University of Strathclyde

Department of Physics

PhD Thesis

Some aspects of measurement of quantum systems

Author:
Andrew Colin

Supervisors:
Prof. Stephen Barnett
Dr. John Jeffers

December 2, 2012

Abstract

This thesis is about two different aspects of measurement.

The first aspect concerns programmed discrimination, where a single binary digit is coded in the relationship between three qbits. There are two different 'program' qbits, and a 'data' qbit which is *guaranteed* to be identical to one or other of the program qbits. The task of programmed discrimination is to discover, as accurately as possible, which of the two possibilities is true.

Most prior work assumes that the two program qbits are randomly placed on the surface of the Bloch sphere. Our own research examines error and failure rates when there is some classical information about the two program qbits. We considered three configurations:

- There is a known overlap between the program qbits

- Both program qbits are on a *known* great circle

- The program qbits are confined to polar caps of a known size.

Our results cover both error rates in optimum recognition and failure rates in unambiguous discrimination, and consider the effects of using multiple copies of the qbits.

The second aspect studies the effects of frequent, but possibly incorrect, measurement on otherwise closed quantum systems with two eigenstates. Previous work has set up a model for this situation, and showed the existence of 'telegraphing', in which the state of the system changes very little over extended periods of time before switching to the opposite state. This is related to the Zeno [1] effect, whereby a system that is measured repeatedly at frequent intervals does not change state. Our work offers a mathematical analysis of the situation, and shows that the analysis conforms well to practical results obtained from a simulator.

The overall arrangement of the thesis is:

- Chapter 1: A general introduction to relevant areas of quantum physics

- Chapter 2: A review of previous work in programmed discrimination

- Chapters 3 and 4: Reports on our own research in programmed discrimination

- Chapter 5: A summary of a recent paper on telegraphing

- Chapter 6: Our work on telegraphing

- Chapter 7: Conclusion.

Our original work is presented in Chapters 3, 4 and 6.

---

[1] Zeno (490-430 BC) held that motion is nothing but an illusion, and suggested various paradoxes, such as that of the tortoise and the hare, to prove his point.

Acknowledgements

Whenever people asked me, "Why are you doing a Ph.D?" I always replied, "For fun". I've had a lot of fun over the last few years, and have begun to explore mysteries of nature which have always filled me with wonder and some dismay. In spite of many popular books I have read, I could not understand Quantum Physics. I still don't understand it, but now I know that no-one else can, either. Richard Feynman said, "Anyone who thinks he understands quantum mechanics, doesn't understand quantum mechanics". So that's OK then.

I have been immensely helped during the study period by many organisations and people. I have to thank the University of Strathclyde for a scholarship that covered my fees, and for allowing me an extra year to get over a time of ill-health. I am grateful to Dr. Alison Yao, with whom I share an office. She has always been friendly, supportive and helpful in many ways, from sorting out my computer connection problems to lending me books. My appreciation is due to fellow students and post-docs in the CNQO group, who provided encouragement and practical help when I was stuck with problems. People of my age are often invisible to younger people, but in the Physics Department I was simply treated as one of their own. I mention Electra, Fiona, Filippo, Nikhil, Robert,and Thomas, all members of 'our' group. My counsellor and long-standing friend Dr. Peter Maas has given me valuable advice and encouragement throughout the study period. My thanks are also due to Professor Pearson, of the Department of Physics and Astronomy, Dickinson College, Carlisle, PA, for permission to use the image in Chapter 1.

My two supervisors, Dr. John Jeffers and Prof. Steve Barnett, were always ready to see me, to read my writings, and enabled me to go to conferences at Imperial College, Strathclyde and Pecs. I met many interesting people and absorbed a number of valuable ideas. I value their collaboration in two of the papers that help to make up this thesis, and I acknowledge the direct contribution that Steve made in suggesting the method of finding density matrices for triads with known overlap that I used in Chapter 3, and the complete analysis of the "Measurement-Centred" method in Chapter 6. I am particularly grateful to Steve for suggesting that I should study for a doctorate at all. We met over a glass of wine when I was completing an Open University Maths degree with a Quantum Physics module. He helped me get a good mark on the current assignment. I thought he was jesting when he offered to admit me to his department as a graduate student, but he replied, "I never joke about such matters!". This thesis is the outcome.

But - most of all - I thank Veronica, my wife of over 50 years. It was her who ceded me the best room in the house as my study, brought me innumerable cups of tea and coffee, put up with my regular getting up at 5.00 am, and encouraged me when I was sick and could no longer see clearly to the end of the work. Without her I would never even have started, let alone finished this thesis. No words can say what I feel.

Thank you all.

Andrew Colin

Contents

Chapter 1

Background to the research

Alice laughed: "There's no use trying," she said; "one can't believe impossible things." "I daresay you haven't had much practice," said the Queen. "When I was younger, I always did it for half an hour a day. Why, sometimes I've believed as many as six impossible things before breakfast." *Lewis Carroll, "Alice in Wonderland"*

1.1 Introduction

This chapter introduces techniques and concepts which are relevant to the subject of the thesis. We have been selective:

Some topics, such as the derivation of the Lindblad equation from first principles, are described in detail.

Certain areas, such as the response of atoms to radiation, are given as background and receive a simple, non-technical description.

Other subjects, important though they are to quantum physics, are simply omitted because they play no part in our research. For example, we say nothing about the Bell inequalities.

1.2 The tools of the trade

Research in theoretical quantum information uses a number of mathematical tools, algorithms and special notations. In this thesis we use:

- Complex numbers. Many of the quantities that occur in quantum physics are best modelled by complex numbers. The complex complement of a number z is usually written as z^*. A useful relation is
$$zz^* = z^*z = |z|^2.$$
.

- Hilbert [1] spaces. A Hilbert space is a home for vectors with complex elements. The space can have any number of dimensions, in which each coordinate is a complex number.

- Vector algebra. Vectors are often written with the *bra-ket* notation, introduced in 1939 by Paul Dirac [1]. Bras and kets are vectors in a Hilbert space. In a *ket* the elements are enclosed by the signs $|$ and $\rangle$, thus : $|a, b, c\rangle$. In a *bra* the elements are enclosed by the signs $\langle$ and $|$, but here each element is the complex conjugate of the corresponding ket. A ket is essentially a column vector, and a bra, a row vector. We may put

$$\langle a, b, c| = (|a^*, b^*, c^*\rangle)^T. \tag{1.1}$$

  It is common to represent the list of numbers in a bra or ket by a single symbol, as in $|\lambda\rangle$. Bras and kets are subject to the normal rules of vector algebra. Thus, as long as λ and κ both occupy the same Hilbert space, the form $\langle\lambda|\kappa\rangle$ is an *inner product*, and evaluates to a complex number known as the *overlap*. Likewise $|\lambda\rangle\langle\kappa|$ is an *outer product* and represents a square matrix.

  In most cases, vectors can be multiplied by arbitrary factors, and it is common to scale the elements so that the magnitude of each vector is 1. Such a vector is said to be *normalised*.

  It is worth noting that:

    - If $|\lambda\rangle$ is normalised, $\langle\lambda|\lambda\rangle = 1$

    - If $|\lambda\rangle$ and $|\kappa\rangle$ are orthogonal to one another, $\langle\lambda|\kappa\rangle = 0$.

- Matrix algebra. By convention every matrix is written with a hat to distinguish it from a scalar with the same name. ($\hat{A} \neq A$). Many matrices in quantum physics, including the matrices of the form $|\lambda\rangle\langle\lambda|$, are *hermitian*.

    - Hermitian matrices are square

    - Their diagonal elements are real, and the off-diagonal elements, which may be complex, follow the rule $a[j, k] = a^*[k, j]$. A hermitian matrix is equal to the transpose of its complex complement. Symbolically,

$$\hat{A} = \hat{A}^\dagger$$

    - Every square matrix $\hat{A}$ of order n has n eigenvalues λ_j (not necessarily distinct) and n corresponding eigenvectors $|\lambda_j\rangle$. These are defined by the equation

$$\hat{A}|\lambda_j\rangle = \lambda_j|\lambda_j\rangle \tag{1.2}$$

    - All the eigenvalues of a hermitian matrix are real

    - All its eigenvectors are orthogonal to one another

---

[1] Names of mathematicians and scientists that are quoted exactly use initial capitals. When the name is modified for use as an adjective, the first letter is in lower case, except when the word starts a new sentence. Thus: *Hilbert* or *Dirac* but *hermitian* or *hamiltonian*

- A hermitian matrix is deemed *positive* if none of its eigenvalues is less than zero

- Equation 1.2 remains true when the components of an eigenvector are multiplied by any constant. In practice, eigenvectors are normalised.

When a matrix $\hat{B}$ is not hermitian, it can always be split into two hermitian components $\hat{X}$ and $\hat{Y}$:

$$\hat{B} = \hat{X} + i\hat{Y}. \tag{1.3}$$

Another useful concept is the *unitary matrix*. This matrix has the property that its inverse is equal to its complex conjugate. Thus $\hat{U}^{-1} = \hat{U}^{\dagger}$ and $\hat{U}\hat{U}^{\dagger} = \hat{1}$ (where $\hat{1}$ is the unit matrix.)

Multiplication of a vector by a unit matrix is equivalent to a rotation about the origin of the relevant Hilbert space.

- Tensor algebra. The use of tensor algebra in this thesis is minimal. The only relevant operator is *tensor multiplication*. This example demonstrates the operation, as applied to two vectors:

$$\begin{pmatrix} a \\ b \end{pmatrix} \otimes \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix}. \tag{1.4}$$

Tensor multiplication of two matrices is equally straightforward:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes \begin{pmatrix} p & q \\ r & s \end{pmatrix} = \begin{pmatrix} ap & aq & bp & pq \\ ar & as & br & bs \\ cp & cq & dp & dq \\ cr & cs & dr & ds \end{pmatrix}. \tag{1.5}$$

- Laplace transforms. The main use of Laplace transforms is in solving linear differential equations. The method works by transforming many functional relations, *including those with derivatives*, into purely algebraic expressions, which can be manipulated by the normal rules of algebra.

- The standard algorithms we used included:

  - *Gaussian Elimination*, to solve large sets of linear simultaneous equations

  - *the Jacobi method* of finding the eigenvalues and eigenvectors of hermitian matrices

  - *Runge Kutta 4th-order integration* to find numerical solutions of differential equations

  - *Golden Section search* to find extreme values of arbitrary functions

All these algorithms were taken from [2]. In addition, Mathematica $^{\text{TM}}$ proved invaluable in finding both symbolic and numerical integrals of complex functions, and Microsoft Excel was useful in plotting graphs.

1.3 Classical and quantum physics

The laws of classical physics were derived from experiments and observations conducted on a 'human' scale. Physicists could devise mental models of nature, based on practical everyday experience, and use them to make valid predictions.

The laws of classical physics proved to hold over a vast range of magnitudes, and were seen to apply, without significant modification, to systems as small as a virus or as large as the solar system. On the other hand, the laws of extremely small systems turned out to be largely different from their classical counterparts, and not intuitively obvious to anyone familiar with classical physics. For example, the only acceptable explanation for certain well-attested observations is that photons are *both* particles and waves, and that they can be in two or more places at the same time. In this context mental models no longer work.

These small systems are termed *quantum systems*, and generally consist of a few elementary particles (typically one or two). Typical quantum systems are photons, electrons, or single atoms.

In reality there is no boundary of size, such that smaller systems follow quantum rules, but larger ones behave classically. Everything follows the laws of quantum physics, but in larger systems all we can observe is the statistical behaviour of immense numbers of quantum systems acting at the same time. So the laws of classical physics are essentially *statistical* in nature. For example, a classical physicist could take a sample of radioactive material, measure the radiation it emits, and show that it is constant over short periods and decays exponentially over time. The quantum physicist would point out that the apparently steady stream of radiation is due to large numbers of atoms individually decaying at completely random times.

The link between quantum phenomena and classical physics is given by Ehrenfest's Law [3], which states that every quantum rule for the *expectation* for some variable (such as momentum) has an exact parallel in classical physics that specifies an actual value for that variable.

1.4 Quantum systems and states

Initially we shall study the way quantum systems behave in isolation; that is, not subject to any external influence. Such systems are said to be *closed*.

Each such system can be described by a number of *observables*: physical properties which depend on the type of system. At any moment every particle has a position and a velocity[2] but otherwise the observables depend on the type of particle. For example

- A photon has a certain polarisation, and a frequency which is directly proportional to its energy

- an electron has a charge, a mass, and a 'spin'[3]

---

[2]Heisenberg's Uncertainty Principle implies that it is impossible, even in principle, to know both the position and velocity of a particle exactly.

[3]The spin of a particle has no direct equivalent in classical physics. It can only take two possible values - up or down. Spin generates a momentum, but there is no related angular velocity.

- an atom has a mass, and several alternative energy states depending on the orbits of its outermost electrons.

We divide these properties into three groups:

Some properties, such as the speed of a photon in free space, the charge on an electron and the mass of a given species of atom at rest with regard to the observer, are all constants of nature. These are of no further interest to us. Other properties, such as position and momentum, are continuously variable. In this thesis they will not be of central interest either. The remaining properties, *when measured*, are discrete: for example, an electron can be either 'spin-up' or 'spin-down', a photon can be either horizontally or vertically polarised, or a two-state atom can be either in a high-energy or low-energy state.

It is worth noting that none of these observables is absolute. The polarisation of a photon depends on the orientation of the measuring device; the spin of an electron is measured in relation to a selected axis, and the energy of an atom is based on an arbitrary 'zero-energy'.

A useful aspect of quantum physics is that all its mathematical laws, except those that refer to measurement, are *linear*. This implies that we can use standard vector and matrix algebra to express these laws, which simplifies what would otherwise be an intractable situation.

