\langle 1|$ that are close in the trace norm simply by effectively "amplifying" their difference. In particular, consider the square root of these two unitaries. The square root of the identity matrix given by this method is the identity, and the square root of $\mathcal{U}_{1-1/2^m}$ is $|0\rangle\langle 0| + e^{2\pi i(2^m-1)/2^m}|1\rangle\langle 1| \approx |0\rangle\langle 0| - |1\rangle\langle 1|$. These two unitaries can be distinguished with high probability with only one call.

The following simple lemma show that $\Omega(2^m)$ calls to $\mathcal{U}$ are necessary to correctly guess which $\mathcal{U}$ we were given with probability greater than $\frac{2}{3}$.

Lemma 6.1. *Suppose we are given a black box $\mathcal{U}$ that implements either $\mathbb{I}$ or $\mathcal{U}_{1-\delta} = |0\rangle\langle 0| + e^{-i\delta}|1\rangle\langle 1|$, for some small $\delta > 0$. Any algorithm that correctly guesses the identity of $\mathcal{U}$ with probability at least $\frac{2}{3}$, for any prior distribution of the two possible values of $\mathcal{U}$, must make $\Omega(\frac{1}{\delta})$ evaluations of $\mathcal{U}$.*

A short proof of this lemma is given in Appendix B.

Since the root $\mathcal{U}_{1-1/2^m}^{\frac{1}{2}}$ acting on the state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ is almost perfectly distinguishable from $\mathbb{I}^{\frac{1}{2}}$, the power finding algorithm must take $\Omega(2^m)$ calls to $\mathcal{U}$ in order to compute $\mathcal{U}^{\frac{1}{2}}$ with (small) constant error. The same holds for any power $t \in (0, 1)$ that is bounded away from from both 0 and 1.

Thus, while more efficient ways to compute fractional powers of $\mathcal{U}$ are possible in special cases, in general one cannot do much better than what has been described here. In the case when $\frac{1}{2^m} < g < \pi$, it is still possible to obtain a lower bound of $\Omega(2^m)$ queries by letting $\mathcal{U}$ be either Z or $|0\rangle\langle 0| - e^{-2\pi i/2^m}|1\rangle\langle 1|$, and computing square roots of $\mathcal{U}^2$ as a means of distinguishing the two cases.

This same argument also shows why a gap assumption is necessary. Without the gap assumption, there is no upper bound on the worst-case complexity of computing the square root of a black box unitary, since such a black box could be used to distinguish $\mathbb{I}$ from $\mathcal{U}_{1-\delta} = |0\rangle\langle 0| + e^{-i\delta}|1\rangle\langle 1|$ for arbitarily small values of δ.

As pointed out earlier, if we have a known gap g without a promise that $g \in \Omega(\epsilon) = \Omega(1/2^m)$, then a complexity of $O(\frac{1}{\epsilon}\log\frac{1}{\epsilon})$ suffices, and the example above shows that $\Omega(\frac{1}{g})$ queries are necessary. If g is unknown then there are several cases one might consider. If we are promised a lower bound $0 < g_{\min} \leq g$, then in the worst case, we must use $O(\frac{1}{g_{\min}}\log\frac{1}{\epsilon})$ calls to the unitary. If instead we have an means of reliably recognizing when the algorithm has succeeded, then trying a

sequence of lower bounds that decreases exponentially until the algorithm succeeds will also give a complexity in $O(\frac{1}{\min\{g,\epsilon\}} \log \frac{1}{\epsilon})$, which is the same complexity as when g is known. If, however, we do not know g and have no means of recognizing when the algorithm has succeeded, then no upper bound is possible.

6.5 Complexity and Error Analysis

Here the upper bounds on the precision of the algorithm are found, which are constrained by the lack of perfect precision in computing the eigenvalue estimates. This analysis is for $t = \frac{1}{2}$, however the same final bound applies for any $t \in (0, 1)$. This is because for t in this range, $t = \frac{1}{2}$ gives the largest error. This can be seen visually in Figure 6.2.

Stage 1 computes a near-optimal eigenvalue estimation (in the sense of nearly optimizing the chance of obtaining an estimate with error at most $\frac{1}{2^{2m}}$). This is done by repeating the "standard" eigenvalue estimation algorithm [CEMM98] a total of $r \in O(m)$ times and take the majority answer, with the constant chosen so that each eigenvalue estimate $e^{2\pi i\widetilde{\lambda_k}}$ of $e^{2\pi i\lambda_k}$ satisfies $|e^{2\pi i\widetilde{\lambda_k}} - e^{2\pi i\lambda_k}| \leq \frac{1}{2^m}$ with probability in $1 - O(\frac{1}{2^{2m}})$. One can fine-tune the optimal eigenvalue estimation procedure further (*e.g.*, [vDDE+07]), however, this procedure is sufficient.

This eigenvalue estimation procedure is described by the transformation

$$|0\rangle \sum_k \alpha_k |\psi_k\rangle \mapsto \sum_k \alpha_k \left(\sum_{\mathbf{y}} \beta_{\mathbf{y},k} |\mathbf{y}\rangle |\lambda_{\mathbf{y}}\rangle \right) |\psi_k\rangle \tag{6.2}$$

where $\beta_{\mathbf{y},k} = \frac{1}{2^m} \prod_{i=1}^{r} \sum_{x=0}^{2^m-1} e^{2\pi i x(\lambda_k - \frac{y_i}{2^m})}$ and the values $\mathbf{y} \in \mathbb{Z}_{2^m}^r$ (where $\mathbb{Z}_{2^m} = \{0, 1, \ldots, 2^m - 1\}$) are r-tuples of integers $\mathbf{y} = (y_1, y_2, ...y_r)$, and $\lambda_{\mathbf{y}}$ is the value obtained by taking the most commonly occurring integer y_{mode} in the r-tuple of integers $\mathbf{y}$ (with some convention for breaking ties) and letting $\lambda_{\mathbf{y}} = y_{\text{mode}}/2^m$.

In Stage II, the intention is to enact the following map, which applies the appropriate eigenvalue phases to each eigenvector:

$$\sum_k \alpha_k \left(\sum_{\mathbf{y}} \beta_{\mathbf{y},k} |\mathbf{y}\rangle |\lambda_{\mathbf{y}}\rangle \right) |\psi_k\rangle \mapsto \sum_k \alpha_k \left(\sum_{\mathbf{y}} \beta_{\mathbf{y},k} |\mathbf{y}\rangle |\lambda_{\mathbf{y}}\rangle \right) e^{2\pi i\lambda_k/2} |\psi_k\rangle , \tag{6.3}$$

but we actually implement the unitary

$$\sum_k \alpha_k \left(\sum_{\mathbf{y}} \beta_{\mathbf{y},k} |\mathbf{y}\rangle |\lambda_{\mathbf{y}}\rangle \right) |\psi_k\rangle \mapsto \sum_k \alpha_k \sum_{\mathbf{y}} \beta_{\mathbf{y},k} |\mathbf{y}\rangle |\lambda_{\mathbf{y}}\rangle e^{2\pi i \lambda_{\mathbf{y}}/2} |\psi_k\rangle . \tag{6.4}$$

Then Stage III uncomputes the eigenvalue estimation, but imperfectly since in Stage II, there is coupling of the control registers with the target registers, leaving not the ideal output,

$$\sum_k \alpha_k |00\ldots0\rangle e^{2\pi i \lambda_k/2} |\psi_k\rangle , \tag{6.5}$$

but instead, as state close to it (assuming the eigenvalue gap):

$$\sum_k \sum_{x'} \alpha_k |x'\rangle f(k,x') e^{2\pi i \lambda_{\mathbf{y}}} |\psi_k\rangle , \tag{6.6}$$

where $f(k,x') = \frac{1}{2^{2m}} \sum_x e^{2\pi i x(\lambda_k - \frac{y}{2^m})} \frac{1}{2^m} \sum_{x'} e^{2\pi i x'(\frac{y}{2^m} - \lambda_k)}$. Post-selecting on the outcome $x' = 0$ gives $f(k,0) = \frac{1}{2^{2m}} \frac{\sin^2(2^m \pi \delta)}{\sin^2(\pi\delta)}$, which approaches a delta function as $m \to \infty$.

The probability that $|e^{2\pi i \lambda_{\mathbf{y}}/2} - e^{2\pi i \lambda_k/2}| > \frac{1}{2^m}$ is in $O(\frac{1}{2^m})$ because the probability that an individual eigenvalue estimate λ_y is within $\frac{1}{2^m}$ of the actual eigenvalue λ_k is $\frac{8}{\pi^2}$ [KLM07]. Since this is greater than $\frac{1}{2}$, repeating the algorithm r times and applying the Chernoff bound improves the probability of the estimate being this accurate to $O(\frac{1}{2^m})$.

So it makes sense to write the state in Stage II as

$$|\Phi\rangle = \sum_k \alpha_k \sum_{\mathbf{y}\in\mathbf{S}} \beta_{\mathbf{y},k} |\mathbf{y}\rangle |\lambda_{\mathbf{y}}\rangle e^{2\pi i \lambda_{\mathbf{y}}/2} |\psi_k\rangle + \sum_k \alpha_k \sum_{\mathbf{y}\in\mathbf{S}'} \beta_{\mathbf{y},k} |\mathbf{y}\rangle |\lambda_{\mathbf{y}}\rangle e^{2\pi i \lambda_{\mathbf{y}}/2} |\psi_k\rangle , \tag{6.7}$$

where the values of $\mathbf{y} \in \mathbf{S}$ include the "good" values of $\mathbf{y}$ that produce estimates $\lambda_{\mathbf{y}}$ satisfying $|e^{2\pi i \lambda_{\mathbf{y}}} - e^{2\pi i \lambda_k}| \leq \frac{1}{2^m}$, and the "bad" values, $\mathbf{y} \in \mathbf{S}'$, are the rest. The norm of the bad part of the state is in $O(\frac{1}{2^{2m}})$ and the norm of the "good" part is $1 - O(\frac{1}{2^{2m}})$.

Since the algorithm is unitary, the errors will propagate linearly through Stage III and the sets S and S' will not mix. Therefore, consider the error introduced in Stage II. Let Λ denote the ideal phase shift operator, and $\tilde{\Lambda}$ be the actual operator.

Thus

$$\Lambda |\Phi\rangle - \tilde{\Lambda} |\Phi\rangle = \sum_k \alpha_k \sum_{\mathbf{y}\in S} \beta_{\mathbf{y},k} |\mathbf{y}\rangle |\lambda_{\mathbf{y}}\rangle (e^{2\pi i\lambda_k/2} - e^{2\pi i\lambda_{\mathbf{y}}/2}) |\psi_k\rangle + \sum_k \alpha_k \sum_{\mathbf{y}\in \mathbf{S}'} \beta_{\mathbf{y},k} |\mathbf{y}\rangle |\lambda_{\mathbf{y}}\rangle (e^{2\pi i\lambda_k/2} - e^{2\pi i\lambda_{\mathbf{y}}/2}) |\psi_k\rangle .$$

The norm of this difference can be bounded by the triangle inequality

$$||\Lambda |\Phi\rangle - \tilde{\Lambda} |\Phi\rangle ||^2 \leq \sum_k \sum_{\mathbf{y}\in\mathbf{S}} |\alpha_k\beta_{\mathbf{y},k}|^2 |\delta_{k,\mathbf{y}}|^2 + \sum_k \sum_{\mathbf{y}\in\mathbf{S}'} |\alpha_k\beta_{\mathbf{y},k}|^2 |\delta_{k,\mathbf{y}}|^2, \tag{6.8}$$

where $|\delta_{k,\mathbf{y}}|^2 = |e^{i\lambda_k/2} - e^{i\lambda_{\mathbf{y}}/2}|^2$.

Noting that $\sum_k |\alpha_k|^2 = 1$, from the previous Chernoff-bounding argument, $\sum_{\mathbf{y}\in\mathbf{S}} \sum_k |\alpha_k\beta_{\mathbf{y},k}|^2 \in 1 - O(\frac{1}{2^{2m}})$, $\sum_{\mathbf{y}\in\mathbf{S}'} \sum_k |\alpha_k\beta_{\mathbf{y},k}|^2 \in O(\frac{1}{2^{2m}})$, and that $|\delta_{k,\mathbf{y}}|^2 \in O(\frac{1}{2^{2m}})$ for the good values, and is at most 1 for the bad values.

Thus

$$||\Lambda |\Phi\rangle - \tilde{\Lambda} |\Phi\rangle ||^2 \in O\left(\frac{1}{2^{2m}}\right). \tag{6.9}$$

Noticing that

$$|| |u\rangle - |v\rangle ||_2^2 = (\langle u| - \langle v|)(|u\rangle - |v\rangle) \geq 2(1 - |\langle u|v\rangle|) \tag{6.10}$$

and using the equality

$$|| |u\rangle \langle u| - |v\rangle \langle v| ||_{\mathrm{Tr}} = 2\sqrt{1 - |\langle u|v\rangle|^2}, \tag{6.11}$$

which implies that

$$|| |u\rangle \langle u| - |v\rangle \langle v| ||_{\mathrm{Tr}} \leq 2|| |u\rangle - |v\rangle ||_2, \tag{6.12}$$

so $||\Lambda(|\Phi\rangle \langle\Phi|)\Lambda - \tilde{\Lambda}(|\Phi\rangle \langle\Phi|)\tilde{\Lambda}||_{\mathrm{Tr}} \in O(\frac{1}{2^m})$. The trace norm of the difference between the two unitary superoperators is simply given by taking the maximization over states $|\Phi\rangle$. Since this holds for any state $|\Phi\rangle$, we have

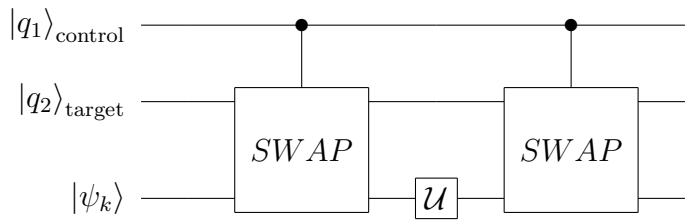
$$||\Lambda - \tilde{\Lambda}||_{\mathrm{Tr}} \in O\left(\frac{1}{2^m}\right). \tag{6.13}$$

This is the case after Stage II of the algorithm. Note that Stage III is a unitary operation that is the same for both the ideal and the actual case. The trace norm is

invariant under unitary transformations, so the trace norm of the unitary operators for the full ideal and actual algorithms have the same bound. In the case of unitary operators this is equivalent to the diamond norm.

6.6 Controlled Unitaries

Note that the eigenvalue estimation algorithm requires the use of a controlled unitary $c-\mathcal{U}$. If an eigenvector of $\mathcal{U}$ with eigenvalue $+1$ is provided, then this transformation can be implemented using a controlled-SWAP (equivalently, a series of Fredkin gates [FT82]), as was shown by Kitaev [Kit95] and is illustrated in the following circuit diagram:



It is easy to verify that the above circuit implements

$$|0\rangle_{\text{control}} |q_2\rangle_{\text{target}} |\psi_k\rangle \mapsto e^{2\pi i\lambda_k} |0\rangle_{\text{control}} |q_2\rangle_{\text{target}} |\psi_k\rangle$$

$$|1\rangle_{\text{control}} |q_2\rangle_{\text{target}} |\psi_k\rangle \mapsto |1\rangle_{\text{control}} (\mathcal{U} |q_2\rangle_{\text{target}}) |\psi_k\rangle .$$

In other words, up to a global phase, this implements the controlled-$(e^{-2\pi i\lambda_k}\mathcal{U})$. If $\lambda_k = 0$, *i.e.*, $|\psi_k\rangle$ has eigenvalue $+1$, then the $c-\mathcal{U}$ has been implemented. Otherwise, one obtains a very similar operation, which may or may not be sufficient, depending on the application. If, for example, any $c-(e^{i\phi}\mathcal{U})$ will suffice, that is, $(e^{i\phi}\mathcal{U})^t$ is what is calculated for some arbitrary global phase $e^{i\phi}$, then as long as the same one is used consistently, this is easy to achieve.

It is not necessary to assume that the eigenvectors $\{|\psi_k\rangle\}$ of $\mathcal{U}$ are known or that a specific eigenvector is given. Instead, any state will do. For simplicity, consider the completely mixed state, which can be decomposed as an equal mixture over the eigenstates. Provided the state that is in the target register is kept, and the resulting state is re-used every time a controlled-$\mathcal{U}$ is constructed between these two registers, the phase (though random) will be the same for the entire computation. This is equivalent to implementing a $c-(e^{i\phi}\mathcal{U})$ consistently throughout the computation,

where ϕ is an unknown random value (whose distribution depends on the weight of the eigenspace of $e^{-i\phi}$ in the initial mixed state). In essence, the target system serves as a phase reference.

6.7 Inverses

The algorithm for generating $\mathcal{U}^t$ uses $\mathcal{U}^{-1}$ when the eigenvalue estimates need to be uncomputed. More specifically, is starts with $|0\rangle|\psi_k\rangle$, and $(QFT_{2^m}^{-1} \otimes \mathbb{I})c - \mathcal{U}^j(QFT_{2^m} \otimes \mathbb{I})$ is applied to compute

$$\sum_j \left(QFT_{2^m}^{-1} \frac{e^{2\pi i j \lambda_k}}{\sqrt{2^m}} |j\rangle \right) |\psi_k\rangle . \tag{6.14}$$

Then, the values in the first register are used to control the phase shift corresponding to $|\psi_k\rangle$. For simplicity, assume that no phase shift is performed (Stage II is omitted), but we still perform Stage III to uncompute the eigenvalue estimates. This is trivial if access is provided to $c-\mathcal{U}^{-1}$ (which can be implemented with a $c-\mathcal{U}$ and $\mathcal{U}^{-1}$), since we just apply $(QFT_{2^m}^{-1} \otimes \mathbb{I})c - \mathcal{U}^{-j}(QFT_{2^m} \otimes \mathbb{I})$ to get back $|0\rangle|\psi_k\rangle$.

However, in some cases, the $\mathcal{U}^{-1}$ operations are not necessary. The unitary $\mathcal{U}^{-1}$ could be simulated by applying $(QFT_{2^m} \otimes \mathbb{I})c - \mathcal{U}^j(QFT_{2^m}^{-1} \otimes \mathbb{I})$ to yield

$$\begin{aligned}(QFT_{2^m} \otimes \mathbb{I})c - \mathcal{U}^j(QFT_{2^m}^{-1} \otimes \mathbb{I})(QFT_{2^m}^{-1} \otimes \mathbb{I})c - \mathcal{U}^j(QFT_{2^m} \otimes \mathbb{I})|0\rangle|\psi_k\rangle \\ = (QFT_{2^m} \otimes \mathbb{I})c - \mathcal{U}^j(M \otimes \mathbb{I})c - \mathcal{U}^j(QFT_{2^m} \otimes \mathbb{I})|0\rangle|\psi_k\rangle \end{aligned} \tag{6.15}$$

where $M = QFT_{2^m}^{-2}$, which implies $M|x\rangle = |2^m - x \bmod 2^m\rangle$.

It is easy to verify that $(M \otimes \mathbb{I})c - \mathcal{U}^j(M \otimes \mathbb{I}) = c - \mathcal{U}^{2^m - j \bmod 2^m}$, and thus the above state equals

$$\begin{aligned}= (QFT_{2^m} \otimes \mathbb{I})(M \otimes \mathbb{I})c - \mathcal{U}^{2^m - j \bmod 2^m}(c - \mathcal{U}^j(QFT_{2^m} \otimes \mathbb{I})|0\rangle|\psi_k\rangle \\ = (QFT_{2^m}^{-1} \otimes \mathbb{I})c - \mathcal{U}^{2^m}(QFT_{2^m} \otimes \mathbb{I})|0\rangle|\psi_k\rangle . \end{aligned} \tag{6.16}$$

Note that since the $c-\mathcal{U}^{2^m}$ no longer depends on the value of the first register, it can be commuted through the QFT operations to get

$$\begin{aligned} c - \mathcal{U}^{2^m}(QFT_{2^m}^{-1} \otimes \mathbb{I})(QFT_{2^m} \otimes \mathbb{I})|0\rangle|\psi_k\rangle &= c - \mathcal{U}^{2^m}|0\rangle|\psi_k\rangle & (6.17) \\ &= e^{2\pi i 2^m \lambda_k}|0\rangle|\psi_k\rangle . & (6.18) \end{aligned}$$

This phase factor of $e^{2\pi i 2^m \lambda_k}$ is in general a problem when the target register is in a superposition of eigenstates, since different eigenstates will pick up a different phase factor.

However, in some cases, it is not a problem. One observation, is that such a transformation is equivalent to applying $\mathcal{U}^{2^m}$. Thus, if the goal is to apply $\mathcal{U}^t$ for $t = \lfloor t \rfloor + \delta$, $0 \leq \delta < 1$, where $t \geq 2^m$, the algorithm for implementing $\mathcal{U}^\delta$ can be applied, but instead with the above modification. This yields an approximation to $\mathcal{U}^{2^m+\delta}$ with error in $O(\frac{1}{2^m})$ without the use of inverse $\mathcal{U}$ operations. Then $\mathcal{U}^{t-2^m}$ can be applied to complete the approximation of $\mathcal{U}^t$ with error in $O(\frac{1}{t}) \subseteq O(\frac{1}{2^m})$.

Also note that, if $\lambda_k = \frac{l_k}{2^m}$ for an integer l_k, then the phase factors $e^{2\pi i 2^m \lambda_k} = 1$ all equal 1.

This problem can also be remedied in any other cases where λ_k can be determined exactly (or much more precisely than error $\frac{1}{2^m}$), since before uncomputing λ_k, one can add an additional phase shift of $e^{-2\pi i 2^m \lambda_k}$ conditional on the value of λ_k. This will eliminate the final unwanted phase factor of $e^{2\pi i 2^m \lambda_k}$ associated with the eigenvector $|\psi_k\rangle$.

6.8 Coping with No Gap Assumption

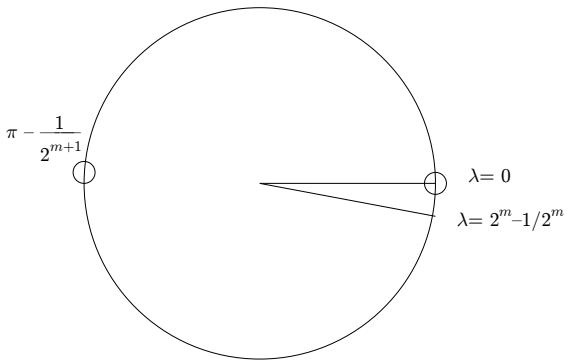


Figure 6.2: An illustration of the ambiguity of dividing a phase angle that is between $2\pi i \frac{2^m-1}{2^m}$ and 0. In the first case, dividing by 2 changes the phase by nearly π and in the second case, dividing by 2 leaves the state unchanged. The circles show the regions that each of the two eigenvalues are mapped to by the operation. Note that they are on nearly opposite sides of the circle.