Dirac's notation is useful for recording and describing the states of quantum systems. Systems are generally named after Greek letters ψ, ϕ, ... and possible states of these systems by arbitrary symbols, such as $\uparrow$ for a particle in the spin-up state, or $\downarrow$ for the spin-down state. The energy states of an atom can be called '0' and '1'. Possible descriptions of states are $|\psi\rangle = |\uparrow\rangle$ or $|\phi\rangle = |1\rangle$.

1.4.1 Pure states

An important subdivision of states is into *pure* and *mixed*.

A fundamental feature of quantum physics is that a system can be in two or more states at the same time. This is called *superposition*. If $|\psi\rangle = |\uparrow\rangle$ and If $|\psi\rangle = |\downarrow\rangle$ are two allowable states of a system, then so is

$$|\psi\rangle = a|\uparrow\rangle + b|\downarrow\rangle \tag{1.6}$$

where a and b are *complex* numbers called *probability amplitudes*.

The actual values of a and b in any instance are generally unknown and unknowable. However, if the system is measured in a suitable way, the measurement leaves it in state $|\uparrow\rangle$ with probability aa^* or $|a|^2$, or in state $|\downarrow\rangle$ with probability $|b|^2$. As one or other of these outcomes is certain, then for a pure system:

$$|a|^2 + |b|^2 = 1. \tag{1.7}$$

If the possible states of a system are known and assigned symbols in a fixed order, the state symbols can be omitted, and the description of the state becomes a normalised ket of probability amplitudes. Thus we can write

$$|\psi\rangle = |a, b\rangle \tag{1.8}$$

with the $|\uparrow\rangle$ and $|\downarrow\rangle$ being implied.

Such a ket is equivalent to a unit vector in a Hilbert space with as many dimensions as there are possible states.

The *outer product* of a bra and a ket usually involves the same state in both vectors. The result is called a *density matrix*:

$$\hat{\rho} = |\psi\rangle\langle\psi|. \tag{1.9}$$

As we have seen, $\hat{\rho}$ is a hermitian matrix. Each row, and each column correspond to a possible state. Every element on the diagonal is a real positive number and indicates the probability that a measurement yields the corresponding state. The trace of the matrix $\hat{\rho}$ is 1, and furthermore it can be shown that the trace of $\hat{\rho}^2$ is also 1.

The density matrix of a system encapsulates everything that we can actually know about the system in its current state.

1.4.2 Mixed states

Suppose we have means of creating both spin-up particles and spin-down particles (both pure states), and mix them in the ratio x spin-ups to y spin-downs. The corresponding density matrix will be the weighted sum of the density matrices for each type of particle:

$$x|\uparrow\rangle\langle\uparrow| + y|\downarrow\rangle\langle\downarrow|. \tag{1.10}$$

The trace of the density matrix for a mixed system is also 1, but the trace of its square is not 1; this is the standard way to distinguish pure from mixed systems.

1.4.3 The Bloch sphere

The Bloch sphere [4] offers a useful way of representing quantum systems that live in a Hilbert space of two dimensions. The sphere (shown in figure 1.1) has unit radius. By analogy with the Earth, we can identify any point on its surface with two angles: ϕ which gives the longitude starting from an arbitrary 'Greenwich' meridian, and θ, which specifies the latitude.[4]

Take a system in a superposition of two possible states:

$$|\psi\rangle = a|0\rangle + b|1\rangle.$$

Every system of this form can be mapped on to a point on the surface of the Bloch sphere with the equation:

$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle \tag{1.11}$$

because:

1. We can extract a common phase factor κ from $|\psi\rangle$, so that:

$$|\psi\rangle = e^{i\kappa}(c|0\rangle + d|1\rangle)) \tag{1.12}$$

---

[4]Actually the geographical latitude is $\pi/2 - \theta$, because the global North Pole is at 90°, but the north pole of the Bloch sphere has $\theta = 0$.

where c is *real*, but d is still complex. The common phase factor has no physical significance and can be ignored

2. We can replace d by $e^{i\phi}g$, where g is *real*

3. As $c^2 + g^2 = 1$, we can replace c by $\cos(\theta/2)$ and g by $\sin(\theta/2)$. This substitution gives equation 1.11 directly.

Equation 1.11 forms a useful alternative to the standard form. Figure 1.1 marks the points that correspond to the states $|0\rangle$ and $|1\rangle$, respectively.
Mixed states are represented by points *inside* the Bloch sphere.
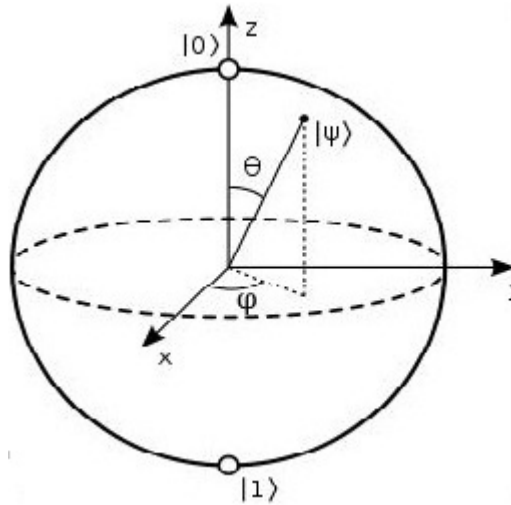
Figure 1.1: The Bloch Sphere.

1.4.4 Entanglement

Consider two systems $|\alpha\rangle = |a,b\rangle$ and $|\beta\rangle = |c,d\rangle$. The overall state of the two systems taken together and viewed as a single entity is the tensor product of α and β.

$$|\alpha\rangle \otimes |\beta\rangle = |ac, ad, bc, bd\rangle. \tag{1.13}$$

However *any* normalised ket $|p, q, r, s\rangle$ can be seen to represent two possible states. If this ket cannot be factorised into the form $|a, b\rangle \otimes |c, d\rangle$ the two states are said to be *entangled*, and any measurement on one will instantaneously affect the other, even if they are physically separate. The simplest examples of entanglement are provided by the four Bell states [5]:

$$\begin{aligned}
|\Psi^-\rangle &= \frac{1}{\sqrt{2}}(|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle) \\
|\Psi^+\rangle &= \frac{1}{\sqrt{2}}(|0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle) \\
|\Phi^-\rangle &= \frac{1}{\sqrt{2}}(|0\rangle \otimes |0\rangle - |1\rangle \otimes |1\rangle) \\
|\Phi^+\rangle &= \frac{1}{\sqrt{2}}(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle).
\end{aligned}$$

$$(1.14)$$

In no case can these states be factorised into a separate states for each system.

Schrődinger considered entanglement to be *the* essence of quantum physics.

1.5 Operators

In simple terms, an operator is a rule that can be applied to a state to produce a (generally) different state. The identity operator $\hat{\mathbf{1}}$ leaves a state unchanged. Operators have various uses in quantum physics. Two of the most important are

- Defining the results of measurements

- Computing the evolution of a state through time.

When we investigate a continuous attribute (such as position or momentum) the relevant operator can be expressed algebraically; but when only a limited number of values are possible, the operator can be written as a square matrix $\hat{A}$.

1.5.1 Operators and observables

Every observable in quantum physics is related to a hermitian operator, whose eigenvalues give the possible values of the observable. For example, suppose we represent spin-up (in a given direction) as +1, and spin-down as -1. Then the 'spin' operator is the matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

This matrix has eigenvalues +1 and -1, and the corresponding eigenvectors are

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} (= |\uparrow\rangle) \text{ and } \begin{bmatrix} 0 \\ 1 \end{bmatrix} (= |\downarrow\rangle). \tag{1.15}$$

The converse is not true; we shall encounter many hermitian matrices that do not correspond to any observable. From this point, when we refer to an 'observable' we imply its operator.

1.5.2 Operators and system evolution

Quantum systems do not generally remain static, but constantly evolve in time. This evolution is governed by the *hamiltonian* operator. Hamiltonians are derived from energy considerations, and the eigenvalues of the operator give the various possible energies of the system. In the case of continuous observables the hamiltonian would be a formula involving kinetic energy and potential energy due to particle interaction, but in the case of discrete values of an observable the hamiltonian is always a hermitian matrix. The derivation of a hamiltonian depends on a detailed knowledge of the system, but we shall give two examples:

First, consider a charged particle in a uniform magnetic field. The particle can be spinning in either direction. Its energy E in the spin-up state is

$$E_{up} = \frac{\gamma_s B \hbar}{2} \tag{1.16}$$

where γ_s is the *spin gyromagnetic ratio*, B is the strength of the magnetic field, and $\hbar$ is Planck's constant [6]. The field can be adjusted, and the other two quantities are constants found by experiment. Similarly, the energy in the spin-down state is

$$E_{down} = -\frac{\gamma_s B \hbar}{2}. \tag{1.17}$$

Both energies are relative to an arbitrary base. The corresponding hamiltonian is:

$$\hat{H} = \begin{pmatrix} \frac{\gamma_s B \hbar}{2} & 0 \\ 0 & -\frac{\gamma_s B \hbar}{2} \end{pmatrix} \tag{1.18}$$

which has the correct eigenvalues.

Another example is taken from a different context. Consider a single atom, which has two possible energy states, called $|0\rangle$ and $|1\rangle$, that correspond to different orbits of its outermost electron. The outer orbit has the higher energy, and when the electron drops to the inner orbit it emits a photon of a frequency that corresponds to the difference in energy. Likewise when an atom in the low orbit state encounters a photon of just the right energy, it absorbs the photon and jumps to the high energy orbit [7]. The hamiltonian governing this situation is:

$$\hat{H} = \begin{pmatrix} 0 & \frac{i\hbar\Omega}{2} \\ -\frac{i\hbar\Omega}{2} & 0 \end{pmatrix} \tag{1.19}$$

where Ω is the frequency of the incoming radiation. The way that a closed system evolves in time is governed by Schrődinger's equation [3]:

$$i\hbar \frac{d|\psi\rangle}{dt} = \hat{H}|\psi\rangle. \tag{1.20}$$

The formal solution of this differential equation, which specifies the state of the system at time t (starting from a known state at time $t = 0$) is

$$|\psi(t)\rangle = \hat{U}(t)|\psi(0)\rangle \tag{1.21}$$

where

$$\hat{U}(t) = \exp\left(-i\frac{\hat{H}t}{\hbar}\right). \tag{1.22}$$

In simple cases $\hat{U}(t)$ can be found analytically, but in most practical cases we must use numerical methods running on a computer.

The properties of $\hat{U}$ mean that evolution of a state in a closed system can be seen as a rotation of the state vector about the origin of the Hilbert space. The norm of the state is preserved. Schrődinger's equation is entirely deterministic. Knowing the starting state $|\psi(0)\rangle$ and the hamiltonian $\hat{H}$ we can calculate the state at any other time, moving forward or backward. Changes to the state are reversible.

When we solve the equations that relate to our single atom, the atom is found to execute oscillations with a frequency Ω. These are known as **Rabi** oscillations [7]. The state of the atom at time t can be described mathematically as

$$|\psi(t)\rangle = \cos\left(\frac{\Omega t}{2}\right)|0\rangle + i\sin\left(\frac{\Omega t}{2}\right)|1\rangle. \tag{1.23}$$

Most of the time, therefore, the atom is in a *superposition* of the two states $|0\rangle$ and $|1\rangle$.

Schrődinger's equation also has a form that applies to density matrices rather than states:

$$\frac{d\hat{\rho}}{dt} = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}] \tag{1.24}$$

where the notation $[\hat{A}, \hat{B}]$ is called a *commutator* and is short for $(\hat{A}\hat{B} - \hat{B}\hat{A})$.

1.6 Measurements

There are many ways of measuring the states of quantum systems. For example, a photon can leave a mark on a photo-sensitive film or trigger a charge-coupled device. It is said that under the right conditions a human can actually *see* a single photon. The polarisation of a photon can be detected by a polaroid filter. Again, as in the Stern-Gerlach experiment [8, 9], spinning particles can be passed through non-uniform magnetic fields to see which way they are deflected.

All quantum measurements can be described by a theoretical framework which we describe below.

All good measurements[5] have certain common features:

- They are probabilistic, so that when a measurement is made of a system which has two (or more) superimposed states, the outcome will (typically) be one of those states. The actual outcome cannot be predicted; only its probability, which depends, as we have explained,

---

[5]The effect of "bad" measurements - that is, measurements which could return incorrect results - is more complex. This area is discussed in Chapters 5 and 6.

on the probability amplitudes in the state of the system. To give an example, if the state of our atom at a given moment is

$$|\psi\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{i}{\sqrt{2}}|1\rangle \qquad (1.25)$$

a measurement might return $|0\rangle$ or $|1\rangle$: both are equally probable and we cannot tell in advance which it will be.

- The measurement inevitably changes the state of the system. Thus the photon that strikes the film or retina is absorbed, and the atom, after the measurement, will typically take on the state which the measurement reported - either $|0\rangle$ or $|1\rangle$. The change is instantaneous, and is known as the 'collapse of the wave function'.

  There is an unanswered philosophical question as to the *true* state of a system before it is measured - or even whether it has a true state at all [10] .

  Figure 1.2 shows the progress of a system through time. The smooth segments are the determinate solutions of Schrödinger's equation, and the sudden jumps, whose outcome is unpredictable, correspond to measurements.
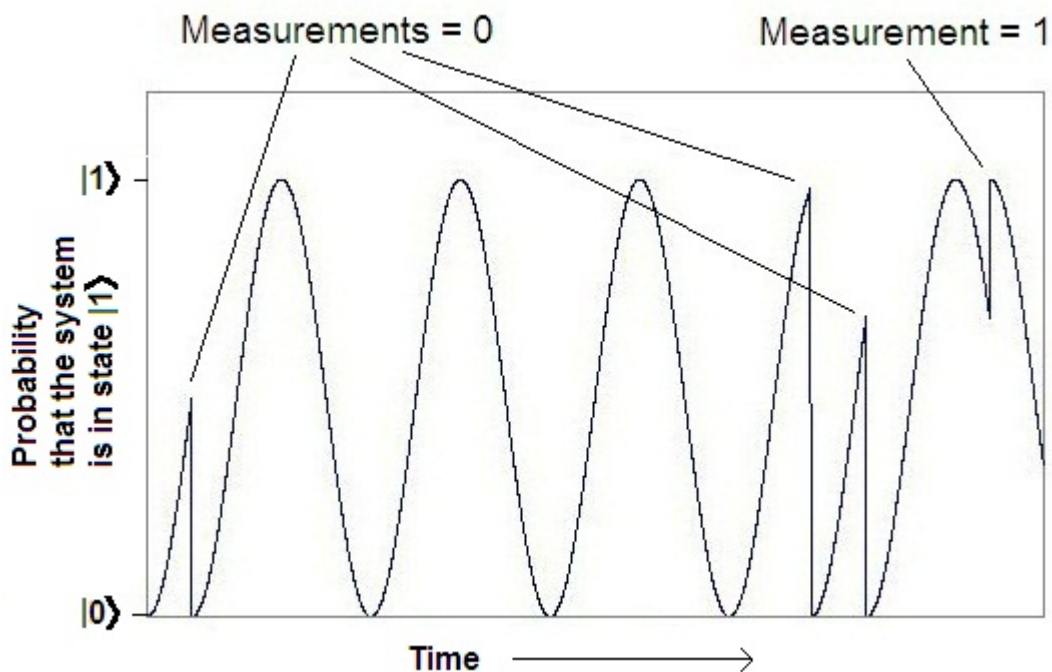

Figure 1.2: State of a quantum system with measurements. The jumps are not predictable.

1.6.1   Von Neumann measurements

The simplest kind of measurement is named after John Von Neumann [11]. As we said above, every non-continuous observable of a quantum system can be represented by a matrix $\hat{A}$, which

has real eigenvalues λ_n and corresponding orthogonal eigenvectors $|\lambda_n\rangle$. We may write,

$$\hat{A} = \sum_n \lambda_n |\lambda_n\rangle\langle\lambda_n|. \tag{1.26}$$

Each eigenvector $|\lambda_j\rangle$ represents a possible state of the system.

Returning to the definition of overlap $\langle\psi|\phi\rangle$, we mentioned that if $|\psi\rangle$ is the same as $|\phi\rangle$ the overlap is 1, but if they are orthogonal the overlap is 0.

Suppose that

$$|\psi\rangle = \sum_n a_j |\lambda_j\rangle \tag{1.27}$$

where the a_j's are the probability amplitudes of the possible states. Then the overlap $\langle\lambda_m|\psi\rangle$ will return a_m, which is precisely the probability amplitude that a measurement on $|\psi\rangle$ will yield $|\lambda_m$. Likewise the overlap $\langle\psi|\lambda_m\rangle$ returns a_m^*, and their product, $\langle\lambda_m|\psi\rangle\langle\psi|\lambda_m\rangle$ gives p_m, the actual probability of this result. But $|\psi\rangle\langle\psi|$ is $\hat{\rho}$, the density matrix of $|\psi\rangle$, so we may put

$$p_m = \langle\lambda_m|\hat{\rho}|\lambda_m\rangle. \tag{1.28}$$

As $\hat{\rho}$ is hermitian, it can be shown that

$$p_m = \text{Trace } (\hat{\rho}|\lambda_m\rangle\langle\lambda_m|). \tag{1.29}$$

The outer product $|\lambda_m\rangle\langle\lambda_m|$ is commonly called a *projector* $\hat{P}_m$.

A set of projectors derived from an 'observable' matrix have several important properties:

- Each projector is hermitian

- Each projector is positive - that is, all its eigenvalues are real and positive or zero

- Each projector corresponds to a possible outcome of the measurement

- The set of projectors is *complete*: that is, $\sum_n \hat{P}_n = \hat{1}$

- The square of a projector (or any positive integral power) is equal to itself

- The projectors in the set are orthonormal; $\hat{P}_j\hat{P}_k = 0$ unless $j = k$.

In the ideal case, we can see the measurement process as presenting the unknown system with such a set of projectors. Just one of them 'clicks' (to use the jargon) and gives the corresponding measurement. The system then adopts the state that corresponds to that measurement. This process is **indeterminate**. All we know in advance are the probabilities of the different outcomes.

A useful notation gives the *expectation* of an observable:

$$\langle A \rangle = \langle\psi|\hat{A}|\psi\rangle. \tag{1.30}$$

This means that if the measurement is applied to a large number of identical copies of $|\psi\rangle$ the *average* result tends towards the real number $\langle A \rangle$. It does **not** mean that the same system is measured many times; for after the first measurement the system will take one or other of the eigenvectors of the matrix, and further measurements will not give any useful information.

1.6.2 Generalised measurements

We have seen how a Von Neumann measurement can be constructed from the projectors of an observable. Such a method is insufficient in many practical cases. For example, once a photon has been "measured" by striking a retina it ceases to exist as an entity, and its "state after the measurement" is meaningless. A more flexible system is known as 'Generalised Measurement' [13]. Like a projective measurement, it consists of a set of operators $\{\hat{\pi}_n\}$ such that

- Each operator corresponds to a possible outcome of the measurement

- Each $\hat{\pi}_j$ is hermitian

- Each $\hat{\pi}_j$ is positive

- The set of operators is *complete*: $\sum_n \hat{\pi}_n = \hat{1}$

It is no longer necessary for operators to be orthogonal. Their number is not bound to the order of the observable, but can be greater or less. The operators that comprise a general measurement are called a *Positive Operator Measure*, often shortened to **POM**.[6]

A simple corollary of Naimark's Theorem [15, 16] shows that any set of operators that obey the rules for a POM are equivalent to a projective measurement in a space with added dimensions. This is equivalent to selecting a hypothetical "ancillary" system with a selected state and finding its tensor product with the state to be measured. The correct selection allows the extended operators to be orthogonal. When a measurement is made with a POM, the probability of outcome p_j is Trace$(\hat{\rho}\hat{\pi}_j)$.

1.6.3 Quantum discrimination

A common problem in quantum physics is to discriminate between two or more possible states of a system. A full discussion of this topic can be found in "Quantum State Discrimination", a review article by S.M. Barnett and S. Croke [12]. If the permitted states are orthogonal and known in advance, a simple projective measurement will suffice; but if they are not, the problem is to find a generalised measurement that is optimised in some respect. Various types of optimisation are possible, but we consider only two:

- *Minimum error discrimination* aims to find a POM such that the average error rate is as small as possible

- *Unambiguous discrimination* seeks to find a POM such that discrimination is sometimes accurate, but at the cost of frequent failure to return any result at all.

To give a very simple illustration of minimum error discrimination, consider a two-dimensional system that can take either of the two states $|\psi_1\rangle$ or $|\psi_2\rangle$, as shown in figure 1.3. We take it that each state is equally likely to occur with a probability of 0.5. These two states are not orthogonal, and cannot be accurately discriminated by a projective measurement.

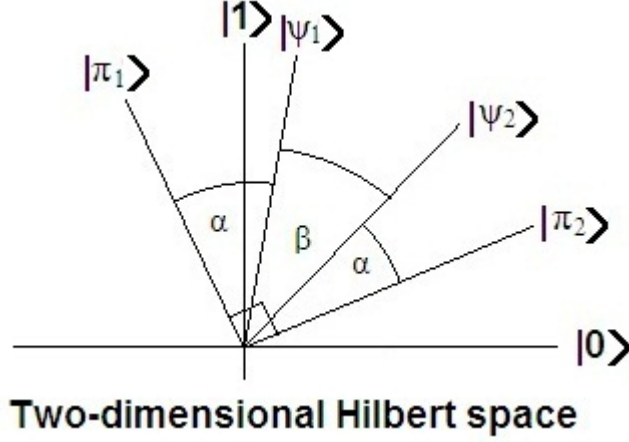[6]Some authors prefer *Positive Operator-valued Measure* or **POVM**

Figure 1.3: Two states to be discriminated optimally.

Consider two *orthogonal* vectors $|\pi_1\rangle$ and $|\pi_2\rangle$ and the corresponding operators $\hat{\pi}_1$ and $\hat{\pi}_2$, where $\hat{\pi}_1 = |\pi_1\rangle\langle\pi_1|$ and $\hat{\pi}_2 = |\pi_2\rangle\langle\pi_2|$. The probability of $|\psi_1\rangle$ causing $\hat{\pi}_1$ to 'click' is its projection on to $|\pi_1\rangle$, giving a correct identification. But because $|\psi_1\rangle$ and $|\pi_2\rangle$ are not orthogonal, there is also the chance that $|\psi_1\rangle$ clicks $|\pi_2\rangle$, leading to a false result. In terms of operators, the overall error rate E is

$$E = \frac{\text{Tr }(\hat{\rho}_1\hat{\pi}_2 + \hat{\rho}_2\hat{\pi}_1)}{2} \tag{1.31}$$

where $\hat{\rho}_1$ and $\hat{\rho}_2$ are density matrices: $\hat{\rho}_1 = |\psi_1\rangle\langle\psi_1|$ and $\hat{\rho}_2 = |\psi_2\rangle\langle\psi_2|$.

Clearly, the error rate is minimised in the symmetrical situation shown in figure1.3. Here $2\alpha + \beta = \pi/2$. Most practical situations are more complicated. There may be *a priori* probabilities for the states, there may be any number of states, and number of dimensions in the underlying Hilbert space can exceed 2. The problem of finding a minimum eror rate has been investigated by C.W. Helstrom in [17], who found a particularly elegant result for two states and any number of dimensions. The following derivation is due to [18].

Let η_1 and η_2 be the *a priori* probabilities of the two states, $\hat{\rho}_1$ and $\hat{\rho}_2$ their associated density matrices, and $\hat{\pi}_1$ and $\hat{\pi}_2$ the components of a POM used to discriminate the two states. We note that $\eta_1 + \eta_2 = 1$ and $\hat{\pi}_1 + \hat{\pi}_2 = \hat{\mathbf{1}}$. Given that the actual state of the system has density matrix $\hat{\rho}_1$, the probability of it being identified correctly is $\text{Tr}(\hat{\rho}_1\hat{\pi}_1)$, and the probability of *incorrect* identification is $\text{Tr}(\hat{\rho}_1\hat{\pi}_2)$. Similar considerations apply to the other state. As every measurement must have *some* outcome, then clearly

$$\sum_{j=1}^{2}\sum_{k=1}^{2}\text{Tr }(\hat{\rho}_j\hat{\pi}_k) = 1. \tag{1.32}$$

The overall probability of correct identification P_{right} is given by

$$P_{right} = \eta_1\text{Tr}(\hat{\rho}_1\hat{\pi}_1) + \eta_2\text{Tr}(\hat{\rho}_2\hat{\pi}_2). \tag{1.33}$$

Using equation 1.32, we can write the overall probability of failure, P_{wrong}, as

$$P_{wrong} = (1 - P_{right}) = \eta_1 \text{Tr}(\hat{\rho}_1 \hat{\pi}_2) + \eta_2 \text{Tr}(\hat{\rho}_2 \hat{\pi}_1). \quad (1.34)$$

We now introduce the hermitian operator

$$\hat{\Lambda} = \eta_1 \hat{\rho}_1 - \eta_2 \hat{\rho}_2. \quad (1.35)$$

Simple algebra (replacing $\hat{\Lambda}$ by its definition) will show that

$$P_{wrong} = \eta_1 + \text{Tr}(\hat{\Lambda}\hat{\pi}_1) = \eta_2 - \text{Tr}(\hat{\Lambda}\hat{\pi}_2). \quad (1.36)$$

The spectral decomposition of $\hat{\Lambda}$ is:

$$\hat{\Lambda} = \sum_{k=1}^{D} \lambda_k |\phi_k\rangle\langle\phi_k| \quad (1.37)$$

where D is the number of dimensions in $\hat{\Lambda}$, $\hat{\lambda}_k$ is an eigenvalue of $\hat{\Lambda}$, and $|\phi_k\rangle$ the corresponding eigenvector. All the eigenvalues are real, but (in general) some may be zero or negative. This lets us rewrite the equation for the probability of error as

$$P_{wrong} = \eta_1 + \sum_{k=1}^{D} \lambda_k \text{ Tr } (|\phi_k\rangle\langle\phi_k|) = \eta_2 - \sum_{k=1}^{D} \lambda_k \text{ Tr } (|\phi_k\rangle\langle\phi_k|). \quad (1.38)$$

The problem of finding the minimum error rate now reduces to finding the POM elements $\hat{\pi}_1$ and $\hat{\pi}_2$ so that the right hand sides of equation(1.38) are minimised, under the constraint that

$$0 \leq \langle\phi_k|\hat{\pi}_j|\phi_k\rangle \text{ for } (j = 1, 2 \text{ and } 1 \leq k \leq D)$$

as each term of the form $\langle\phi_k|\hat{\pi}_j|\phi_k\rangle$ is a non-negative probability. This can be achieved by ensuring that $\langle\phi_k|\hat{\pi}_1|\phi_k\rangle = 1$ and $\langle\phi_k|\hat{\pi}_2|\phi_k\rangle = 0$ are true for eigenstates with negative eigenvalues. Likewise $\langle\phi_k|\hat{\pi}_1|\phi_k\rangle = 0$ and $\langle\phi_k|\hat{\pi}_2|\phi_k\rangle = 1$ should be true for eigenstates with non-negative eigenvalues. The best forms for $\hat{\pi}_1$ and $\hat{\pi}_2$ follow directly; $\hat{\pi}_1$ is the sum of the projectors associated with negative eigenvalues, and $\hat{\pi}_2$ is the sum of the projectors with zero or positive eigenvalues.