In general, the gap between the largest and smallest eigenvalues modulo 2π could be arbitrarily small. In such cases there is no value of ϕ where $\phi \in (2\pi(1-g),\ 2\pi)$

for $g \geq \frac{1}{2^m}$, or if the available resources are limited so that $g \leq \frac{1}{2^m}$ for the value of m chosen, then the *average-case* performance can be bounded if there is some freedom allowed when defining the root of $\mathcal{U}$.

If one is only interested in the average-case performance (*e.g.*, averaging over all possible input states to $\mathcal{U}^{\frac{1}{2}}$ according to the Haar measure), one is still faced with the potential problem that most or all of the spectrum of $\mathcal{U}$ might be located in the difficult region (see Section 6.4.2) of $e^{i\phi} \approx 1$, $\pi < \phi \leq 2\pi$. But, if all that is required is a square root of $e^{i\theta}\mathcal{U}$ for a random $\theta \in [0, 2\pi)$, then the average-case error (for any input distribution) is still in $O(\frac{1}{2^m})$ for an algorithm that only uses the square root of $e^{i\theta}\mathcal{U}$ simulated by the algorithm once.

For an algorithm that uses $(e^{i\theta}\mathcal{U})^{\frac{1}{2}}$ a total of k times, the error could get magnified to $O(\frac{k^2}{2^m})$. Thus, $m' = m + \lceil 2 \log k \rceil$ could be selected for each simulation of $(e^{i\theta}\mathcal{U})^{\frac{1}{2}}$.

To illustrate why the factor in the error is k^2 instead of k, consider the following algorithm where substituting an exact square root with the approximate square root exhibits a Grover-search-type amplitude amplification and thereby yields an error in $\Omega(\frac{k^2}{2^m})$. Consider an operator $\mathcal{U}$ that has every 2^mth root of unity $e^{2\pi i \ell / 2^m}$ as an eigenvalue with equal multiplicities. Consider an algorithm that does eigenvalue estimation on the identity operator with precision in $\Theta(1/2^m)$, and uses quantum searching to find an eigenvector with eigenvalue -1. Clearly, this algorithm should not find such an eigenvalue since all the eigenvalues are $+1$. Contrast this with the situation that occurs when $U_1 = (e^{i\theta}\mathcal{U})^{-\frac{1}{2}}$ and let U_2 be our approximation to $(e^{i\theta}\mathcal{U})^{\frac{1}{2}}$ for the same θ. (For simplicity, assume it is unitary, even though in practice it will be a map that is almost unitary.) The value θ can be expressed in the form $2\pi\ell/2^m - \epsilon$, for some integer ℓ and non-negative value $\epsilon \leq 1/2^m$. Note that $U_1(e^{i\theta}\mathcal{U})^{\frac{1}{2}} = I$, but $U_1 U_2$ will have most eigenvalues near $+1$, and proportion $1/2^m$ of the eigenvalues near -1. A -1 eigenvalue occurs with probability greater than $O(\frac{1}{2^m})$ if an eigenvalue phase being estimated falls in the range $2\pi\left(1 - \frac{3}{2^{m+2}}\right) < \lambda_k < 2\pi$. If $\lambda_k \leq \frac{3}{2^{m+2}}$, the probability that an eigenvalue which should be mapped by the square-root operation to a value near π will instead be estimated to be 0 and mapped to 0 is, for a single eigenvalue estimation, less than one half. Repeating this m times reduces this probability to $O\left(\frac{1}{2^m}\right)$ by the Chernoff bound. Because θ is distributed randomly in the range $[0, 2\pi)$, the error happens to a fraction of the eigenvalues

$$\frac{1}{2\pi}\left(2\pi - 2\pi\left(1 - \frac{3}{2^{m+2}}\right)\right) = O\left(\frac{1}{2^m}\right). \tag{6.19}$$

Thus, a generalized version of quantum searching given in the next section will

generate these near -1 eigenvectors with an amplitude in $O(\frac{k}{\sqrt{2^m}})$ using only k evaluations of U_1U_2. Recall that with the ideal $(e^{i\theta}\mathcal{U})^{\frac{1}{2}}$ (instead of U_2), there would be no such eigenvectors for the search algorithm to find, thus the part of the wave function that found these -1 eigenvalues is an error. This implies an error with probability in $O(\frac{k^2}{2^m})$.

6.8.1 A Generalization of Quantum Search

In the quantum search algorithm, the first step is to take the initial state, usually denoted $|000...0\rangle$ and map it to a state that has an inner product of size $\frac{1}{\sqrt{N}}$ with the state being searched for. In the original presentation, the elements being searched over are computational basis states (binary strings) so, for example, the operator that performs a single qubit Hadamard on each qubit will do this. If this inner product is much smaller than $\frac{1}{\sqrt{N}}$ the algorithm will not run in $O(\sqrt{N})$ calls to the oracle. If is is zero, the algorithm will fail.

In cases where it is not promised that the solution is a computational basis state, ensuring that the algorithm proceeds using an operation which creates overlap with the solution state requires additional thought. Here an appropriate operator for this task, A, is given.

Suppose a black box O_ϕ is given that flags an unknown quantum state $|\phi\rangle \in H_N$. In other words, $O_\phi |\phi\rangle |b\rangle = |\phi\rangle |b \oplus 1\rangle$ and $O_\phi |\phi'\rangle |b\rangle = |\phi'\rangle |b\rangle$ for any $|\phi'\rangle$ orthogonal to $|\phi\rangle$. For brevity, let us directly use the phase flip version of this black box $O_\phi |\phi\rangle = -|\phi\rangle$ and $O_\phi |\phi'\rangle = |\phi'\rangle$ for $\langle\phi'|\phi\rangle = 0$ (see Section 8.1 of [KLM07] for a more detailed discussion of th.e relationship between these black boxes). This is the oracle that "marks" the solution state.

If, as in ordinary quantum searching, one is given a unitary operation A that maps $|00\ldots0\rangle \mapsto \sin(\theta)|\phi\rangle + \cos(\theta)|\phi'\rangle$ where $\langle\phi|\,\phi'\rangle = 0$, then it is possible immediately to define the slight generalization of a quantum search iterate,

$$Q = -AU_{00\ldots0}A^{-1}O_\phi, \tag{6.20}$$

where $U_{00\ldots0} = I - 2|00\ldots0\rangle\langle00\ldots0|$ and note that $Q^k|00\ldots0\rangle = \cos((2k+1)\theta)|\phi\rangle + \sin((2k+1)\theta)|\phi'\rangle$. Thus, if with an algorithm A that "guesses" $|\phi\rangle$ with probability $p = \sin^2(\theta)$, then after $O(\frac{1}{\sqrt{p}})$ calls to Q, one obtains $|\phi\rangle$ with probability very close to 1.

In the standard description of the quantum searching algorithm, the marked

state is some computational basis state, so it is easy to produce a unitary with this property A such that $|\langle\phi| A |00\ldots0\rangle| = \frac{1}{\sqrt{N}}$, in particular, any operator that maps $|00\ldots0\rangle \mapsto \sum_x \frac{1}{\sqrt{N}} |x\rangle$ will work.

However, for an arbitrary quantum state $|\phi\rangle$, it is non-trivial to produce such an initial state if one is not given such an A. A maximally mixed state will have the required overlap, but since amplitude amplification requires the inverse of the preparation operation A, this is not a possibility. Generating pure states that for practical purposes are "random" (*e.g.*, [AE07]) to use as the initial state could work, however, in this case it is not necessary because there is another approach that will work.

Let A act on an N^2 dimensional Hilbert space and maps $|00\ldots0\rangle |00\ldots0\rangle \mapsto \sum_x \frac{1}{\sqrt{N}} |x\rangle |x\rangle$. Note that for any basis $\{|\phi_j\rangle\}$, this maximally entangled state can be rewritten as $\sum_j \frac{1}{\sqrt{N}} |\phi_j\rangle \left|\phi_j^*\right\rangle$ (where the coefficients of $\left|\phi_j^*\right\rangle$ are the complex conjugates of those of $|\phi\rangle$). Notice that the maximally entangled state has equal support for every eigenvector in the basis of eigenvalues of the black box unitary, thus ensuring that the required inner product with the final state is obtained. Assume, without loss of generality, that $|\phi_0\rangle = |\phi\rangle$. Applying amplitude amplification, and using $Q = -AU_{00\ldots0,0\ldots0}A^{-1}(O_\phi \otimes I)$, then it is easy to verify that

$$\begin{aligned} &Q^k |00\ldots0\rangle |00\ldots0\rangle \\ &\quad = \sin((2k+1)\theta) |\phi_0\rangle |\phi_0^*\rangle + \cos((2k+1)\theta) \left(\sum_{j=1}^{N-1} \frac{1}{\sqrt{N-1}} |\phi_j\rangle \left|\phi_j^*\right\rangle \right), \end{aligned} \quad (6.21)$$

where $\sin(\theta) = \frac{1}{\sqrt{N}}$. Thus, by choosing $k \approx \frac{\pi}{4\theta} \in O(\frac{1}{\sqrt{p}})$, one obtains $|\phi_0\rangle |\phi_0^*\rangle$ with high fidelity. For applications where the extra $|\phi^*\rangle$ system will pose a problem, this approach will not work, and techniques for generating pure states that behave like random quantum states could be used instead.

This technique also works if the black box flags a subspace of dimension d. Again, assume without loss of generality that this subspace is spanned by the vectors $|\phi_0\rangle, |\phi_1\rangle, \ldots, |\phi_{d-1}\rangle$. Then the maximally entangled state can be rewritten as

$$\sin(\theta) \left(\sum_{j=0}^{d-1} \frac{1}{\sqrt{d}} |\phi_j\rangle \left|\phi_j^*\right\rangle \right) + \cos(\theta) \left(\sum_{j=d}^{N-1} \frac{1}{\sqrt{N-d}} |\phi_j\rangle \left|\phi_j^*\right\rangle \right) \quad (6.22)$$

where $\sin^2(\theta) = \frac{d}{N}$.

Thus $Q^k |00\ldots0\rangle |00\ldots0\rangle \approx \sum_{j=0}^{d-1} \frac{1}{\sqrt{d}} |\phi_j\rangle \left|\phi_j^*\right\rangle$ for $k \approx \frac{\pi}{4}\sqrt{\frac{N}{d}}$, assuming $\sqrt{\frac{d}{N}}$ is much less than one.

6.9 Large Powers and an Example of Exponential Improvement

In general, when computing larger powers of $\mathcal{U}$ such that $t > 1$, the method proposed in the preceding sections is not the most accurate or efficient approach. Rather, it is better to first apply $\mathcal{U}$ directly a total of $\lfloor t \rfloor$ times, and then use the algorithm introduced in Section 6.4 in order to compute any remaining fractional power of $\mathcal{U}$. This is because the imprecision in the eigenvalue estimates is magnified when t is large. If, instead of the exact eigenvalue $e^{2\pi i \lambda_k}$, the algorithm finds an estimate $e^{2\pi i \widetilde{\lambda_k}}$, where $|\widetilde{\lambda_k} - \lambda_k| = \epsilon$, then mapping $|\psi_k\rangle \mapsto e^{2\pi i \widetilde{\lambda_k} t} |\psi_k\rangle$ will lead to an error in the phase parameter of size $t\epsilon$. For arbitrary $\mathcal{U}$, it takes $O(\frac{1}{\epsilon})$ calls to $\mathcal{U}$ in order to get an estimate with error at most ϵ with high probability. Thus to achieve an error of $\delta = t\epsilon$ in the application of $\mathcal{U}^t$, we need $\epsilon = \delta/t$ precision in the estimate, amounting to $O(\frac{t}{\delta})$ calls to $\mathcal{U}$, whereas applying the operation $\mathcal{U}$ directly as many times as possible only requires $\lfloor t \rfloor + O(\frac{1}{\delta})$ uses of the oracle for $\mathcal{U}$.