It is convenient to reorder the eigenvalues and their eigenvectors so that λ_1 to λ_q are negative, and λ_{q+1} to λ_D are zero or positive. Then

$$\hat{\pi}_1 = \sum_{k=1}^{q} |\phi_k\rangle\langle\phi_k| \text{ and } \hat{\pi}_2 = \sum_{k=q+1}^{D} |\phi_k\rangle\langle\phi_k|. \quad (1.39)$$

Inserting these operators in equation (1.34) , the minimum error is

$$P_{wrong} = \eta_1 - \sum_{k=1}^{q} |\lambda_k| = \eta_2 - \sum_{k=q+1}^{D} |\lambda_k|. \quad (1.40)$$

Adding these equations, and noting that $\eta_1 + \eta_2 = 1$,

$$2P_{wrong} = 1 - \sum_{k=1}^{D} |\lambda_k|. \quad (1.41)$$

This leads directly to Helstrom's rule:

Construct $\hat{\Lambda} = \eta_1 \hat{\rho}_1 - \eta_2 \hat{\rho}_2$ and calculate z, the sum of the absolute values of the eigenvalues of $\hat{\Lambda}$. Then the optimum error rate E, the smallest error rate that can be achieved, is given by

$$E = \frac{1-z}{2} \tag{1.42}$$

This relationship is used extensively in our own work, discussed in Chapters 3 and 4.

The possibility of unambiguous discrimination was pointed out by I.D. Ivanovic, D. Dieks and A. Peres [19, 20, 21] in 1987 and further elaborated by A. Chefles and S.M. Barnett in 1998 [22]. The argument can be stated in terms of a system with two possible states, $|\psi_1\rangle$ and $|\psi_2\rangle$, with corresponding density matrices $\hat{\rho}_1$ and $\hat{\rho}_2$, and a POM with two component operators $\hat{\pi}_1$ and $\hat{\pi}_2$, derived from the vectors $|\pi_1\rangle$ and $|\pi_2\rangle$. Suppose that $|\pi_1\rangle$ is made orthogonal to $|\psi_2\rangle$, and $|\pi_2\rangle$ is made orthogonal to $|\psi_1\rangle$, as shown in figure 1.4. Evidently, as $\text{Trace}(\hat{\rho}_1\hat{\pi}_2) = 0$, state $|\psi_1\rangle$ can *never* make $\hat{\pi}_2$ click. If $\hat{\pi}_2$ does click, the state of the system *must* be $|\psi_2\rangle$. Likewise, if $\hat{\pi}_1$ clicks, the state must be $|\psi_1\rangle$.
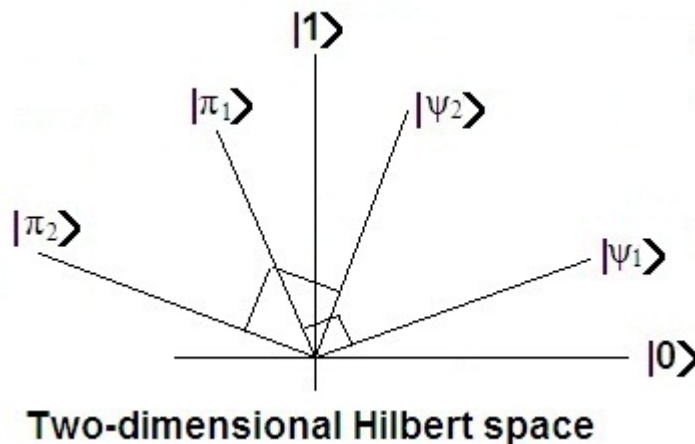
Figure 1.4: Two states to be discriminated unambiguously.

Unfortunately there remains a snag. As $|\psi_1\rangle$ and $|\psi_2\rangle$ are not orthogonal to one another, nor are the vectors $|\pi_1\rangle$ and $|\pi_2\rangle$. The resulting operators $\hat{\pi}_1$ and $\hat{\pi}_2$ do not sum to the unit matrix, and so violate one of the essential conditions of a POM. It is necessary to introduce a *third* operator, $\hat{\pi}_0$, such that

$$\hat{\pi}_0 = \hat{1} - \hat{\pi}_1 - \hat{\pi}_2 \tag{1.43}$$

$\hat{\pi}_0$ must be a *positive* operator ; that is, all its eigenvalues must be equal to or greater than zero. To ensure this condition, it is necessary to scale down the other two operators, which decreases the likelihood that they will click with a given state. This third operator corresponds to the result 'don't know', and its presence is the penalty that must be paid for intermittent, unambiguous

22

discrimination. Furthermore, the measurement process, like all quantum measurements, destroys the system being measured, so one cannot go back and measure it again. In practice, the 'don't know' result is far more frequent than a successful discrimination. Unambiguous discrimination also features in Chapters 3 and 4.

1.6.4   Practical measurements

To this point, we have described the measurement of quantum systems from a purely theoretical point of view. We have presented a Von Neumann measurement as a set of orthogonal projectors, and a general measurement as a group of π operators. Making the corresponding practical measurements is hard[7] because the experimenter must deal with single quantum systems which are many orders of magnitude smaller than those normally encountered in laboratories.

Most work reported in the literature [12, 23, 24] uses single photons in various states of polarisation as the quantum systems to be measured. A fully equipped quantum optics laboratory might contain

1. Light sources, which must generate precisely directed beams of single photons of a well-defined wavelength, and polarisation. This can be done by lasers with strongly attenuating filters, or quantum dots. Sometimes non-linear crystals are used to generate pairs of entangled photons moving in different directions

2. Photon detectors, which can react to single photons and emit a recordable electric pulse

3. Coincidence sensors, which can detect simultaneous pulses from two photon detectors

4. Wave-plates, which rotate the polarisation of a beam. They are described as half- or quarter- wave plates depending on the angle of rotation

5. Mirrors, to change the direction of the light

6. Beam splitters, which are basically semi-silvered mirrors, reflecting half the incident light and allowing the other half to pass without change. A polarising beam splitter can separate the horizontally and vertically polarised components of the incident light. The Glan-Thomson polariser deflects both output beams symmetrically.

Figure 1.5 is a typical quantum optics laboratory. The various components are fastened to a heavy steel table which keeps them in alignment. They include mirrors, beam-splitters,and a photo-diode, and are apparently arranged as a Mach-Zehnder interferometer [26].

Figure 1.6 shows the schematic symbols used for the various components, and figure 1.7 represents an experiment due to R.B.Clarke, A. Chefles, S.M. Barnett and E. Riis [24]. The aim

---

[7]This area has been described to me as "a young person's job". Regretfully I have no personal experience of it.
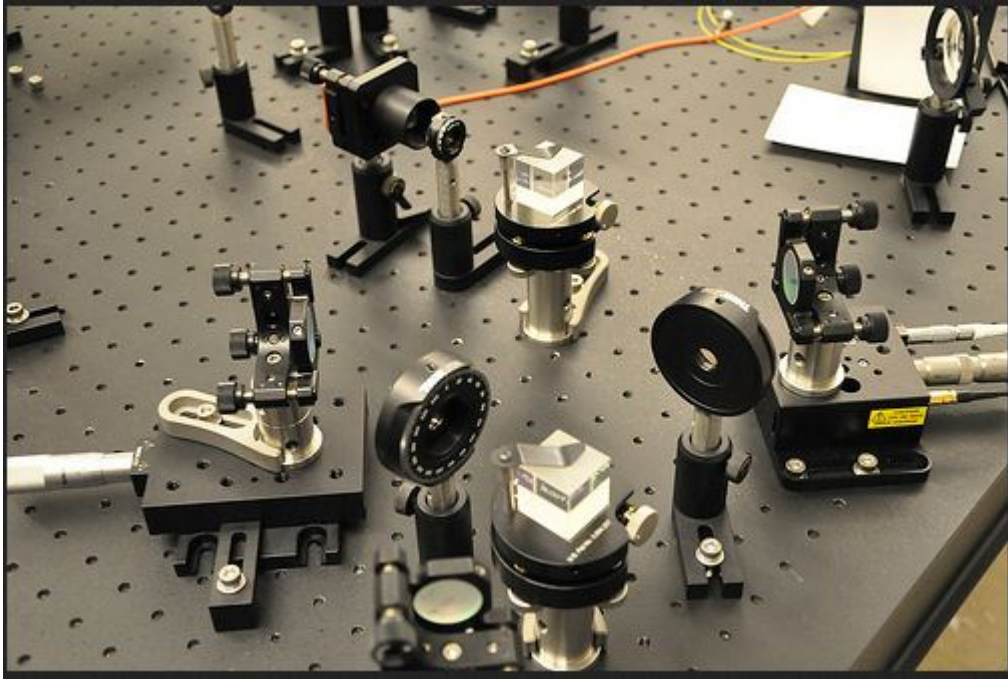
Figure 1.5: Quantum optics laboratory (Courtesy of Department of Physics and Astronomy, Dickinson College, Carlisle, PA).

is to distinguish between three possible states of a polarised photon:

$$
\begin{aligned}
|\psi_1\rangle &= |H\rangle \\
|\psi_2\rangle &= -\frac{1}{2}|H\rangle + \frac{\sqrt{3}}{2}|V\rangle \\
|\psi_3\rangle &= -\frac{1}{2}|H\rangle - \frac{\sqrt{3}}{2}|V\rangle
\end{aligned}
\tag{1.44}
$$

where $|H\rangle$ and $|V\rangle$ stand for the horizontally and vertically polarised states, respectively. In figure 1.7, the unknown photon is injected at the bottom, and is seen by one of the three detectors PD1, PD2 and PD3. The accuracy is that predicted by the Helstrom limit.

It is often very difficult to devise a configuration that will carry out a generalised (that is, a non Von Neumann) measurement directly. However this can always be done by taking another quantum system known as an 'ancilla', reacting it with the system being measured so that they become entangled, and making a Von Neumann measurement on the ancilla. The result will give information about the original system.

1.7   Qbits

An important category of quantum systems includes those which have certain observables with only two eigenvalues, which we call $|0\rangle$ and $|1\rangle$. As superposition is allowed, the state of such a
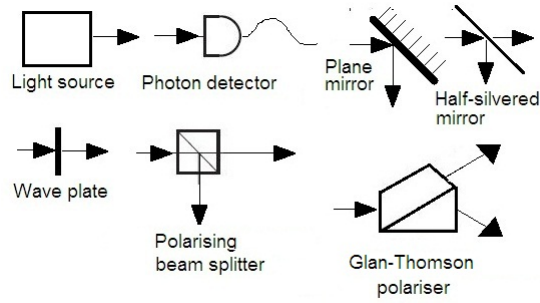
Figure 1.6: Optical components
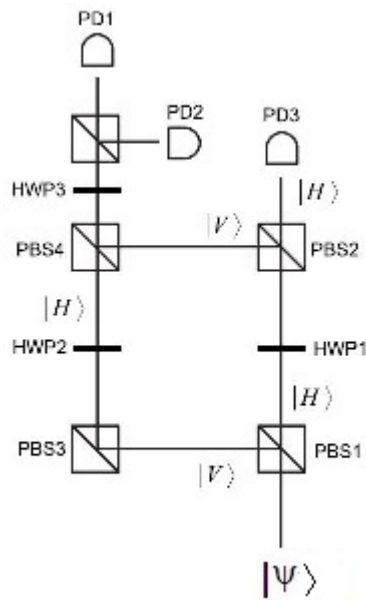
Figure 1.7: Discriminator for three states.

system can be represented by the equation

$$|\psi\rangle = a|0\rangle + b|1\rangle. \tag{1.45}$$

The category includes photons with vertical or horizontal polarisation, particles in the 'spin-up' or 'spin-down' state, and certain species of atom which can exist in one of two energy states. These systems are somewhat analogous to classical binary digits, and form a valuable resource in quantum information studies. They are called *qbits* [8].

There are several important differences between a qbit and a classical bit:

- Classical bits are stable; there is no reason that they should not endure indefinitely. On the

---

[8]This word has alternative spellings, and many authors use "qubits". I prefer the version used by [27].

other hand qbits are subject to corruption by external noise, and can be quite extremely unstable. Error correction techniques offer some mitigation to this problem, but not much.

- Classical bits have only two possible values, but as superposition exists, a qbit can possess an uncountable infinity of different states. The only states that can be measured with certainty are those where there is no superposition.

- Classical bits can be replicated without limit, but it is in general impossible to 'clone' a qbit [25]. Here is a simple demonstration of the fact, as given in [13]:

We could start with an unknown qbit $|\psi\rangle$ and a 'blank' qbit $|B\rangle$. Suppose that cloning works for the two special states $|\psi\rangle = |0\rangle$ and $|\psi\rangle = |1\rangle$. Then

$$|0\rangle \otimes |B\rangle \rightarrow |0\rangle \otimes |0\rangle$$
$$|1\rangle \otimes |B\rangle \rightarrow |1\rangle \otimes |1\rangle.$$

Now consider the more general state $|\psi\rangle = a|0\rangle + b|1\rangle$. It will be transformed as

$$(a|0\rangle + b|1\rangle) \otimes |B\rangle \rightarrow a(|0\rangle \otimes |0\rangle) + b(|1\rangle \otimes |1\rangle). \tag{1.46}$$

This is not the same as a pair of copies of the original state, which is

$$(a|0\rangle + b|1\rangle) \otimes (a|0\rangle + b|1\rangle) = a^2(|0\rangle \otimes |0\rangle) + ab(|0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle) + b^2(|1\rangle \otimes |1\rangle). \tag{1.47}$$

The impossibility of cloning has important practical consequences. As we have seen, it is (in general) impossible to measure the state of a qbit accurately. If cloning were possible, one could generate a large number of clones, measure them all, and use statistics to determine the state of the original qbit. But this cannot be done.

It is convenient to write the state of a qbit in terms of angles on the Bloch sphere:

$$|\psi\rangle = \cos(\theta/2)|0\rangle + e^{i\phi}\sin(\theta/2)|1\rangle. \tag{1.48}$$

1.8 The behaviour of atoms under radiation

Many techniques in quantum physics rely on illuminating atoms or ions with radiation. There follows a very brief, simplified overview of the resulting phenomena.

Radiation consists of a stream of photons. A photon behaves both as a massless particle and as a wave of a certain frequency. The energy of the photon is $\hbar\omega$, where ω is the angular frequency of the photon and $\hbar$ is Planck's constant.

An atom consists of a massive, positively charged central nucleus, surrounded by negative electrons arranged in orbits or *shells*. The number of electrons depends on the species of the atom and varies from 1 (Hydrogen) to over 100 for some heavy elements. The number of electrons in each shell is limited, and all except the outermost shell are full. If any electrons are removed from the outer shell, or extra ones added, the atom becomes ionised. The atom acquires an

electric charge (which may be of either sign), and its chemical properties, which depend on the number of electrons in the outermost orbit, are profoundly modified.

When an atom or ion is in its 'ground state', the electrons in the outermost shell are in their lowest energy states. However, when the atom is disturbed, by radiation or other means, an electron may jump or 'transit' to an orbit of higher energy. Each change of orbit has an associated energy difference. Once it has reached the higher energy orbit, the atom may remain stable for a while, but eventually it will drop back to the lower energy, and emit a photon of exactly the right frequency to take the extra energy away. This is called 'fluorescing'.

One way of producing an 'upward' transition to a higher energy level is to illuminate the atom with a photon with precisely the right energy needed. If the atom is illuminated continuously it will alternate between the high and low energy states, and execute a Rabi cycle [7], with a frequency that depends on the strength of the illumination.

1.9   Random telegraph signals

When a two-valued variable - such as a voltage in a digital circuit - switches unpredictably between its two states, the phenomenon is called *Random telegraphing*.

Many species of atom have more than one excited state, each with its own characteristic energy gap above the ground state. In certain atoms, such as ionised Barium, one energy level is close to the ground state, so that a weak transition is excited by a relatively low energy photon ("red") while the other is higher, and needs a more energetic photon ("blue") to cause a strong transition. This is shown in figure 1.8. Strong transitions are relatively easy to detect; when the atom is illuminated with photons of the right energy it fluoresces with such brilliance that the light can be seen with the naked eye. Weak transitions are much harder to observe.

Dehmelt [14] set out to detect weak transitions, by illuminating a three-state atom with both strong blue and weak red light. For much of the time the atom responds to the blue light, and fluoresces visibly, just like a two-state atom. However, when the atom responds to the red light and makes a weak transition, it can no longer react to the blue light, because a direct transition from one excited state to another cannot occur. Fluorescing stops until the atom drops spontaneously back to the ground level.

A photocell sensitive to blue light will observe bursts of radiation separated by periods of darkness. Each dark period corresponds to the time that the atom is in the lower excted state, and gives an indication of a weak transition. This alternation between light and dark is called *telegraphing*. The duration of each state (light vs. dark) is referred to as a *dwell*.

Suppose the probability of a transition during a time period δt is $p \, \delta t$. As δt tends to 0, this gives a random dwell duration with a negative exponential distribution of e^{-p} .
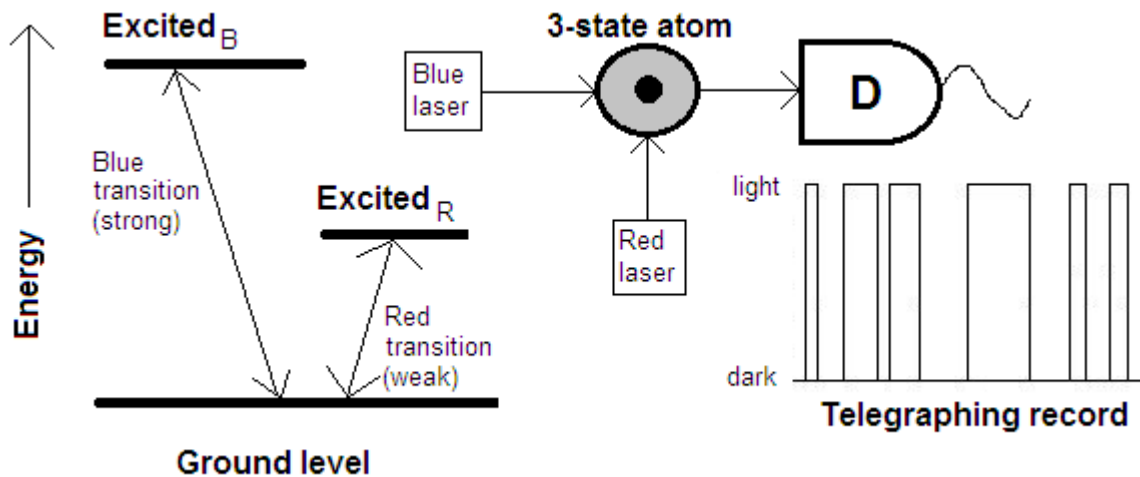
Figure 1.8: States of a three-state atom

1.10   Imperfect measurements

We have seen that a perfect measurement of a system in the state

$$|\psi\rangle = a|0\rangle + b|1\rangle$$

will return state $|0\rangle$ with probability aa^* and state $|1\rangle$ with probability bb^*, and leave the system in one of the two eigenstates. However, an *imperfect* measurement, which can sometimes deliver results not in keeping with this rule, will leave the system in an intermediate state. In particular, a measurement so bad that it gives hardly any information about the system will alter its state very little.

There is, of course, no practical purpose in making bad measurements. On the other hand, sequences of imperfect measurements that make only minute changes to the state of a system can act as a surrogate for other types of interference, such as stray photons or proximity to other systems, while being more straightforward to handle mathematically. For this reason, the analysis of sequences of imperfect measurements is a topic of current interest, and relevant to part of this thesis.

1.11   Open systems

Schrődinger's equation describes the reversible evolution of a *closed* quantum system which has no interaction with the environment. We have seen that whenever a system is measured, it undergoes a sudden change, and Schrődinger's equation no longer holds.

Whenever a system has any link with its surroundings, either through the transfer of energy or by measurement, it is called *open*, and its evolution through time is generally irreversible. It is governed by a *Master equation*.

There are several varieties of Master equation. Perhaps the simplest is the *Lindblad form* [28] , which depends on two key assumptions:

- The environment is so huge compared to the system under investigation that it is not significantly altered by the evolution of the system

- The evolution is *Markovian*; that is, it depends solely on the current state of the system, and in no way on its past history.

The Lindblad form of the Master equation defines the rate of change of $\hat{\rho}_S$ with time, as a function of its current state and various time-independent operators. Following Salo, Stenholm, Kurizki and Kofman [29] the basic form of the Lindblad equation can be derived by using several key considerations:

- Quantum mechanics is linear, so we seek a *linear* form

- Probability is conserved; the trace of a density matrix at any time t is always 1 $(\mathrm{Tr}(\hat{\rho}(t)) = 1 \, \forall t)$

- $\hat{\rho}$ is always hermitian $(\hat{\rho}(t) = \hat{\rho}^{\dagger}(t) \, \forall t)$

- $\hat{\rho}$ is always positive $(\langle \psi | \hat{\rho} | \psi \rangle \geq 0 \, \forall \psi)$.

Applying these constraints, [30] it can be shown that the *most general* linear transformation $\mathcal{T}$ is:

$$\mathcal{T}(\hat{\rho}) = \hat{B}\hat{\rho} + \hat{\rho}^{\dagger}\hat{B}^{\dagger} + \sum_{j} \hat{K}_{j}\hat{\rho}\hat{K}_{j}^{\dagger}. \tag{1.49}$$

The $\hat{K}$s are called *Kraus* operators. $\hat{B}$ and the product $\hat{K}_{j}\hat{K}_{j}^{\dagger}$ must be hermitian, but this does not necessarily apply to the individual operators $\hat{K}_{j}$.

We now specify a transformation that yields the derivative of $\hat{\rho}$ with respect to time.

$$\mathcal{T}(\hat{\rho}) = \frac{d\hat{\rho}}{dt}. \tag{1.50}$$

As the trace of $\hat{\rho}$ is always 1, the trace of its derivative must be zero. Symbolically:

$$\mathrm{Tr}(\mathcal{T}(\hat{\rho})) = 0. \tag{1.51}$$

Next, we replace $\hat{B}$ by $(\hat{P} + i\hat{H})$ and $\hat{B}^{\dagger}$ by $(\hat{P} - i\hat{H})$ where $\hat{P}$ and $\hat{H}$ are hermitian. We can now use the cyclic property of the Trace operator, and the fact that $\mathrm{Tr}([\hat{A}, \hat{B}]) = 0$ $(\forall \hat{A}, \hat{B})$ to show that

$$\mathcal{T}(\hat{\rho}) = \frac{d\hat{\rho}}{dt} = \frac{-i}{\hbar}[\hat{H}, \hat{\rho}] + \sum_{j}(\hat{K}_{j}\hat{\rho}\hat{K}_{j}^{\dagger} - \tfrac{1}{2}\hat{K}_{j}^{\dagger}\hat{K}_{j}\hat{\rho} - \tfrac{1}{2}\hat{\rho}\hat{K}_{j}^{\dagger}\hat{K}_{j}). \tag{1.52}$$

This is the required Lindblad form. The term $\frac{-i}{\hbar}[\hat{H}, \hat{\rho}]$ represents the closed Schrődinger evolution, $\hat{H}$ being the hamiltonian. The other term gives the effect of interaction with the environment. pecification of the Kraus operators is an *ad hoc* process, depending on the problem under consideration.

29

1.12   History of quantum information

The era of Quantum Physics began in the first half of the twentieth century, and was based on the work of Planck, Bohr, Einstein, Heisenberg, Schrődinger, Ehrenfest, Dirac, von Neumann, and others. Initially experiments were relatively crude, and could only be made by studying the behaviour of large ensembles of particles. The properties of individual quantum systems were little more than far-sighted hypotheses, just as, a couple of centuries earlier, Lavoisier and Dalton originated atomic theory on the basis of simple experiments with substantial quantities of various substances.

Practical developments during the first part of the era all harnessed the quantum properties of ensembles of particles. Examples include atomic reactors, atomic weapons, the transistor, and the laser.

From about 1980, technical developments made it possible to observe the properties of individual systems. For example, photo-detectors became sensitive enough to respond to individual photons, and a single atom could be suspended in a vacuum and subjected to radiation of various kinds. This led to the possibility that individual quantum systems could carry information. Thus a photon can be either vertically or horizontally polarised, and its state can represent a single binary digit, commonly called a *qbit*.

One valuable application of quantum information is in communication. Alice[9] transmits a stream of information to Bob by a sequence of photons, polarised in one or the other direction. Bob measures them to recover the information.

A key feature of quantum measurement is that it cannot generally be done without changing the system that is measured. This is used in various protocols [31], to ensure that if there is an eavesdropper, the fact is known; the stream can be discarded and another one sent in its place. Of course this only makes sense if the stream of information is not a valid message in its own right, but merely a set of random digits which will form the key for encoding a 'real' message that will be transmitted by classical means. If the key is as long as the message itself, the coded form will be, in principle, unbreakable.

The photons that Alice sends to Bob travel along fibre-optic cables which attenuate the signal. This limits the distance over which transmission can take place. Intermediate amplifiers cannot be used because, as we have seen, it is impossible to clone a quantum state without error.

At the time of writing, quantum communication is limited to a range of some tens of kilometres, but the quality of fibre optic cables, and the distance that they can reliably carry single photons, is being continuously improved.

Another potential application of quantum information is in quantum computation. Qbits can be input to logic gates which invert and combine them in various ways, to calculate useful results. A major theoretical advantage is that using superposition, a quantum computer can effectively carry out many calculations at the same time, making it much faster than any classical computer. There are, however, two key problems to be overcome before this can be put to good use.

---

[9]Alice and Bob are accepted symbols for entities (people or machines) exchanging confidential information. Eve tries to read or corrupt this information in transit (because she 'eavesdrops').

- Quantum computation requires the accurate storage of qbits for significant lengths of time. With current technology, the probability of a qbit *decohering* and changing its state increases rapidly with time. It is possible, even now, to build very small quantum computers with perhaps a dozen qbits, which have a significant chance of completing a calculation before any of the qbits decohere sufficiently to invalidate the result; but any serious application would need hundreds or even thousands of qbits. The stability of a qbit must increase by several orders of magnitude before quantum computers become of any practical use.

- There are very few known problems where a quantum computer would gain a significant advantage over a classical machine. Two examples are

  - Grover's Search Algorithm [32]
  - Shor's Factorisation algorithm [33].

At present, the most widely used method for secret communication over public networks is public key encryption coding [34]. The key used is the product of two very large prime numbers (several hundred decimal digits). The code would be broken if the key could be factorised, but there is no known classical algorithm that could do this in an acceptable time-scale. However, Shor's algorithm could do the task if it could be implemented on a quantum computer. This has led to intense interest in quantum computing from governments anxious to read messages sent by terrorists and other methods.