Somewhat counterintuitively, however, for some special conditions on the oracle $\mathcal{U}$, proceeding by determining the eigenvalues and applying the rotations in the algorithm only, with no direct applications of $\mathcal{U}$, can be more efficient. In fact, there can be values of $t > 1$ such that it is possible to determine $\mathcal{U}^t$ in even fewer than $\lfloor t \rfloor$ calls. This happens when the oracle allows for more accuracy in the eigenvalue estimates. For example, for black boxes that map $|j\rangle \mapsto (-1)^{X_j} |j\rangle$ for some $j \in \{0,1\}^n$ and $X_1 X_2 \ldots X_j \ldots X_N \in \{0,1\}^N$, the eigenvalue parameter X_j can be exactly determined with only one call to the black box. As discussed later is Section 6.10, it is easy to exactly determine the eigenvalue of an eigenvector of the quantum Fourier transform, since the QFT has order four ($QFT^4 = \mathbb{I}$). In both of these examples, the technique of the previous sections does not help is not the most efficient way of computing higher integer powers of $\mathcal{U}$ since $\mathcal{U}$ has very low order, say r, and so computing $\mathcal{U}^t$ is equivalent to computing $\mathcal{U}^{t \mod r}$, where $t \mod r$ is the unique element of $t - r\mathbb{Z}$ that lies in in the interval $[0, r)$. For the second example, $QFT^3 = QFT^7 = QFT^{11}$ *etc.*, so if $t = 4n - 1$ where n is an integer, the most efficient approach is to make three direct calls to $\mathcal{U}$. However it is possible to construct an example where $\mathcal{U}$ has very high order, but one can still compute precise estimates of its eigenvalues with relatively few calls to $\mathcal{U}$, and thus exponentiate $\mathcal{U}$ with much fewer than t calls, even for t less than the order of $\mathcal{U}$.

Suppose the eigenvalues of the desired unitary $\mathcal{U}$ are all of the form $e^{2\pi i \frac{\ell_k}{p_k}}$, for $k \in \{1, 2, \ldots, b\}$, where ℓ_k is an integer, $p_k \in \{p_1 = 2, p_2 = 3, p_3 = 5, ..., p_b | p_i =$

i^{th} prime}, and b is a natural number. Then $\mathcal{U}^B = \mathbb{I}$, where $B = \prod_{k=1}^{b} p_k$, and (assuming that each p_k occurs at least once with a non-zero l_k) no lower power of $\mathcal{U}$ is equal to the identity. If t is on the order of B, then the number of calls to the unitary required for the straightforward implementation of $\mathcal{U}^t$ is in $O(B) = O(\prod_{k=1}^{b} k \log k)$. This product is called a *primorial* [FMRC62], and it is in $O(e^{(1+o(1))b \log b}) = O(b^{(1+o(1))b})$. Suppose also that access is not given to $\mathcal{U}$ inverse in this case: since the .

Now it is demonstrated how the basic algorithm introduced in Section 6.4 may be updated to find approximate $\mathcal{U}^t$ with exponentially fewer calls to $\mathcal{U}$ than the number of calls it takes to find $\mathcal{U}^t$ via a direct application of t unitaries.

The method used for this problem employs the continued fraction algorithm, which was also used in the order-finding part of Shor's algorithm [Sho97]. The eigenvalue estimation algorithm will return an estimate $\frac{h}{2^m}$ of the eigenvalue parameter $\frac{\ell_k}{p_k}$ to within an error at most $\frac{1}{2^m}$ with probability at least $\frac{8}{\pi^2}$. The error probability can be amplified down to $\frac{1}{2^m}$ by repeating the estimation $O(m)$ times and taking the majority vote. If m is chosen such that $m \in O(\log b)$, so that $2^m > 2p_b p_{b-1}$, then the fraction $\frac{\ell_k}{p_k}$ will be the only fraction, or *convergent*, in the continued fraction expansion of $\frac{h}{2^m}$ that has distance at most $\frac{1}{2^m}$ from $\frac{h}{2^m}$ and has a denominator at most p_b.

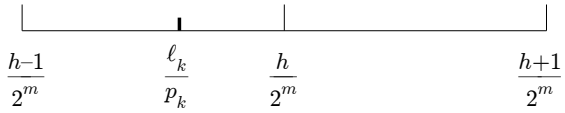


Figure 6.3: A diagram illustrating the procedure. The estimate is $\frac{h}{2^m}$. The smallest gap between any two eigenvalues is $\frac{1}{p_b p_{b-1}}$ and this must be larger than $\frac{2}{2^m}$ to ensure that only one rational fraction with denominator no greater than p_b will be found in the region $\frac{h-1}{2^m}$ to $\frac{h+1}{2^m}$.

This convergent $\frac{\ell_k}{p_k}$ can be found with $O(m)$ arithmetic operations over integer numbers with at most $\log m$ bits each (*e.g.*, no more than $O(m^3)$ binary operations using trivial algorithms for division and multiplication) using the continued fraction algorithm. Once we have the exact value of the eigenvalue parameter $\frac{\ell_k}{p_k}$, with probability $\frac{1}{2^m}$, we can correctly compute *any* positive power t of $\mathcal{U}$ with error in $O(\frac{1}{2^m})$ using $poly(b, \log t)$ calls to $\mathcal{U}$ and other elementary operations. This is done by mapping, in superposition, $|\ell_k\rangle |p_k\rangle |\psi_k\rangle \mapsto e^{2\pi i t \frac{\ell_k}{p_k} + O(\frac{1}{2^m})} |\ell_k\rangle |p_k\rangle |\psi_k\rangle$, which can be done by standard reversible computing methods and m-bit arithmetic operations

(the $O(\frac{1}{2^m})$ error is due to precision in the arithmetic operations). Then, as usual, one uncomputes the computation of $|\ell_k\rangle |p_k\rangle$. As alluded to before, this gives an overall complexity of $O(m2^m) = O(b \log b)$ to perform the operation up to error $O(\frac{1}{2^m})$. This follows from equation (6.13) in Section 6.5.

This complexity is polynomial in $\log t$ (vs $poly(t)$) because we only need to compute an m-bit approximation to $\frac{t\ell_k}{p_k}$. This also allows us to uncompute the eigenvalue estimates found in the first stage of the algorithm without ever needing to use the operation $\mathcal{U}^{-1}$. Note that t could be very large, for example $t \sim B/2 = 2^{b-1}$. In this case, we can apply $\mathcal{U}^t$ to any state of the appropriate dimension with the output state containing an error term with probability amplitude $O(\frac{1}{b})$ with only $O(b \log b) \ll t \in O(2^b)$ calls. The error probability can efficiently be made exponentially small by repeating the phase estimation part.

6.10 Fractional Quantum Fourier Transform

In classical computation, roots of the Fourier transform are used for efficient filtering of the frequency noise that is some non-constant function of a conjugate variable, such as time [Alm94]. Given the role the QFT plays in the quantum computation, and the importance of roots of Fourier transforms in classical computation, there may as well be other interesting techniques based on the fractional powers of the quantum Fourier transform.

Note that $QFT^4 = \mathbb{I}$. This symmetry of the eigenvalues allows the algorithm to be very efficient in the case $\mathcal{U} = QFT$. To find a fractional power t, $t < 1$ of the QFT using the algorithm introduced in Section 6.4, observe the fact that the QFT has four eigenvalues that are all of the form $e^{2\pi i \frac{\ell}{2^2}}$ where ℓ is an integer: $+1$, i, -1 and $-i$. This means that only two ancilla bits are required for the eigenvalue estimation part of the exponentiation algorithm, and such estimation results in the exact values. During the eigenvalue estimation algorithm, the number of calls to the controlled-QFT is three. The uncomputation requires another three calls to the controlled-QFT. This means that the query complexity of the computation of a fractional power of the QFT is exactly six queries to the controlled-QFT. Moreover, during such computation we introduced no errors due to the approximations of eigenvalues, and thus, given all gates are perfect, the computed fractional power is exact.

6.11 Conclusions

In this work the connection between continuous time black boxes (*i.e.*, "Hamiltonian oracles") and their discrete counterparts, unitary black boxes, has been further developed. Previous work has mostly focused on important special cases of this relationship. Our work pushes the boundary in a new direction, where the unknown $\mathcal{U}$ is a unitary acting on an arbitrarily large Hilbert space. It also demonstrates constructively that an algorithm that makes use of a continuous oracle can also be implemented using only a discrete oracle that that is equivalent to the continuous oracle for some fixed time, assuming that if the eigenvalues of the Hamiltonian implemented by the oracle are between 0 and E_{max}, then the fixed time interval should be bounded below $2\pi/E_{max}$ so that the gap assumption is satisfied.

For the case that t is a positive non-integer, and there are no assumptions on the eigenvalues of $\mathcal{U}$ other than the gap assumption, the complexity of this algorithm is substantially higher than that in the continuous setting, where one has access to a Hamiltonian H satisfying $\mathcal{U} = e^{-iH}$. If access to the unknown Hamiltonian, H, itself is given then in principle it is easy to perform any unitary of the form e^{-iHt} with a cost t in terms of calls to the Hamiltonian. For an algorithm with total query complexity $t < 1$ in the Hamiltonian black box model, it is hard to imagine a reasonable way to effect e^{-iHt} with fewer than one evaluation of $\mathcal{U}$, and indeed, the new algorithm makes $O(\frac{1}{\epsilon}\log\frac{1}{\epsilon})$ calls to $\mathcal{U}$ to effect e^{-iHt} with precision ϵ (a lower bound for this is shown in Section 6.4.2), which is notably lower than the number of calls required in the obvious approach of doing process tomography on $\mathcal{U}$, approximating $\mathcal{U}^t$, and then synthesizing a circuit to implement such an approximation. In particular, this approach does *not* depend on the dimension of the Hilbert space that $\mathcal{U}$ acts on.

An algorithm has been presented for finding the powers of unknown unitary operations raised to any real power $t > 0$. The complexity of the algorithm was found to be in $O(\lfloor t \rfloor + 2^m m)$ where the error is in $O(\frac{1}{2^m})$. For large integers t we presented a non-trivial example where the new method is exponentially more efficient than the direct repeated application of $\mathcal{U}$ a total of t times. Note that the same exponential speed-up is possible also in the continuous case if the Hamiltonian has some structure that enables eigenvalue estimation much better than the standard worst-case limits; in such cases, this procedure, using a modest amount of ancilla workspace, can simulate the evolution of H for a time τ in time polynomial in $\log(\tau)$.

This algorithm could be of practical use in a situation where there is very

precise clocks and control, with error less than a very small positive ϵ, but for some technical reason the unknown Hamiltonian must be run for a minimum time $T_{min} >> \epsilon$. Conceivably, this might occur if there is a minimum reset time on a of apparatus used to simulate or execute the Hamiltonian, and T_{min} is still smaller than $2\pi/E_{max}$ (assuming the energies are non-negative). In this case, there is a smallest amount of time for which a given Hamiltonian can be run in practice. If there is a need to run the Hamiltonian for less than that amount of time, then the method presented here could be used.

Further, if the Hamiltonian that solves some problem is time dependent, as is in general the case in quantum circuits or in models based on the permutation or braiding of particles on a surface, it may be of interest to apply some fraction of this total transformation to turn the solution to the problem into a subroutine of a larger algorithm. The method we propose provides a general way of obtaining this fractional application.