The technique of quantum computing was first suggested by Richard Feynman [35] in the early 1980's, but progress has been slow. The best known achievement to date has been the factorisation of 15 to give ($15 = 5 \times 3$).

Chapter 2

Programmed Discrimination between Quantum States

" Everything should be made as simple as possible, but not simpler" *Albert Einstein*

This chapter describes recent work most relevant to our own research. It offers a time-line of developments in Programmed Discrimination, and a detailed summary of a paper that was key to our work.

2.1 Classical and quantum computing

The earliest digital computing machines were designed to carry out fixed functions, such as multiplication and calculation of square roots, by mechanical means. A major step in their evolution occurred when Charles Babbage [36] designed the Analytical Engine, a machine that could be adapted to carry out a wide range of different functions, using a fixed mechanism and supplied with a set of instructions as *data*. This is the fundamental configuration used in all present-day (classical) computers.

It might be that quantum computers will follow a similar line of evolution, but every system proposed at present depends on hardware dedicated to solving one specific problem such as searching or factorisation.

The term *programmable quantum computer* has received two different interpretations. On one understanding, it relates to *universal* quantum gates that can be switched, by classical control inputs, to perform a flexible range of different functions. Significant advances in this area have already been reported [37] . The other understanding refers to assemblies of gates whose combined unitary function is controlled entirely by quantum inputs, so retaining the adaptability of a general-purpose computer without the need to carry out measurements of intermediate results. It is this latter interpretation that is of interest here. No design for a general-purpose quantum computer has yet been proposed, but it seems likely that programmed discrimination will play a major role in extracting the result of any algorithm that such a machine might carry out.

2.2 History of programmed discrimination

In the first chapter of this thesis, we discussed measurement of quantum systems in absolute terms, such as finding the actual spin of a particle or the current energy status of an atom. We also considered discrimination between two or more known states. In general, measurements cannot be both exact and fully determined, but there is a choice to be made. Either

- each measurement has a definite error rate, determined by the Helstrom limit, or

- error-free measurements can sometimes be made, but at a high price, as the most common outcome of the process is to yield no information at all.

These two modes are termed **optimal** [17] and **unambiguous**[19, 20, 21], respectively. Both play major parts in this thesis.

When groups of two or more systems are involved, we can use different forms of measurement: we can try to determine the *relationship* between the systems. For example, consider a pair of qbits, A and B. The result of a measurement on the pair might be that A is in the same state as B, or not in the same state. This is information about the pair as a whole, and tells us nothing about the actual states of A and B as individual systems. Nevertheless such information can be extremely useful. To give an example, a qbit which is transmitted over a long distance, or stored for an appreciable length of time, may undergo unitary evolution which will render attempts to recognise its original state useless. On the other hand a pair of qbits in the same circumstances will both undergo the same evolution, so the relationship between them will be maintained and can yield reliable information.

The idea of such measurement first emerged in the 1990s, following initial developments in quantum information studies. In an early paper Jaeger and Shimony [38] considered the discrimination of two quantum states that are not necessarily orthogonal, and suggested procedures for the task, together with estimates of the probability of success. In 2003 the topic was taken up by Barnett, Chefles and Jex [39, 49]. They were the first to propose the use of POMs which can sometimes deliver unambiguous measurements, at the cost of frequent inconclusive results.

The notion of *programmable* state discrimination was introduced by Bergou and Hillery [40] in 2005. To quote from their paper:

We shall, therefore, consider the following problem which is perhaps the simplest version of a programmable state discriminator. The program consists of the two qubit states that we wish to distinguish. In other words, we are given two qubits: one in the state $|\psi_1\rangle$ and another in the state $|\psi_2\rangle$. We have no knowledge of the states $|\psi_1\rangle$ and $|\psi_2\rangle$. Then we are given a third qubit that is guaranteed to be in one of these two program states, and our task is to determine, as best as we can, in which one. We are allowed to fail, but not make a mistake. What is the best procedure to accomplish this?

The paper presents the task as a measurement optimisation problem which may, or may not, lead to an unambiguous result. The authors also include another innovation: instead of considering the two possibilities to be equally likely, they assign a variable probability, η_1, that the third qbit is in the same state as $|\psi_1\rangle$, and $\eta_2 = (1 - \eta_1)$ to the other possibility. This leads to a range of results, with the interesting property that when η_1 is near to one or zero, only one of the possibilities can be recognised unambiguously; the other will never be recognised at all.

In 2006 Hayashi, Horibe and Hashimoto [41] extended the subject to consider multiple copies of the program qbits. 2006 also saw the publication of a paper by Bergou, Bužek, Feldman, Herzog and Hillery [42] which tied together all existing results and considered several new aspects of the problem, including

- *Optimal discrimination* (as opposed to unambiguous discrimination). An optimal procedure is not allowed to fail, but it may return an incorrect result. The problem here is to make the error rate as small as possible

- Using multiple copies of the various qbits

- Assuming that one of the program qbits is either known exactly, or completely unavailable.

This paper proved so important for our own work that we summarise its findings in the final section of this chapter.

Another useful paper by Bergou [43] appeared in 2007. It discusses the application of both discrimination strategies to the B92 quantum key distribution protocol [44]. In 2008, Herzog and Bergou [45] again extended the area to consider discrimination of *qudit* states; that is, of systems that can have n different internal states, where $n \geq 2$. A paper by Hillery, Andersson, Barnett and Oi entitled *Decision problems with black boxes* [46] appeared in 2010. It drew the analogy between the states $|\psi_1\rangle$ and $|\psi_2\rangle$ and operators, represented by black boxes.

Finally we mention our own papers, which contribute to this area. With the exceptions of [40] and [42], which mention the case where the program qbits have a known average overlap, all previous work assumes that this relationship is unknown, and presents symbolic results as averages over all possible cases. This generally allows analytic methods to be used. In contrast, both our papers assume that we do have classical knowledge of the relationship between the program qbits. We have selected three particular types of relationship:

- A *known* overlap between the reference qbits

- Both reference qbits on the same *known* great circle of the Bloch sphere

- Both qbits confined to *known* areas centred on the poles of the Bloch sphere.

In 2011 we published *Programmed discrimination of qbits with added classical information* [47]. This paper uses numerical methods to explore both optimal and unambiguous recognition rates for each type of relationship. In 2012, in *Programmed discrimination of multiple sets of qbits*
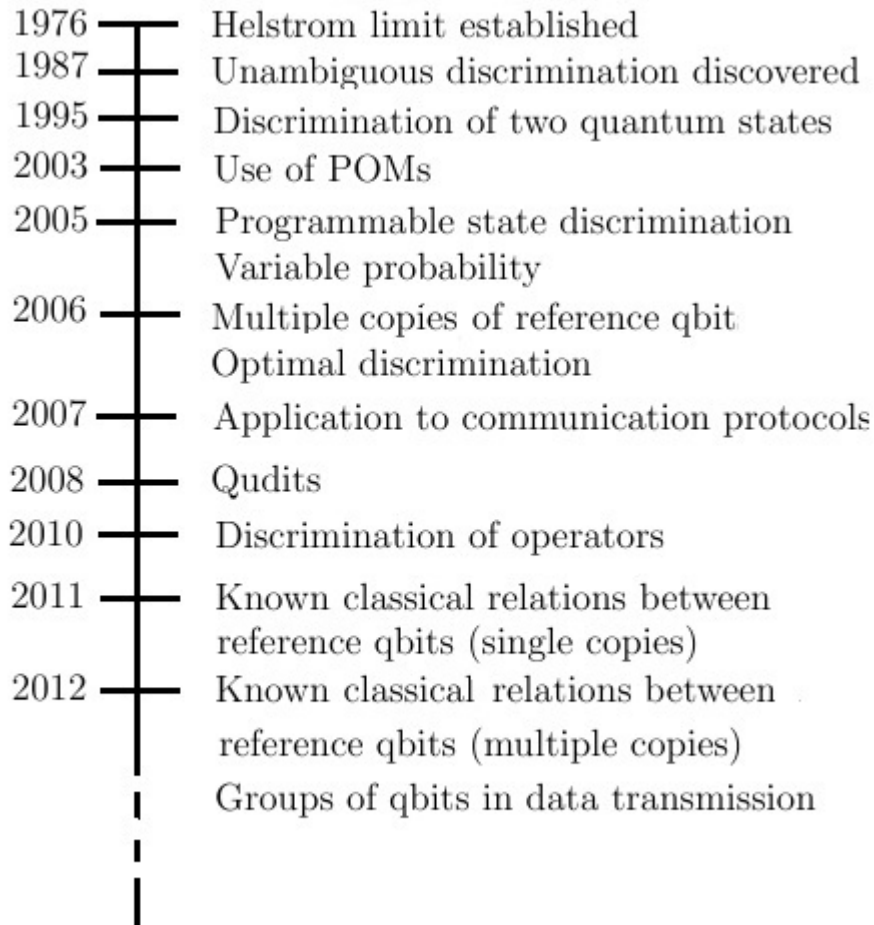
34

1976 — Helstrom limit established
1987 — Unambiguous discrimination discovered
1995 — Discrimination of two quantum states
2003 — Use of POMs
2005 — Programmable state discrimination
Variable probability
2006 — Multiple copies of reference qbit
Optimal discrimination
2007 — Application to communication protocols
2008 — Qudits
2010 — Discrimination of operators
2011 — Known classical relations between
reference qbits (single copies)
2012 — Known classical relations between
reference qbits (multiple copies)
Groups of qbits in data transmission

Figure 2.1: Developments in programmed discrimination

with added classical information [48] we extend this study to multiple qbits, including non-symmetrical arrangements, and consider the performance of groups of qbits in the context of data transmission. These two papers constitute much of the work reported in this thesis, and form the basis for Chapters 3 and 4. Figure 2.1 gives a timeline of developments in programmed discrimination.

2.3 Summary of a key paper

In this section we present a brief summary of the paper entitled *Programmable quantum-state discriminators with simple programs* by Bergou, Bužek, Feldman, Herzog and Hillery [42], whom we term the *Hunter Group* by reason of their academic affiliation. The starting point is a *triad* of qbits. Two of them, $|\psi_1\rangle$ and $|\psi_3\rangle$ are the *program* qbits, and the third, $|\psi_2\rangle$), is the *data* qbit. The data qbit is guaranteed to be identical to one or other of the program qbits, and the task is

to discover which. The paper discusses the theoretical limits of such a discrimination process.

Consider a Triad. As $|\psi_2\rangle$ must be the same as either $|\psi_1\rangle$ or $|\psi_3\rangle$, there are only two possibilities for the input state of the discrimination device:

$$
\begin{aligned}
|\Psi_1^{in}\rangle &= |\psi_1\rangle_A |\psi_1\rangle_B |\psi_3\rangle_C \\
|\Psi_2^{in}\rangle &= |\psi_1\rangle_A |\psi_3\rangle_B |\psi_3\rangle_C.
\end{aligned}
\tag{2.1}
$$

As both $|\psi_1\rangle$ and $|\psi_3\rangle$ are unknown, the Hunter Group finds two density matrices, both averaged over the Bloch sphere:

$$
\begin{aligned}
\overline{\hat{\rho}}_1 &= \overline{|\Psi_1^{in}\rangle\langle\Psi_1^{in}|} \\
\overline{\hat{\rho}}_2 &= \overline{|\Psi_2^{in}\rangle\langle\Psi_2^{in}|}.
\end{aligned}
\tag{2.2}
$$

The states $|\psi_1\rangle$ and $|\psi_3\rangle$ are independent of one another, so $\overline{\hat{\rho}}_1$ can be written as the tensor product $\overline{\hat{\lambda}}_1 \otimes \overline{\hat{\mu}}_1$, where

$$
\begin{aligned}
\overline{\hat{\lambda}}_1 &= \overline{|\psi_1\rangle|\psi_1\rangle\langle\psi_1|\langle\psi_1|} \\
\overline{\hat{\mu}}_1 &= \overline{|\psi_3\rangle\langle\psi_3|}.
\end{aligned}
\tag{2.3}
$$

Both $\overline{\hat{\lambda}}_1$ and $\overline{\hat{\mu}}_1$ are averaged over the whole Bloch sphere.

For a general qbit $|\psi_1\rangle = \cos(\beta/2)|0\rangle + e^{i\alpha}\sin(\beta/2)|1\rangle$, the product state is the four-component vector:

$$
\begin{aligned}
|\psi_1\rangle|\psi_1\rangle &= (\cos(\beta/2)|0\rangle + e^{i\alpha}\sin(\beta/2)|1\rangle)^{\otimes 2} \\
&= \begin{pmatrix} \cos^2(\beta/2) & |00\rangle \\ +e^{i\alpha} \ \sin(\beta/2)\cos(\beta/2) & |01\rangle \\ +e^{i\alpha} \ \sin(\beta/2)\cos(\beta/2) & |10\rangle \\ +e^{2i\alpha} \ \sin^2(\beta/2) & |11\rangle \end{pmatrix}
\end{aligned}
\tag{2.4}
$$

so that $\overline{\hat{\lambda}}_1$ is the 4×4 matrix which is the tensor product of this vector and its complex conjugate, where each element is averaged over the Bloch sphere. Computing this average gives:

$$
\begin{aligned}
\overline{\hat{\lambda}}_1 &= \frac{1}{4\pi}\int_0^{2\pi} d\alpha \int_0^{\pi} d\beta |\psi_1\rangle|\psi_1\rangle\langle\psi_1|\langle\psi_1| \\
&= \frac{1}{6}\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix},
\end{aligned}
\tag{2.5}
$$

where the rows and columns correspond, in order, to the states $|00\rangle$, $|01\rangle$, $|10\rangle$ and $|11\rangle$.