This algorithm could serve as a useful subroutine in other discrete time model quantum programs. This technique may also be applied to compute other functions of $\mathcal{U}$, by replacing each eigenvalue $e^{i\lambda_k}$ with $e^{if(\lambda_k)}$ for any function $f(\cdot)$. This may be useful in contexts such as Childs' method [Chi08] for simulating continuous time quantum walk Hamiltonians with discrete time walks. This requires applying phases that are the sines or arcsines of the eigenvalues of the original Hamiltonian. The method proposed here could achieve this for an unknown Hamiltonian to be simulated. The properties of the function $f(\cdot)$ will determine the efficiency as a function of the precision. Also, in the case when $f(\cdot)$ is continuous with bounded first derivative, and 2π-periodic, then we can also drop the assumption that there must be a gap in the spectrum, since this assumption was only needed to avoid the consequences of small errors in the eigenvalue estimation becoming large ones when the function of $\mathcal{U}$ which transforms the eigenvalues has a discontinuity.

In addition, this algorithm provides some intuition about a potential strategy for constructing a quantum public key cryptography scheme. There are a number of classical cryptography schemes that rely on the computational difficulty of finding roots modulo some number, such as RSA. One might attempt to construct quantum cryptography schemes where knowledge of the "secret key" corresponds to the ability to compute a fractional power (*i.e.*, a root) of some unitary operator. However, since this algorithm provides a way of implementing such roots, it seems that such an approach would be unlikely to yield a secure quantum cryptography scheme.

Harrow, Hassidim, and Lloyd [HHL08] investigated a similar technique to attack the problem of solving linear systems of equations in a more narrow setting and were able to obtain an exponential improvement over the best classical algorithm in that case. Specifically, they wanted to find and expectation value of the vector $\mathbf{x}$, $\mathbf{x}^\dagger M\mathbf{x}$ such that $\mathbf{x}$ is the solution to the equation $A\mathbf{x} = \mathbf{b}$ where $\mathbf{b}$ is a vector given in the problem and A is a sparse matrix of largest dimension N with a suitably small condition number κ whose description is also provided. Their optimal algorithm takes $poly(\log N)$ time, as opposed to $O(N)$, the time required classically, if the *condition number* κ, which is the ratio of the largest eigenvalue of A to the smallest, is $poly(\log N)$. The complexity of the algorithm goes with κ^2 (assumed to be $poly(\log N)$) and an intuition for why this should be is that as κ increases the matrix A becomes more difficult to invert.

This condition number constraint does relate somewhat to the need for a gap assumption in this work. The reason the gap is required is to ensure that errors do not occur too close to the vicinity of the discontinuity in the function $f(\theta) : \theta \mapsto \frac{\theta \mod 2\pi}{2}$ at zero. Likewise, if κ is very large then the renormalizing so that the largest eigenvalue does not exceed one makes the smallest eigenvalue very close to zero. However, the function of the eigenvalues Harrow *et al.* implement is $f(\lambda) = \frac{1}{\lambda}$, which is indeterminate at zero. Thus in both cases, there is an need to avoid small errors made in the eigenvalue estimation steps from becoming unduly magnified through constraining assumptions.

Note that one important difference between their work and ours is that they have a description of A and it is sparse and either Hermitian or adapted to be Hermitian in the algorithm so that it is easy to implement e^{iAt}. In the problem presented here there is no such description and this complicates the problem and in general prevents the exponential speed up seen in Harrow *et al.*'s work.

Chapter 7

Conclusions

Contents

7.1 Summary of Results

The results given here bring together ideas from the theory of quantum information processing, particularly reference frames, conservation laws, and symmetries, and algorithms.

For the use of phase references to maintain coherence within qubits in a quantum computation, as considered in Chapter 4, different requirements were laid out and methods suggested. The main important points to be realized are that systems used for quantum references need not be large - in fact for the proper choice of encoding, they can add a sublinear ancilliary system overhead, in terms of the space requirement of the quantum computation. Also, there are many ways of managing reference systems in a quantum computer, and some will be superior to others based on technological constraints, but some are definitely inferior in terms of reducing error in the computation. Error correction can be used against errors

caused by imperfect gate implementations due to imperfect references. Assuming the proper choice of strategy for managing the reference(s), these errors can be independent and stochastic, so standard error correction procedures will work. All gates will require energy to be performed, but these fields need not be kept coherent, and there is no reason in principle this energy could not be recycled from one gate to the next. Therefore, the energy for such gates is expected to be inversely proportional to the choice of error parameter to first order. These requirements, while presenting some technical challenges, do not present more serious problems to quantum computing than other known technical obstacles.

While references managed in the proper way with the use of error correction, can preserve information in a quantum setting, a quantum directional reference which is isolated except for being used for measurement or quantum operational tasks will degrade with use. A system such as this could be part of a device for measuring individual spin states in a solid state nuclear magnetic resonance quantum computer. The investigation in Chapter 3 sheds light on the behaviour of spin systems used repeatedly as directional references and gives bounds on their longevity as references before they need to be replenished. This is done though the means of studying the evolution of the moments of the angular momentum operator J_z. This gives expressions for fidelity for the reference state after many uses. A general expression is found for all rotationally invariant maps in terms of a limited number of parameters. These parameters are sufficient to study the action of the repeated map and draw some general conclusions about the longevity of the reference if the map and reference meet three conditions. The conditions are satisfied by many cases of physical interest, and if they are obeyed the number of times the reference can be used before it falls below and error threshold for a task is of order j^2 where the reference system is a spin-j system (that is, is described by a $2j+1$ dimensional density matrix). Various example cases are explicitly explored.

In the vein of quantum algorithms, in Chapter 5, a new form of the discrete time walk on a line was proposed and explored: a quantum walk with a pair of entangled walkers. This provided the opportunity to explore new dynamics in a simple quantum walk. The correlation functions, expected separations and probabilities for being found at a particular vertex and explored and compared to the case of walkers with a separable initial state. The use of antisymmetric initial conditions is found to give a linear increase in the expected separation of the walkers over a separable initial condition for two walkers.

The relation between continuous and discrete time quantum evolutions was considered in Chapter 6 and a new algorithm was developed for simulating fractional

evolution of a unitary based on eigenvalue estimation. The algorithm applies an approximation of the operation $\mathcal{U}^t$ for any real t where $\mathcal{U}$ is a black box. The algorithm is found to require $O\left(\frac{1}{\epsilon}\log\frac{1}{\epsilon}\right)$ calls to the unitary $\mathcal{U}$, provided a gap condition is fulfilled. An example where this algorithm could give an exponential speed up over the naive application of $\mathcal{U}$ to a system repeatedly was explored. The success of this hinged on a promise on the eigenvalues. The result is particularly interesting in the context of other recent results [CGM$^+$08, HHL08], as it seems to indicate that the complexity of inverting or taking roots of unitary matrices depends on the form of their eigenvalues, but not on their eigenvectors. Potential applications in cryptography and simulations were proposed.

Appendix A

Proof of Theorem 3.1

The proof of Theorem 3.1 is inspired by ideas from [KW99, D'A04] and is shown through three lemmas, which together demonstrate that all rotationally covariant maps can be expressed in the form $\xi(\rho_j) = \sum_{n=0}^{2j} q_n \zeta^{\circ n}(\rho_j)$. First it is shown that the map ζ is covariant with respect to the representations of the group $SU(2)$ on $2j+1$ dimensions, recalling that the group $SO(3)$ of space rotations and the group $SU(2)$ of unitary transformations are locally isomorphic up to a phase, so that the spatial covariance condition can be replaced by covariance with respect to $SU(2)$. This implies that ξ is a rotationally covariant map. The second lemma indicates the general form that any covariant map must take and shows that there are $2j+1$ independent parameters required to specify such a map. The third and final lemma demonstrates that the proposed expression for the map ξ is composed of $2j+1$ linearly-independent operators, each with a coefficient q_n. Lastly it is argued that these coefficients must be real and therefore, the $2j+1$ parameters required to express any covariant map can be expressed as the $2j+1$ independent real coefficients q_n. Together, these facts imply that any rotationally covariant map can be written $\xi(\rho_j) = \sum_{n=0}^{2j} q_n \zeta^{\circ n}(\rho_j)$ with the appropriate choice of real parameters q_n.

First, let ρ_j be a density operator on the Hilbert space of a spin-j system. Consider the map

$$\zeta(\rho_j) = \frac{1}{\lambda}(J_x \rho J_x + J_y \rho J_y + J_z \rho J_z), \qquad [3.9]$$

where $\lambda = j(j+1)$, and J_x, J_y and J_z are the angular momentum operators in the x, y and z directions for some arbitrary Cartesian frame. First, notice that the

map is unital, *i.e.*

$$\zeta(\mathbb{I}_j) = \frac{1}{\lambda}\sum_{i=1}^{3} J_i^\dagger J_i = \frac{1}{\lambda}J^2 = \mathbb{I}_j, \tag{A.1}$$

where J is the total angular momentum operator.

Lemma A.1. *The map ζ is covariant with respect to the spin-j irreducible representation R_j of $SU(2)$.*

Proof. If R_j is the $d = 2j+1$ dimensional irreducible representation (irrep) of the group $SU(2)$ generated by $\{J_x, J_y, J_z\}$, then it can be expressed in the form

$$R_j(\Omega) = e^{-i\theta_1 J_z} e^{-i\theta_2 J_y} e^{-i\theta_3 J_z}, \tag{A.2}$$

for some real parameters θ_1, θ_2 and θ_3 (which are strongly related to the Eulers angles). By the symmetry of ζ with respect to the y and z-axes, if it is proved that

$$R_z(\theta)^{-1}\zeta(R_z(\theta)\rho_j R_z(\theta)^{-1})R_z(\theta) = \zeta(\rho_j), \qquad \forall\ \theta \in [0, 2\pi), \tag{A.3}$$

for any rotations around the z-axis only, then (A.3) with any rotations around the y-axis follows immediately. From (A.2), deduce that (A.3) is satisfied for any rotations. Such z rotations will map

$$J_x \ \rightarrow\ \cos\theta J_x + \sin\theta J_y \tag{A.4}$$

$$J_y \ \rightarrow\ -\sin\theta J_x + \cos\theta J_y, \tag{A.5}$$

and inserting the transformed operators into equation 3.9 it can be readily computed that ζ is invariant with respect to z rotations. Therefore,

$$R_j(\Omega)^{-1}\zeta(R_j(\Omega)\rho_j R_j(\Omega)^{-1})R_j(\Omega) = \zeta(\rho_j) \qquad \forall\ \Omega \in SU(2). \tag{A.6}$$

The map ζ is therefore invariant. □

Because a composition of invariant maps is also invariant, then any map of the form

$$\xi(\rho) = q_0\rho + \sum_{k=1}^{2j} q_k \zeta^{\circ k}(\rho), \tag{A.7}$$

where the values q_i are real numbers and $\zeta^{\circ k} = \zeta \circ \zeta \circ \ldots \circ \zeta$, is also invariant with respect to $SU(2)$. It remains to show that any invariant map with respect to $SU(2)$ can be written on the form (A.7) .

To prove that every invariant map can be written as above, use is made of the Liouville representation of a superoperator [Hav03]. Upon representing a $d \times d$ density matrix ρ into a d^2 long column vector $|\rho\rangle\rangle$ by stacking the columns of the density matrix, the action of any superoperator ξ can be represented as a $d^2 \times d^2$ matrix $\mathcal{K}(\xi)$, such that

$$|\xi(\rho)\rangle\rangle = \mathcal{K}(\xi)|\rho\rangle\rangle. \tag{A.8}$$

This representation is necessarily basis dependent. If a given process has Kraus operators $\{E_k\}$, the Liouville representation takes the form

$$\mathcal{K}(\xi) = \sum_k E_k^* \otimes E_k \tag{A.9}$$

where $*$ represents the complex conjugate with respect to the chosen basis.