Alternatively, $|\Psi_1^{in}\rangle$ is *symmetrical* in the two components $|\psi_1\rangle_A|$ and $\psi_1\rangle_B$. In general, any system of two components $|\psi\rangle_A \otimes |\psi\rangle_B$ can be in a superposition of four possible states:

$$
\begin{aligned}
|u_1\rangle &= |0\rangle|0\rangle \\
|u_2\rangle &= |1\rangle|1\rangle \\
|u_3\rangle &= \frac{1}{\sqrt{2}}(|0\rangle|1\rangle + |1\rangle|0\rangle) \\
|u_4\rangle &= \frac{1}{\sqrt{2}}(|0\rangle|1\rangle - |1\rangle|0\rangle)
\end{aligned}
\tag{2.6}
$$

where $\sum_{j=1}^4 u_j = \hat{\mathbf{1}}$. But the symmetry of $|\Psi_1^{in}\rangle$ implies that $|\psi_1\rangle_A$ and $\psi_1\rangle_B$ can be interchanged, so the coefficient of $|u_4\rangle$, which is usually referred to as the asymmetric Bell state $|\psi^-\rangle$, must be 0. It follows that $\overline{\hat{\lambda}}_1$ can be expressed in the normalised form

$$
\overline{\hat{\lambda}}_1 = \frac{1}{3}(\hat{\mathbf{1}} - |\psi^-\rangle\langle\psi^-|)
\tag{2.7}
$$

$\overline{\hat{\lambda}}_1$ is therefore proportional to the projector on to the *symmetric* two-qbit subspace.

A similar calculation shows that $\overline{\overline{\mu}}$ is proportional to the identity operator:

$$
\begin{aligned}
\overline{\overline{\mu}}_1 &= \frac{1}{4\pi} \int_0^{2\pi} d\alpha \int_0^{\pi} d\beta |\psi_3\rangle\langle\psi_3| \\
&= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
\end{aligned}
\tag{2.8}
$$

Tensor multiplication of these gives

$$
\overline{\overline{\rho}}_1 = \overline{\hat{\lambda}}_1 \otimes \overline{\overline{\mu}}_1
\tag{2.9}
$$

This density matrix applies to the first of the two states, where the first and the second qbits are identical. Similarly, if the second and third qbits are identical,

$$
\overline{\overline{\rho}}_2 = \overline{\overline{\mu}}_1 \otimes \overline{\hat{\lambda}}_1
\tag{2.10}
$$

These two density matrices form the basis for calculating optimum error rates for the minimum error strategy.

2.3.1 Minimum error

Suppose that η_1 is the *a priori* probability that the data qbit is the same as $|\psi_1\rangle$, and η_2 the probability that the data qbit is the same as $|\psi_3\rangle$. Clearly, $\eta_1 + \eta_2 = 1$. To find the minimum error the Hunter Group compute the difference matrix:

$$
\hat{\Lambda} = (\eta_1\overline{\overline{\rho}}_1 - \eta_2\overline{\overline{\rho}}_2)
\tag{2.11}
$$

Following Helstrom [17], the lowest probability of error (P.E.) that can be achieved in the discrimination is given by :

$$
\text{P.E.} = \frac{1}{2}(1 - z)
\tag{2.12}
$$

where z is the sum of the absolute values of the eigenvalues of $\hat{\Lambda}$. It is worth noting that when $\eta_1 = \eta_2 = 1/2$, so that there is no bias either way, and in the absence of any classical data, P.E. ≈ 0.356. As η_1 tends to 1 (or zero) the probability of error falls accordingly, in a nearly linear way.

2.3.2 Unambiguous discrimination

The Hunter Group's method is an adaptation of the process described in [40], which we summarise below. The method depends on finding an unambiguous discrimination POM that consists of the set of operators $\{\hat{\pi}_0, \hat{\pi}_1, \hat{\pi}_2\}$ such that $\hat{\pi}_1$ is *orthogonal* to $\hat{\rho}_2$, $\hat{\pi}_2$ is *orthogonal* to $\hat{\rho}_1$, and $\hat{\pi}_0 = I - \hat{\pi}_1 - \hat{\pi}_2$. Possible measurement outcomes mean:

- $\hat{\pi}_1$: The system is certainly in state $\hat{\rho}_1$

- $\hat{\pi}_2$: The system is undoubtedly in state $\hat{\rho}_2$

- $\hat{\pi}_0$: No information about the state is available, and the original state is changed by the measurement so that further measurements are useless.

As $\hat{\rho}_1$ is symmetric in qbits $|\psi_1\rangle_A$ and $|\psi_1\rangle_B$, $\hat{\pi}_2$ can simply be a multiple of the projector on to the (zero) asymmetric sub-state of these two qbits, so that $\langle \Psi_1 | \hat{\pi}_2 | \Psi_1 \rangle = \mathrm{Tr}(\hat{\rho}_1 \hat{\pi}_2) = 0$.

This projector is

$$|\psi^-\rangle_{AB\,AB}\langle\psi^-| \tag{2.13}$$

where $|\psi^-\rangle_{AB}$ is the antisymmetric state:

$$|\psi^-\rangle_{AB} = \frac{1}{\sqrt{2}}(|\psi_1\rangle_A |\psi_1^\perp\rangle_B - |\psi_1^\perp\rangle_A |\psi_1\rangle_B) \tag{2.14}$$

.

Given that the system is in state $|\Psi_j^{in}\rangle$, (j=1,2) the overall probability of correct identification, p_j, is

$$p_j = \langle \Psi_j^{in} | \hat{\pi}_j | \Psi_j^{in} \rangle \tag{2.15}$$

Substituting for $\hat{\pi}_j$ in this equation, and simplifying, it can be shown that

$$p_j = c_j/2(1 - |\langle \psi_1 | \psi_3 \rangle|^2)(j = 1, 2) \tag{2.16}$$

where c_1 and c_2 are unknown real positive constants and $\langle \psi_1 | \psi_3 \rangle$ is the overlap between the two program qbits.

It remains only to find optimum values for c_1 and c_2, while ensuring that $\hat{\pi}_0$ remains a positive operator. The outcome depends on η_1 and η_2, the *a priori* probabilities that the system is in state $|\Psi_1^{in}\rangle$ or $|\Psi_2^{in}\rangle$. For a broad range, centred on $\eta_1 = 1/2$, the analysis shows that

$$c_1 = \frac{2}{3}(2 - \sqrt{\eta_2/\eta_1}) \tag{2.17}$$

$$c_2 = \frac{2}{3}(2 - \sqrt{\eta_1/\eta_2}) \tag{2.18}$$

The overall probability of correct unambiguous identification of the qbit is given by

$$P = \eta_1 p_1 + \eta_2 p_2 \qquad (2.19)$$

Substitution from equations (14)-(17) shows that

$$P = \frac{2}{3}(1 - \sqrt{\eta_1 \eta_2})(1 - |\langle \psi_1 | \psi_3 \rangle|^2) \qquad (2.20)$$

For this to be valid both c_1 and c_2 must be non-negative, and this is only true when $\frac{1}{5} \leq \eta_1 \leq \frac{4}{5}$. Outwith this range it is better to use a projective von Neumann measurement aligned with the more frequent possibility. Here the probability of correct discrimination for $\eta_1 > \frac{4}{5}$ is

$$P_{high\ \eta_1} = \frac{1}{2}\eta_1(1 - |\langle \psi_1 | \psi_3 \rangle|^2) \qquad (2.21)$$

Conversely, if $\eta_1 < \frac{1}{5}$,

$$P_{low\ \eta_1} = \frac{1}{2}\eta_2(1 - |\langle \psi_1 | \psi_3 \rangle|^2). \qquad (2.22)$$

In the case considered by the Hunter group, the position of the program qbits is undefined, and their average overlap is $\frac{1}{\sqrt{2}}$. This leads directly to their formula for failure to make an unambiguous discrimination in the central range of η of

$$Q_F = \frac{2 + \sqrt{\eta_1 \eta_2}}{3}. \qquad (2.23)$$

This completes our review of some of the Hunter Group's principal results. They will be fundamental for our own findings, which we report in the following chapters of the thesis. In particular, we note that the success of unambiguous discrimination depends only on two factors: the *a priori* frequencies η_1 and η_2, and the overlap between the program qbits.

The configuration of qbits directly of interest to our work all include two program qbits (possibly in multiple copies) and one data qbit (also possibly in multiple copies).

Table 2.1 summarises the analytic error rates results derived by Bergou and his colleagues. In the table we refer to a configuration that has j copies of program qbit $|\psi_1\rangle$, k copies of the data qbit, and m copies of the other program qbit as $\{j, k, m\}$. η_{min} is the *smaller* of $\{\eta_1, \eta_2\}$ and η_{max} the *larger*.

Table 2.2 shows the failure rates for unambiguous discrimination.

These expressions are useful in our own research, as they represent special cases of the configurations that we consider and offer confirmation of our numerical computations.

Configuration	Condition
$\{1,1,1\}$	$P.E = \eta_{min}\left(1 - \frac{1}{2}\frac{\eta_{max}}{\eta_{max}-\eta_{min}+\sqrt{1-\eta_{max}\eta_{min}}}\right)$
$\{1,n,1\}$	$P.E. = \eta_{min}\left[1 - \frac{n}{n+1}\frac{\eta_{max}}{\eta_{max}-\eta_{min}+\sqrt{(\eta_{max}-\eta_{min})^2+\frac{4n(n+2)}{(n+1)^2}\eta_{min}\eta_{max}}}\right]$
$\{n,1,n\}$	$P.E = \eta_{min}\left[1 - \frac{n}{n+1}\frac{\eta_{max}}{\eta_{max}-\eta_{min}+\sqrt{(\eta_{max}-\eta_{min})^2+\frac{4n\eta_{min}\eta_{max}}{(n+1)}}}\right]$

Table 2.1: Probable error rates for optimal discrimination

Configuration	Range	Failure rate
$\{1,1,1\}$	$\eta_1 < \frac{1}{5}$	$1 - \eta_2/4$
$\{1,1,1\}$	$\frac{1}{5} \le \eta_1 \le \frac{4}{5}$	$\frac{2+\sqrt{\eta_1\eta_2}}{3}$
$\{1,1,1\}$	$\eta_1 > \frac{4}{5}$	$1 - \eta_1/4$
$\{1,n,1\}$	$\eta_1 < \frac{1}{1+(n+1)^2}$	$1 - \eta_2 n/(2n+2)$
$\{1,n,1\}$	$\frac{1}{1+(n-1)^2} \le \eta_1 \le \frac{(n+1)^2}{1+(n+1)^2}$	$1 - \frac{1}{n+2}\left[\frac{n+1}{2} - \sqrt{\eta_1\eta_2}\right]$
$\{1,n,1\}$	$\eta_1 > \frac{(n+1)^2}{1+(n+1)^2}$	$1 - \eta_1 n/(2n+2)$
$\{n,1,n\}$	$\eta_1 < \frac{2}{2+(n+1)(n+2)}$	$\eta_1 + \frac{\eta_2}{n+1}$
$\{n,1,n\}$	$\frac{2}{2+(n+1)(n+2)} \le \eta_1 \le \frac{2(n+1)}{2(n+1)(n+2)}$	$\frac{\eta_1}{2} + \frac{\eta_2}{n+2} + \sqrt{\frac{2\eta_1\eta_2}{(n+1)(n+2)}}$
$\{n,1,n\}$	$\eta_1 > \frac{2(n+1)}{2+(n+1)(n+2)}$	$\eta_1\frac{n+2}{2n+2} + \eta_2\frac{2}{n+2}$

Table 2.2: Probable failure rates for unambiguous discrimination

Chapter 3

Programmed discrimination of qbits with added classical information

"Quantum particles are the dreams that stuff is made of." *Anonymous*

3.1 Introduction

This chapter describes the work published in *Programmed discrimination of qbits with added classical information* by Colin, Barnett and Jeffers [47]. Starting from the Hunter Group's analysis described in Chapter 2, we consider programmed discrimination when there is a known classical relationship between the two program qbits.

We aim to find how these classical relationships affect the best attainable error and failure rates. We consider three specific cases, illustrated in figure 3.1:

- A known overlap between the two program qbits (figure 3.1a)

- The knowledge that both program qbits lie on a *given* great circle on the Bloch sphere. (figure 3.1b)

- The knowledge that one program qbit is restricted to a cap centered on the north pole of the Bloch sphere, and the other to a similar cap centered on the south pole. (figure3.1c)

The classical relationships are expressed as constraints on the values of the two program qbits. In keeping with the Hunter group, we also take into account the effect of different *a priori* probabilities of the two outcomes.

We refer to the set of three qbits **A**, **B** and **C**, prepared in the states { $|\psi_1\rangle_A, |\psi_2\rangle_B, |\psi_3\rangle_C$} as a *Triad*, where the subscripts {A,B,C} identify the position of the qbits in the Triad state, with the data qbit **B** placed between the other two. The state $|\psi_2\rangle$ will, of course, be either $|\psi_1\rangle$ or $|\psi_3\rangle$. η_1 and η_2 are the *a priori* probabilities of the two possible outcomes, where $\eta_1 + \eta_2 = 1$.
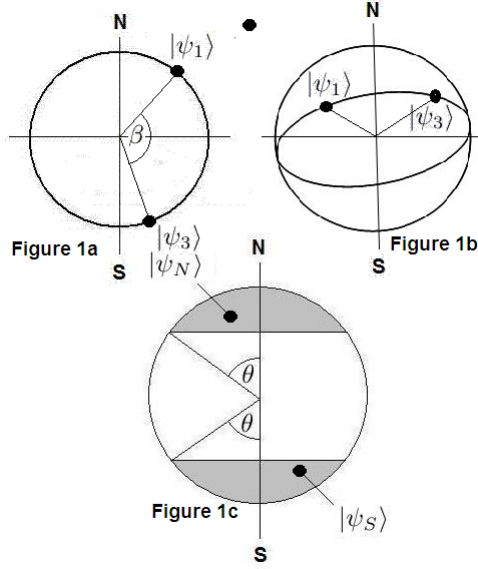
Figure 3.1: Three types of classical knowledge about a qbit

3.2 Brief overview of the Hunter Group's method

The Hunter Group's work is the paradigm on which our own research is based. To clarify the rest of the chapter we restate the Group's methods here.