Lemma A.2. *The Liouville representation of any map that is invariant with respect to $SU(2)$ has the form*

$$\mathcal{K}(\xi) = (e^{-i\pi J_y} \otimes \mathbb{I}) \sum_{k=0}^{2j} c_k \Pi_k (e^{i\pi J_y} \otimes \mathbb{I}), \tag{A.10}$$

where $c_k \in \mathbb{C}$ and Π_k is the projector into the $2k+1$ dimensional subspace of total angular momentum k.

Proof. The condition on a map to be invariant can thus be expressed as

$$(R_j^*(\Omega) \otimes R_j(\Omega))\mathcal{K}(\xi) = \mathcal{K}(\xi)(R_j^*(\Omega) \otimes R_j(\Omega)), \qquad \forall\ \Omega \in SU(2), \tag{A.11}$$

so the condition on the operators in this new representation is that $\mathcal{K}(\xi)$ must commute with $R_j^* \otimes R_j$. If the group $SU(2)$ is represented in the $\{|j,m\rangle_z\}$ basis (i.e. the eigenstates of J_z), using equation A.2 and noting the way rotations about one axis update other operators as in equations A.4 and A.5, then due to the fact that $e^{-i\pi J_y} e^{-i\theta_1 J_z} e^{i\pi J_y} = e^{i\theta_1 J_z}$ (note the similarly to spin-echo techniques), $R_j^*(\Omega) = e^{-i\pi J_y} R(\Omega) e^{i\pi J_y}$. This implies

$$\begin{aligned}(R_j(\Omega) \otimes R_j(\Omega))&(e^{i\pi J_y} \otimes \mathbb{I})\mathcal{K}(\xi)(e^{-i\pi J_y} \otimes \mathbb{I}) \\ &= (e^{i\pi J_y} \otimes \mathbb{I})\mathcal{K}(\xi)(e^{-i\pi J_y} \otimes \mathbb{I})(R_j(\Omega) \otimes R_j(\Omega)) \ \ \forall\ \Omega \in SU(2).\end{aligned} \tag{A.12}$$

Note that $(R_j(\Omega) \otimes R_j(\Omega))$ is the *collective* representation of $SU(2)$ on two spin-j systems. A Clebsch-Gordan decomposition gives the irreducible representation of

the group of all collective rotations $\mathcal{G}(R_j \otimes R_j)$ which take the form

$$\mathcal{G}(R_j \otimes R_j)(\Omega) \simeq \bigoplus_{k=0}^{2j} R_k(\Omega), \qquad \forall\ \Omega \in SU(2)b, \tag{A.13}$$

where "$\simeq$" denotes "unitarily equivalent" and R_k is the spin-k irreducible representation of $SU(2)$ which has multiplicity one.

Because $\mathcal{K}(\xi)$ is required to commute with $R_j(\Omega) \otimes R_j(\Omega)$ for any $\Omega \in SU(2)$, it must then commute with all irreducible representations of $\mathcal{G}(R_j \otimes R_j)$. By Schur's lemma,

$$(e^{i\pi J_y} \otimes \mathbb{I})\mathcal{K}(\xi)(e^{-i\pi J_y} \otimes \mathbb{I}) \simeq \bigoplus_{k=0}^{2j} c_k \mathbb{I}_k = \sum_{k=0}^{2j} c_k \Pi_k, \tag{A.14}$$

where $c_k \in \mathbb{C}$, $\mathbb{I}_k$ is the $2k+1$ dimensional identity operator, and Π_k is the projector into the $2k+1$ dimensional subspace of total angular momentum k. □

Thus Lemma A.2 shows that there are $2j+1$ independent projectors forming $\mathcal{K}(\xi)$. To characterize every possible invariant mapping, $2j+1$ independent operators are required. Lemma A.3 demonstrates that this is true for the expression A.7.

Lemma A.3. *The matrices $(\sum_i J_i^* \otimes J_i)^k$ for $0 \leqslant k \leqslant 2j$ are linearly independent and form a complete basis to represent any invariant map of the form (A.10).*

Proof. Rewriting equation (A.7) in the Liouville representation, the map becomes

$$\mathcal{K}(\xi) = \sum_{k=0}^{n} \frac{q_k}{\lambda^k} \left(\sum_{i \in \{x,y,z\}} J_i^* \otimes J_i \right)^k. \tag{A.15}$$

The Hermitian matrices $(\sum_i J_i^* \otimes J_i)^k$ are diagonal in the same basis, and the eigenvalues of $(\sum_i J_i^* \otimes J_i)^k$ are ν_l^k, where ν_l is an eigenvalue of $\sum_i J_i^* \otimes J_i$.

To find all of the eigenvalues $\{\nu_l\}$, expand the total angular momentum operator for two spin-j systems $\mathcal{J}^2 = \sum_i (J_i \otimes \mathbb{I} + \mathbb{I} \otimes J_i)^2$ to get a new expression for $\sum_i J_i \otimes J_i$:

$$\begin{aligned} \sum_i J_i \otimes J_i &= \frac{1}{2}\Big(\sum_i (J_i \otimes \mathbb{I} + \mathbb{I} \otimes J_i)^2 - \sum_i J_i^2 \otimes \mathbb{I} - \mathbb{I} \otimes \sum_i J_i^2\Big) \\ &= \frac{1}{2}(\mathcal{J}^2 - J^2 \otimes \mathbb{I} - \mathbb{I} \otimes J^2) \\ &= \frac{1}{2}(\mathcal{J}^2 - 2j(j+1)\mathbb{I}) \end{aligned} \tag{A.16}$$

With the relation $\sum_i J_i^* \otimes J_i = -e^{-i\pi J_y} \otimes \mathbb{I} \left(\sum_i J_i \otimes J_i\right) e^{i\pi J_y} \otimes \mathbb{I}$, this implies that $\sum_i J_i^* \otimes J_i$ and $\sum_i J_i \otimes J_i$ will have the same eigenvalues up to a negative sign.

From equation (A.16), it can be shown that

$$\nu_l = -\frac{1}{2}\left(l(l+1) - 2j(j+1)\right), \tag{A.17}$$

where $l(l+1)$ is the eigenvalue resulting from the vector addition of the two spin-J systems, which implies that l can range from 0 to $2j$ and has multiplicity $2l+1$. There are $2j+1$ different values of l, so there is $2j+1$ different eigenvalues. By the fundamental theorem of algebra, this implies that there are no polynomials of degree $2j$ that have the eigenvalues $\{v_l\}$ as roots. This implies the matrices $(\sum_i J_i^* \otimes J_i)^k$ for $0 \leqslant k \leqslant 2j$ are linearly independent. By counting the number free parameters, it is proven that the operators $(\sum_i J_i^* \otimes J_i)^k$ form a complete basis to represent any map represented by equation (A.10). □

The last requirement necessary to show that the equation (A.7) is a representation of all possible invariant maps on a spin-j system, it is necessary to show that the coefficients q_i are real — that there are $2j+1$ parameters in this representation and not $4j+2$, which would be the case if there were a real and an imaginary parameter for each coefficient. First, note that

$$\zeta(\rho_j) = \frac{1}{2\lambda}(2J_z\rho_j J_z + J_+\rho_j J_+^\dagger + J_-\rho_j J_-^\dagger), \tag{A.18}$$

where $J_\pm = J_x \pm iJ_y$. Also, from equation (A.18), the only contribution to ${}_{\hat{n}}\langle j,-j|\,\xi(|j,j\rangle_{\hat{n}\hat{n}}\langle j,j|)\,|j,-j\rangle_{\hat{n}}$ is from $q_{2j}\zeta^{\circ 2j}(|j,j\rangle_{\hat{n}\hat{n}}\langle j,j|)$. So,

$${}_{\hat{n}}\langle j,-j|\,\xi(|j,j\rangle_{\hat{n}\hat{n}}\langle j,j|)\,|j,-j\rangle_{\hat{n}} = q_{2j}\ {}_{\hat{n}}\langle j,-j|\,\zeta^{\circ 2j}(|j,j\rangle_{\hat{n}\hat{n}}\langle j,j|)\,|j,-j\rangle_{\hat{n}}\,. \tag{A.19}$$

Since $\xi(\rho_j)$ and $\zeta^{\circ 2j}$ must be positive, then q_{2j} is a non-negative real number. The contributions to ${}_{\hat{n}}\langle j,-j+1|\,\xi(|j,j\rangle_{\hat{n}\hat{n}}\langle j,j|)\,|j,-j+1\rangle_{\hat{n}}$ are from the terms $q_{2j}\zeta^{\circ 2j}(|j,j\rangle_{\hat{n}\hat{n}}\langle j,j|)$ and $q_{2j-1}\zeta^{\circ 2j-1}(|j,j\rangle_{\hat{n}\hat{n}}\langle j,j|)$. And since $\xi(\rho_j)$ is positive and q_{2j} is real, then q_{2j-1} must also be real. Note that q_{2j-1} could be negative. The rest of the argument follows this same chain. By induction, it is clear that all the coefficients q_i must be real if ξ is to be a valid map on density matrices.

Appendix B

Proof of Lemma 6.1

Proof that $\Omega\left(\frac{1}{\delta}\right)$ calls are required to distinguish $\mathbb{I}$ from $U_{1-\delta}$ with high probability, where $\mathcal{U} \in \{U_{1-\delta} = |0\rangle\langle 0| + e^{-i\delta}|1\rangle\langle 1| \mid \delta \in [0, \pi)\}$:

The probability of distinguishing two admissible superoperators (that is, superoperators $\Phi(\rho) = \sum_1^n A_i \rho A_i^\dagger$ for some positive integer n where the Kraus operators obey $\sum_{i=1}^n A_i^\dagger A_i = \mathbb{I}_\rho$) is:

$$\frac{1}{2} + \frac{1}{4}\|\Phi_0 - \Phi_1\|_\diamond. \tag{B.1}$$

For unitary superoperators, $\Phi_0(\rho) = U\rho U^\dagger$ and $\Phi_1(\rho) = V\rho V^\dagger$:

$$\|\Phi_0 - \Phi_1\|_\diamond = \max_u\{\|\Phi_0(uu^\dagger) - \Phi_1(uu^\dagger)\|_{\mathrm{Tr}} : u \in \mathcal{S}(\mathcal{X})\} = \min_u 2\sqrt{1 - |u^\dagger U^\dagger V u|^2}$$

where $\mathcal{S}(\mathcal{X}) = \{u \in \mathcal{X} : |u| = 1\}$ and $\mathcal{X}$ is a Hilbert space.

So the probability of distinguishing $\mathbb{I}$ from $U_{1-\delta}$ with k calls is

$$\frac{1}{2} + \frac{1}{4}\min_{\mathbf{a}} 2\sqrt{1 - |\mathbf{a}^\dagger \mathbb{I}^\dagger F(U_{1-\delta})\mathbf{a}|^2} = \frac{1}{2} + \frac{1}{4}\min_{\mathbf{a}} 2\sqrt{1 - |\mathbf{a}^\dagger U_{1-\delta}\mathbf{a}|^2}$$

where $F(U)$ is some function of k applications of U on any number of qubits. So we need to discover the value of

$$\min_{\mathbf{a}} |\mathbf{a}^\dagger F(U_{1-\delta})\mathbf{a}|.$$

Consider first the case where $k = 1$. Then it suffices to consider one qubit as the input to either $\mathbb{I}$ or $U_{1-\delta}$ and $F(U) = U_{1-\delta}$. We need to find values a_1, a_2 such that

the following is minimized

$$\left|\begin{pmatrix} a_1^* & a_2^* \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & e^{-i\delta} \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}\right| = \left|(|a_1|^2 + e^{-i\delta}|a_2|^2)\right|$$

The right hand side is smallest when $|a_1|^2 = |a_2|^2 = \frac{1}{2}$, so that the phase is relative rather than global. In that case the probability of distinguishing these operators is equal to

$$\frac{1}{2} + \frac{1}{4} \times 2\sqrt{1 - \left|\frac{1}{2} + e^{-i\delta}\frac{1}{2}\right|^2} = \frac{1}{2} + \frac{1}{2}\sqrt{1 - \cos^2\left(\frac{\delta}{2}\right)} = \frac{1}{2}\left(1 + \sin\left(\frac{\delta}{2}\right)\right).$$

The solution oscillates between one half and one depending on the choice of $\frac{\delta}{2}$. If we now consider k applications of U, there are many ways of arranging the k applications on some number of qubits ranging from 1 to k. Additional qubits beyond k will not help since U contains a relative phase compared to $\mathbb{I}$ rather than just a global one.

Consider how repeated applications of Φ_U, the map which is a single application of the unitary, alter the diamond norm. We have just found that:

$$\|\Phi_U - \Phi_{\mathbb{I}}\| = 2\sin\left(\frac{\delta}{2}\right) \tag{B.2}$$

And so applying a second map Φ_U:

$$\|\Phi_U \circ \Phi_U - \Phi_{\mathbb{I}} \circ \Phi_{\mathbb{I}}\| = \|\Phi_U \circ \Phi_U - \Phi_U \circ \Phi_{\mathbb{I}} + \Phi_U \circ \Phi_{\mathbb{I}} - \Phi_{\mathbb{I}} \circ \Phi_{\mathbb{I}}\| \tag{B.3}$$

And by the triangle inequality

$$\begin{aligned} \|\Phi_U \circ \Phi_U - \Phi_{\mathbb{I}} \circ \Phi_{\mathbb{I}}\| &\leq \|\Phi_U \circ \Phi_U - \Phi_U \circ \Phi_{\mathbb{I}}\| + \|\Phi_U \circ \Phi_{\mathbb{I}} - \Phi_{\mathbb{I}} \circ \Phi_{\mathbb{I}}\| & \text{(B.4)} \\ &\leq 4\sin\left(\frac{\delta}{2}\right) & \text{(B.5)} \end{aligned}$$

Repeating this k times, the probability of distinguishing the operations is less than:

$$\frac{1}{2}\left(1 + 2k\sin\left(\frac{\delta}{2}\right)\right),$$

which is greater than some fixed constant when $k = \Omega\left(\frac{1}{\delta}\right)$. Therefore, in order to distinguish the operations $\mathbb{I}$ and $U_{1-\delta}$ with high probability, we require $\Omega(\frac{1}{\delta})$ calls to $U_{1-\delta}$.

References

[Aar07] S. Aaronson. The learnability of quantum states. *Proceedings of the Royal Society of London A*, **463**(2088):3089 – 3114, 2007. DOI:10.1098/rspa.2007.0113. EPRINT arXiv:quant-ph/0608142v3.

[ADZ93] Y. Aharonov, L. Davidovich, and N. Zagury. Quantum random walks. *Physical Review A*, **48**(2):1687 – 1690, 1993. DOI:10.1103/PhysRevA.48.1687.

[AE07] A. Ambainis and J. Emerson. Quantum t-designs: t-wise independence in the quantum world. In *Twenty-Second Annual IEEE Conference on Computational Complexity*, pp. 129 – 140, 2007. DOI:10.1109/CCC.2007.26.

[AGP06] P. Aliferis, D. Gottesman, and J. Preskill. Quantum accuracy threshold for concatenated distance-3 codes. *Quantum Information and Computation*, **6**(2):97 – 165, 2006. EPRINT arXiv:quant-ph/0504218v3.

[Alm94] L. B. Almeida. The fractional fourier transform and time-frequency representations. *IEEE Transactions Signal Processing*, **42**(11):3084 – 3091, 1994. DOI:10.1109/78.330368.

[Amb07] A. Ambainis. Quantum walk algorithm for element distinctness. *SIAM Journal on Computing*, **37**(1):210–239, 2007. DOI:10.1109/FOCS.2004.54. EPRINT arXiv:0311001 [quant-ph].

[AS67] Y. Aharonov and L. Susskind. Charge superselection rule. *Physical Review*, **155**(5):1428 – 1431, 1967. DOI:10.1103/PhysRev.155.1428.

[BB04] O. Buerschaper and K. Burnett. Stroboscopic quantum walks, 2004. EPRINT arXiv:quant-ph/0406039v2.

[BMK+99] D. Bouwmeester, I. Marzoli, G. P. Karman, W. Schleich, and J. P. Woerdman. Optical galton board. *Physical Review A*, **61**(1):013410, 1999. DOI:10.1103/PhysRevA.61.013410.

[Boa83] M. Boas. *Mathematical Methods in the Physical Sciences*. Wiley, 1983.

[BOS+02] M. D. Bowdrey, D. K. L. Oi, A. J. Short, K. Banaszek, and J. A. Jones. Fidelity of single qubit maps. *Physics Letters A*, **294**(5 – 6):258 – 260, 2002. DOI:10.1016/S0375-9601(02)00069-5.

[BRS04] S. Bartlett, T. Rudolph, and R. Spekkens. Optimal measurements for relative quantum information. *Physical Review A*, **70**:032321, 2004. DOI:10.1103/PhysRevA.70.032321.

[BRS07] S. Bartlett, T. Rudolph, and R. Spekkens. Reference frames, superselection rules, and quantum information. *Reviews of Modern Physics*, **79**:555 – 606, 2007. DOI:10.1103/RevModPhys.79.555.

[BRST06] S. Bartlett, T. Rudolph, R. Spekkens, and P. Turner. Degradation of a quantum reference frame. *New Journal of Physics*, **8**:58, 2006. DOI:10.1088/1367-2630/8/4/058.

[BSLB08] J.-C. Boileau, L. Sheridan, M. Laforest, and S. D. Bartlett. Quantum reference frames and the classification of rotationally invariant maps. *Journal of Mathematical Physics*, **49**(3):032105, 2008. DOI:10.1063/1.2884583. EPRINT arXiv:0709.0142 [quant-ph].

[BW99] J. P. Barnes and W. S. Warren. Decoherence and programmable quantum computation. *Physical Review A*, **60**(6):4363 – 4374, 1999.

[BYG06] Z. Bar-Yossef and M. Gurevich. Random sampling from a search engine's index. In *Proceedings of the 15th international conference on World Wide Web*, p. 367. Association for Computing Machinery, 2006. DOI:doi.acm.org/10.1145/1135777.1135833.

[CCD+03] A. M. Childs, R. Cleve, E. Deotto, E. Farhi, S. Gutmann, and D. A. Spielman. Exponential algorithmic speedup by quantum walk. In *Proceedings of the 35th ACM Symposium on the Theory of Computing*, pp. 59 – 68, 2003. DOI:10.1145/780542.780552.

[CEMM98] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca. Quantum algorithms revisited. *Proceedings of the Royal Society of London A*, **454**(1969):339–354, 1998. DOI:10.1098/rspa.1998.0164. EPRINT arXiv:quant-ph/9708016.

[CGM+08] R. Cleve, D. Gottesman, M. Mosca, R. Somma, and D. Yonge-Mallo. Efficient discrete-time simulations of continuous-time quantum query algorithms, 2008. EPRINT arXiv:0811.4428v1 [quant-ph].

[Chi08] A. M. Childs. On the relationship between continuous- and discrete-time quantum walk, 2008. EPRINT arXiv:0810.0312v1 [quant-ph].

[CI00] N. J. Cerf and S. Iblisdir. Optimal n-to-m cloning of conjugate quantum variables. *Physical Review A*, **62**:040301, 2000. DOI:10.1103/PhysRevA.62.040301.

[CL08] B. A. Chase and A. J. Landahl. Universal quantum walks and adiabatic algorithms by 1d hamiltonians, 2008. EPRINT arXiv:0802.1207v1 [quant-ph].

[CN97] I. Chuang and M. Nielsen. Prescription for experimental determination of the dynamics of a quantum black box. *Journal of Modern Optics*, **44**(11 – 12):2455 – 2467, 1997. DOI:10.1080/09500349708231894.

[CS51] E. U. Condon and G. Shortley. *The Theory of Atomic Spectra*. Cambridge University Press, 1951.

[CS05] Y. C. Cheng and R. J. Silbey. A study on the noise threshold of fault-tolerant quantum error correction. *Physical Review A*, **72**:012320, 2005. DOI:10.1103/PhysRevA.72.012320. EPRINT arXiv:quant-ph/0412168v1.

[CW05] J. Combes and H. M. Wiseman. States for phase estimation in quantum interferometry. *Journal of Optics B: Quantum and Semiclassical Optics*, **7**:14 – 21, 2005. DOI:10.1088/1464-4266/7/1/004.

[D'A04] G. M. D'Ariano. Extremal covariant quantum operations and POVM's. *Journal of Mathematical Physics*, **45**:3620 – 3635, 2004. DOI:10.1063/1.1777813. EPRINT arXiv:quant-ph/0310024v4.

[DFK91] M. Dyer, A. Frieze, and R. Kannan. A random polynomial-time algorithm for approximating the volume of convex bodies. *Journal of the ACM*, **38**:1–17, 1991.

[Dir58] P. A. M. Dirac. *The Principle of Quantum Mechanics*. Clarendon Press Oxford, 1958.

[DRKB03] W. Dür, R. Raussendorf, V. M. Kendon, and H.-J. Briegel. Quantum walks in optical lattices. *Physical Review A*, **66**:052319, 2003. DOI:10.1103/PhysRevA.66.052319.

[EAZ05] J. Emerson, R. Alicki, and K. Zyczkowski. Scalable noise estimation with random unitary operators. *Journal of Optics B: Quantum and Semiclassical Optics*, **7**(10):S347 – S352, 2005. DOI:10.1088/1464-4266/7/10/021.

[FG98] E. Farhi and S. Gutmann. Analog analogue of a digital quantum computation. *Physical Review A*, **57**:2403, 1998. DOI:10.1103/PhysRevA.57.2403.

[FGG00] E. Farhi, J. Goldstone, and S. Gutmann. A numerical study of the performance of a quantum adiabatic evolution algorithm for satisfiability, 2000. EPRINT arXiv:quant-ph/0007071v1.

[FMRC62] A. Fletcher, J. C. P. Miller, L. Rosenhead, and L. J. Comrie. *An Index of Mathematical Tables*, volume 1–2. Blackwell, Oxford and Addison-Wesley, Reading, MA, 2nd edition, 1962.

[FT82] E. Fredkin and T. Toffoli. Conservative logic. *International Journal of Theoretical Physics*, **21**(3–4):219–253, 1982. DOI:10.1007/BF01857727.

[GB02] J. Gea-Banacloche. Minimum energy requirements for quantum computation. *Physical Review Letters*, **89**(21):217901, 2002. DOI:10.1103/PhysRevLett.89.217901.

[GBO06] J. Gea-Banacloche and M. Ozawa. Minimum-energy pulses for quantum logic cannot be shared. *Physical Review A*, **74**(6):060301, 2006. DOI:10.1103/PhysRevA.74.060301. EPRINT arXiv:0611137v1 [quant-ph].

[GG84] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**(6):721 – 741, 1984. DOI:10.1109/TPAMI.1984.4767596.

[Gis96] N. Gisin. Hidden quantum nonlocality revealed by local filters. *Physics Letters A*, **210**(3):151 – 156, 1996. DOI:10.1016/S0375-9601(96)80001-6.

[Gro46] H. J. Groenewold. On the principles of elementary quantum mechanics. *Physica*, **12**(7):405 – 460, 1946. DOI:10.1016/S0031-8914(46)80059-4.

[Gro96] L. K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of 28th Annual ACM Symposium on the Theory of Computing*, pp. 212–219, 1996. DOI:10.1145/276698.276712.

[Hal06] B. C. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Springer, 2006.

[Hav03] T. F. Havel. Procedures for converting among Lindblad, Kraus and matrix representations of quantum dynamical semigroups. *Journal of Mathematical Physics*, **44**(2):534 – 557, 2003. DOI:10.1063/1.1518555. EPRINT arXiv:quant-ph/0201127.

[HBF03] M. Hillery, J. Bergou, and E. Feldman. Quantum walks based on an interferometric analogy. *Physical Review A*, **68**:032314, 2003. DOI:10.1103/PhysRevA.68.032314.

[Hec00] K. E. Hecht. *Quantum Mechanics*. Springer, 2000.

[HHH99] M. Horodecki, P. Horodecki, and R. Horodecki. General teleportation channel, singlet fraction, and quasidistillation. *Physical Review A*, **60**(3):1888 – 1898, 1999. DOI:10.1103/PhysRevA.60.1888.

[HHL08] A. Harrow, A. Hassidim, and S. Lloyd. Quantum algorithm for solving linear systems of equations, 2008. EPRINT arXiv:0811.3171v2 [quant-ph].

[Ioa02] L. Ioannou. Continuous-time quantum algorithms: Searching and adiabatic computation. Master's thesis, University of Waterloo, 2002. URL `http://uwspace.uwaterloo.ca/handle/10012/1129`.

[JPK04] H. Jeong, M. Paternostro, and M. S. Kim. Simulation of quantum random walks using the interference of a classical field. *Physical Review A*, **69**:012310, 2004. DOI:10.1103/PhysRevA.69.012310.

[KBLW01] J. Kempe, D. Bacon, D. A. Lidar, and K. B. Whaley. Theory of decoherence-free fault-tolerant universal quantum computation. *Physical Review A*, **63**:042307, 2001. DOI:10.1103/PhysRevA.63.042307.

[Kem03] J. Kempe. Quantum random walks - an introductory overview. *Contemporary Physics*, **44**(4):307 – 327, 2003. DOI:10.1080/0010715103100011077 6. EPRINT arXiv:quant-ph/0303081v1.

[Kem04] A. Kempf. Amath673 lecture notes, 2004.

[KGBO09] T. Karasawa, J. Gea-Banacloche, and M. Ozawa. Gate fidelity of arbitrary single-qubit gates constrained by conservation laws. *Journal of Physics A: Mathematical and General*, 2009. EPRINT arXiv:0809.3095v2 [quant-ph].

[Kit95] A. Kitaev. Quantum measurements and the Abelian stabilizer problem, 1995. eprint arXiv:quant-ph/9511026v1.

[KLM07] P. Kaye, R. Laflamme, and M. Mosca. *An Introduction to Quantum Computing*. Oxford University Press, 2007.

[KLZ98] E. Knill, R. Laflamme, and W. Zurek. Resilient quantum computation: Error models and thresholds. *Proceedings of the Royal Society of London A*, **454**(1969):365 – 384, 1998. doi:10.1098/rspa.1998.0166. eprint arXiv:quant-ph/9702058v1.

[KMO08] G. Kimura, B Meister, and M. Ozawa. Quantum limits of measurements induced by multiplicative conservation laws: Extension of the wigner-araki-yanase theorem. *Physical Review A*, **78**:032106, 2008. doi:10.1103/PhysRevA.78.032106.

[KR03] A. Klappenecker and M. Rötteler. Quantum software reusability. *International Journal on Foundations of Computer Science*, **14**(5):777–796, 2003. doi:10.1142/S0129054103002023. eprint arXiv:0309121 [quant-ph].

[KW99] M. Keyl and R. F. Werner. Optimal cloning of pure states, testing single clones. *Journal of Mathematical Physics*, **40**:3283 – 3299, 1999. doi:10.1063/1.532887. eprint arXiv:quant-ph/9807010.

[LCC+02] Q. Lv, P. Cao, E. Cohen, K. Li, and S. Shenker. Search and replication in unstructured peer-to-peer networks. In *Proceedings of the 16th ACM International Conference on Supercomputing*, p. 84, 2002. doi:10.1145/514191.514206.

[LCW98] D. A. Lidar, I. L. Chuang, and K. B. Whaley. Decoherence-free subspaces for quantum computation. *Physical Review Letters*, **81**:2594, 1998. doi:10.1103/PhysRevLett.81.2594.

[LOYT98] R. C. Liu, B. Odom, Y. Yamamoto, and S. Tarucha. Quantum interference in electron collision. *Nature*, **391**:263, 1998. doi:10.1038/34611.

[Moc07] C. Mochon. Hamiltonian oracles. *Physical Review A*, **75**:042313, 2007. doi:10.1103/PhysRevA.75.042313.

[MR02] C. Moore and A. Russell. Quantum walks on the hypercube. In *Randomization and Approximation Techniques in Computer Science*, *Lecture Notes in Computer Science*, volume 2483, p. 952. Springer, 2002. doi:10.1007/3-540-45726-7_14. eprint arXiv:quant-ph/0104137v1.

[Nai43] M. A. Naimark. On a representation of additive operator set functions. *C. R. (Dolkady) Academy of Sciences URSS*, **41**:359 – 361, 1943.

[NC97] M. A. Nielsen and I. L. Chuang. Programmable quantum gate arrays. *Physical Review Letters*, **79**:321 – 324, 1997. doi:10.1103/PhysRevLett.79.321.

[NC00] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[Nie02] M. A. Nielsen. A simple formula for the average gate fidelity of a quantum dynamical operation. *Physics Letters A*, **303**:249, 2002. DOI:10.1016/S0375-9601(02)01272-0.

[NW08] D. Nagaj and P. Wocjan. Hamiltonian quantum cellular automata in 1d. *Physical Review A*, **78**:032311, 2008. DOI:10.1103/PhysRevA.78.032311.

[OPBV02] Y. Omar, N. Paunković, S. Bose, and V. Vedral. Spin-space entanglement transfer and quantum statistics. *Physical Review A*, **65**:062305, 2002. DOI:10.1103/PhysRevA.65.062305.

[OPSB06] Y. Omar, N. Paunković, L. Sheridan, and S. Bose. Quantum walk on a line with two entangled particles. *Physical Review A*, **74**:042304, 2006. DOI:10.1103/PhysRevA.74.042304.

[Oza02] M. Ozawa. Conservative quantum computing. *Physical Review Letters*, **89**:057902, 2002. DOI:10.1103/PhysRevLett.89.057902.

[PA07] P. K. Pathak and G. S. Agarwal. Quantum random walk of two photons in separable and entangled states. *Physical Review A*, **75**:032351, 2007. DOI:10.1103/PhysRevA.75.032351.

[POBV02] N. Paunković, Y. Omar, S. Bose, and V. Vedral. Entanglement concentration using quantum statistics. *Physical Review Letters*, **88**:187903, 2002. DOI:10.1103/PhysRevLett.88.187903.

[Pre98] J. Preskill. Reliable quantum computers. *Proceedings of the Royal Society of London A*, **454**(1969):385 – 410, 1998. DOI:10.1098/rspa.1998.0167. EPRINT arXiv:quant-ph/9705031v3.

[PY07] D. Poulin and J. Yard. Dynamics of a quantum reference frame. *New Journal of Physics*, **9**:156, 2007. DOI:10.1088/1367-2630/9/5/156.

[RBMC04] D. Rugar, R. Budakian, H. J. Mamin, and B. W. Chui. Single spin detection by magnetic resonance force microscopy. *Nature*, **430**(6997):329 – 332, 2004. DOI:10.1038/nature02658.

[RC03] J. Roland and N. J. Cerf. Quantum-circuit model of Hamiltonian search algorithms. *Physical Review A*, **68**:062311, 2003. DOI:10.1103/PhysRevA.68.062311. EPRINT arXiv:quant-ph/0302138v1.

[RC04] C. P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer Texts in Statistics. Springer, 2004.

[Rei06] B. W. Reichardt. *Automata, Languages and Programming*, *Lecture Notes in Computer Science*, volume 4051, chapter Fault-tolerance threshold for a distance-three quantum code, pp. 50 – 61. Springer, 2006. DOI:10.1007/11786986_6. EPRINT arXiv:quant-ph/0509203v2.

[Rit05] W. G. Ritter. Quantum channels and representation theory. *Journal of Mathematical Physics*, **46**(8):082103, 2005. DOI:10.1063/1.1945768.

[Ros57] M. E. Rose. *Elementary Theory of Angular Momentum*. John Wiley & Sons, Inc., 1957.

[SBTK03] B. C. Sanders, Stephen D. Bartlett, Ben Tregenna, and Peter L. Knight. Quantum quincunx in cavity quantum electrodynamics. *Physical Review A*, **67**:042305, 2003. DOI:10.1103/PhysRevA.67.042305.

[Sch05] I. Schur. *Neue Begründung der Theorie der Gruppencharaktere*. Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften zu Berlin, 1905.

[SGB+95] J. A. Sidles, J. L. Garbini, K. J. Bruland, D. Rugar, O. Züger, S. Hoen, and C. S. Yannoni. Magnetic resonance force microscopy. *Reviews of Modern Physics*, **67**(1):249 – 265, 1995. DOI:10.1103/RevModPhys.67.249.

[Sho97] P. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, **26**(5):1484 – 1509, 1997. DOI:10.1137/S0097539795293172.

[SKW03] N. Shenvi, J. Kempe, and K. B. Whaley. Quantum random-walk search algorithm. *Physical Review A*, **67**:052307, 2003. DOI:10.1103/PhysRevA.67.052307.

[SLB87] M. Scully, W. Lamb, and A. Barut. On the theory of the Stern-Gerlach apparatus. *Foundations of Physics*, **17**:575 – 583, 1987. DOI:10.1007/BF01882788.

[SPOB06] L. Sheridan, N. Paunković, Y. Omar, and S. Bose. Discrete time quantum walk on a line with two particles. *International Journal of Quantum Information*, **4**(3):573 – 583, 2006. DOI:10.1142/S0219749906002006.

[Ste96] A. M. Steane. Error correcting codes in quantum theory. *Physical Review Letters*, **77**(5):793 – 797, 1996. DOI:10.1103/PhysRevLett.77.793.

[Sze04] M. Szegedy. Quantum speed-up of Markov chain based algorithms. In *Proceedings of the 45th IEEE Symposium on the Foundations of Computer Science*, pp. 32 – 41. IEEE, 2004. DOI:10.1109/FOCS.2004.53.

[TM02] B. C. Travaglione and G. J. Milburn. Implementing the quantum random walk. *Physical Review A*, **65**:032310, 2002. DOI:10.1103/PhysRevA.65.032310.

[vDDE+07] W. van Dam, G. M. D'Ariano, A. Ekert, C. Macchiavello, and M. Mosca. Optimal quantum circuits for general phase estimation. *Physical Review Letters*, **98**:090501, 2007. DOI:10.1103/PhysRevLett.98.090501.

[vEK01] S. J. van Enk and H. J. Kimble. On the classical character of control fields in quantum information processing. *Quantum Information and Computation*, **2**(1):1 – 13, 2001. EPRINT arXiv:quant-ph/0107088v1.

[Wat01] J. Watrous. Quantum simulations of classical random walks and undirected graph connectivity. *Journal of Computer and System Sciences*, **62**(2):376 – 391, 2001. DOI:10.1006/jcss.2000.1732.

[WWW52] G. C. Wick, A. S. Wightman, and E. P. Wigner. The intrinsic parity of elementary particles. *Physical Review*, **88**(1):101 – 105, 1952. DOI:10.1103/PhysRev.88.101.

[Zal98] C. Zalka. Simulating quantum systems on a quantum computer. *Proceedings of the Royal Society of London A*, **454**:313 – 322, 1998. DOI:10.1098/rspa.1998.0162.