The starting point is the input to the discrimination system:

$$|\Psi_1^{in}\rangle = |\psi_1\rangle_A|\psi_1\rangle_B|\psi_3\rangle_C$$
$$|\Psi_2^{in}\rangle = |\psi_1\rangle_A|\psi_3\rangle_B|\psi_3\rangle_C. \tag{3.1}$$

The next stage is to find two density matrices, averaged over the Bloch sphere:

$$\overline{\hat{\rho}_1} = \overline{|\Psi_1^{in}\rangle\langle\Psi_1^{in}|}$$
$$\overline{\hat{\rho}_2} = \overline{|\Psi_2^{in}\rangle\langle\Psi_2^{in}|}. \tag{3.2}$$

The *minimum error* rate uses Helstrom's rule:

Take

$$\hat{\Lambda} = (\eta_1\overline{\hat{\rho}_1} - \eta_2\overline{\hat{\rho}_2}). \tag{3.3}$$

Then the smallest attainable probability of error is given by

$$P.E. = \frac{1}{2}(1-z) \tag{3.4}$$

where z is the sum of the absolute values of the eigenvalues of $\hat{\Lambda}$.

Unambiguous discrimination relies on finding a set of three measurement operators $\{\hat{\pi}_0, \hat{\pi}_1, \hat{\pi}_2\}$, such that:

- $\hat{\pi}_1$ is orthogonal to $\overline{\hat{\rho}_2}$, so if it clicks the system is certainly in state $\overline{\hat{\rho}_1}$

- $\hat{\pi}_2$ is orthogonal to $\overline{\hat{\rho}_1}$, so if it clicks the system is certainly in state $\overline{\hat{\rho}_2}$

- $\hat{\pi}_0$ is a positive operator such that $(\hat{\pi}_0 + \hat{\pi}_1 + \hat{\pi}_2) = \hat{\mathbf{1}}$. If it clicks, the discrimination process has failed.

The Hunter Group's paper goes on to show that the success rate P in unambiguous discrimination depends only on the overlap between the program qbits and the values of η_1 and η_2. For the central range of η_1 and η_2 they give the relation

$$P = \frac{2}{3}(1 - \sqrt{\eta_1 \eta_2})(1 - |\langle \psi_1 | \psi_3 \rangle|^2). \tag{3.5}$$

When the *a priori* probability is close to 1 (or 0) this equation is no longer valid, and the best success rate for unambiguous discrimination is given by testing directly for the more common possibility. This implies that the less common possibility will *never* be detected.

In the Hunter Group's research, which is based on one data qbit and two program qbits, equation 3.5 is valid in the range $(0.2 \leq \eta_1 \leq 0.8)$. When the qbits are replicated (as we discuss in Chapter 4) the range varies. We use these rules extensively in computing our own results.

3.3 Known overlap between the program qbits

3.3.1 Finding the density matrices

Here we calculate density matrices $\overline{\hat{\rho}}_1$ and $\overline{\hat{\rho}}_2$, analogous to those used by the Hunter Group, but modified to account for the known overlap between the program qbits. We address this problem in two stages: firstly we solve a simple case where the position of one of the program qbits is fixed, and secondly, we extend this solution to cover the general situation.

To find $\overline{\hat{\rho}}_1$, we provisionally place one of the program qbits at the north pole of the Bloch sphere

$$(|\psi_1\rangle = |0\rangle).$$

The other program qbit can be now written in the general form

$$|\psi_3\rangle = \cos(\beta/2)|0\rangle + e^{i\alpha} \sin(\beta/2)|1\rangle. \tag{3.6}$$

The overlap is given by:

$$\begin{aligned} \langle \psi_1 | \psi_3 \rangle &= (\cos(\beta/2)\langle 0|0\rangle + e^{i\alpha} \sin(\beta/2)\langle 0|1\rangle) \\ &= \cos(\beta/2). \end{aligned} \tag{3.7}$$
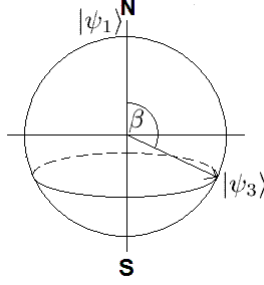
Figure 3.2: Bloch sphere with locus of constant overlap with $|0\rangle$

The overlap is independent of α, as shown in Figure 3.2. For all values of β except 0 and π, $|\psi_3\rangle$ lies on a circle of latitude given by $(0 \leq \alpha < 2\pi)$.

We first find the provisional density matrix $\bar{\hat{\rho}}_{1(prov)}$. We use

$$
\begin{aligned}
|\psi_1\rangle &= |0\rangle \\
|\psi_3\rangle &= \cos(\beta/2)|0\rangle + e^{i\alpha}\sin(\beta/2)|1\rangle.
\end{aligned}
\tag{3.8}
$$

Assuming that $|\psi_2\rangle$, the data qbit, is the same as $|\psi_1\rangle$, we get the product ket

$$
\begin{aligned}
|\Psi_{1(prov)}\rangle &= |\psi_1\rangle|\psi_1\rangle|\psi_3\rangle \\
&= (\cos(\beta/2)|000\rangle + e^{i\alpha}\sin(\beta/2)|001\rangle).
\end{aligned}
\tag{3.9}
$$

The provisional density matrix $\bar{\hat{\rho}}_{1(prov)}$ is the value of $|\Psi_{1(prov)}\rangle\langle\Psi_{1(prov)}|$ averaged over the circle of constant overlap. This requires us to average over all values of the azimuthal angle α. Before averaging,

$$
\hat{\rho}_{1(prov)} = (\cos(\beta/2)|000\rangle + e^{i\alpha}\sin(\beta/2)|001\rangle)(\cos(\beta/2)\langle000| + e^{-i\alpha}\sin(\beta/2)\langle001|).
\tag{3.10}
$$

But $\int_0^{2\pi} d\alpha\, e^{i\alpha} = 0$. When the provisional density matrix is averaged, the off-diagonal terms vanish, leaving

$$
\bar{\rho}_{1(prov)} = \cos^2(\beta/2)|000\rangle\langle000| + \sin^2(\beta/2)|001\rangle\langle001|.
\tag{3.11}
$$

To find $\bar{\hat{\rho}}_{2(prov)}$, where $|\psi_2\rangle = |\psi_3\rangle$, we use a similar argument, but this time placing $|\psi_3\rangle$ at $|0\rangle$, instead of $|\psi_1\rangle$. We obtain

$$
\bar{\rho}_{2(prov)} = \cos^2(\beta/2)|000\rangle\langle000| + \sin^2(\beta/2)|100\rangle\langle100|.
\tag{3.12}
$$

Next we generalise these results by removing the restriction that one of the program qbits must be at the north pole. To make the appropriate transformations we replace $|0\rangle$ by the general form

$$
|0'\rangle = \cos(\theta/2)|0\rangle + e^{i\psi}\sin(\theta/2)|1\rangle
\tag{3.13}
$$

and $|1\rangle$ by the state orthogonal to $|0'\rangle$, namely,

$$|1'\rangle = \sin(\theta/2)|0\rangle - e^{i\psi}\cos(\theta/2)|1\rangle. \tag{3.14}$$

The generalised forms of $|\Psi_1\rangle$ and $|\Psi_1\rangle$ are now given by

$$\begin{aligned}
|\Psi_1\rangle &= \cos(\beta/2)|0'0'0'\rangle + e^{i\alpha}\sin(\beta/2)|0'0'1'\rangle \tag{3.15}\\
|\Psi_2\rangle &= \cos(\beta/2)|0'0'0'\rangle + e^{i\alpha}\sin(\beta/2)|1'0'0'\rangle.
\end{aligned}$$

We can write $|\Psi_1\rangle$ as the sum of two column vectors:

$$|\Psi_1\rangle = |p\rangle + |q_1\rangle \tag{3.16}$$

where $|p\rangle$ is the expansion, in terms of the computational basis states $|0\rangle$ and $|1\rangle$, of the term $\cos(\beta/2)|0'0'0'\rangle$, and $|q_1\rangle$ is the expansion of $e^{i\alpha}\sin(\beta/2)|0'0'1'\rangle$. Similarly,

$$|\Psi_2\rangle = |p\rangle + |q_2\rangle \tag{3.17}$$

where $|q_2\rangle$ is the expansion of the term $e^{i\alpha}\sin(\beta/2)|1'0'0'\rangle$. The full forms of these vectors are:

$$|p\rangle = \cos(\beta/2)\begin{pmatrix}
\cos^3(\theta/2) & |000\rangle \\
+e^{i\psi}\sin(\theta/2)\cos^2(\theta/2) & |001\rangle \\
+e^{i\psi}\sin(\theta/2)\cos^2(\theta/2) & |010\rangle \\
+e^{i\psi}\sin(\theta/2)\cos^2(\theta/2) & |100\rangle \\
+e^{2i\psi}\sin^2(\theta/2)\cos(\theta/2) & |011\rangle \\
+e^{2i\psi}\sin^2(\theta/2)\cos(\theta/2) & |101\rangle \\
+e^{2i\psi}\sin^2(\theta/2)\cos(\theta/2) & |110\rangle \\
+e^{3i\psi}\sin^3(\theta/2) & |111\rangle
\end{pmatrix} \tag{3.18}$$

$$|q_1\rangle = \sin(\beta/2)\begin{pmatrix}
\cos^2(\theta/2)\sin(\theta/2) & |000\rangle \\
-e^{i\psi}\cos^3(\theta/2) & |001\rangle \\
+e^{i\psi}\cos(\theta/2)\sin^2(\theta/2) & |010\rangle \\
+e^{i\psi}\cos(\theta/2)\sin^2(\theta/2) & |100\rangle \\
-e^{2i\psi}\cos^2(\theta/2)\sin(\theta/2) & |011\rangle \\
-e^{2i\psi}\cos^2(\theta/2)\sin(\theta/2) & |101\rangle \\
+e^{2i\psi}\sin^3(\theta/2) & |110\rangle \\
-e^{3i\psi}\cos(\theta/2)\sin^2(\theta/2) & |111\rangle
\end{pmatrix} \tag{3.19}$$

$$|q_2\rangle = \sin(\beta/2)\begin{pmatrix}
\cos^2(\theta/2)\sin(\theta/2) & |000\rangle \\
+e^{i\psi}\cos(\theta/2)\sin^2(\theta/2) & |001\rangle \\
+e^{i\psi}\cos(\theta/2)\sin^2(\theta/2) & |010\rangle \\
-e^{i\psi}\cos^3(\theta/2) & |100\rangle \\
+e^{2i\psi}\sin^3(\theta/2) & |011\rangle \\
-e^{2i\psi}\cos^2(\theta/2)\sin(\theta/2) & |101\rangle \\
-e^{2i\psi}\cos^2(\theta/2)\sin(\theta/2) & |110\rangle \\
-e^{3i\psi}\cos(\theta/2)\sin^2(\theta/2) & |111\rangle
\end{pmatrix} \tag{3.20}$$

Note that the computational states are not listed in the order of the corresponding binary numbers, but are grouped according to the number of $|1\rangle$s they contain; this will generate density matrices with compact sub-matrices on the main diagonal and zeros elsewhere.

Each term in this expansion is the product of the appropriate elements of $|0\rangle$ and $|1\rangle$. For example, the coefficient of $|101\rangle$ in the state vector $|q_1\rangle$ is the product of: $e^{i\psi}\sin(\theta/2)|1\rangle$, $\cos(\theta/2)|0\rangle$, and $-e^{i\psi}\cos(\theta/2)|1\rangle$.

Let $\mathbf{A}$ denote the outer product $|p\rangle\langle p|$ and $\mathbf{B}$ denote $|q_1\rangle\langle q_1|$. The density matrix $\bar{\bar{\rho}}_1$ is the average, over the Bloch sphere, of the weighted sum of two matrices $\mathbf{A}$ and $\mathbf{B}$:

$$\bar{\bar{\rho}}_1 = \frac{1}{4\pi}\left(\int_S \mathbf{A}\,dS\right) + \int_S \mathbf{B}\,dS. \tag{3.21}$$

The integral $\int_S dS$ denotes integration over the surface of the Bloch sphere.

Similarly, using the definition of $\bar{\bar{\rho}}_2$ we find

$$\bar{\bar{\rho}}_2 = \frac{1}{4\pi}\left(\int_S \mathbf{A}\,dS\right) + \int_S \mathbf{C}\,dS \tag{3.22}$$

where $\mathbf{C}$ is the outer product $|q_2\rangle\langle q_2|$.

The expanded forms of these integrals are.

$$X = \frac{1}{4\pi}\int_S A\,dS = \frac{1}{12}\begin{pmatrix} 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix} \tag{3.23}$$

$$Y = \frac{1}{4\pi}\int_S B\,dS = \frac{1}{12}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.24}$$

46

$$Z = \frac{1}{4\pi} \int_S C dS = \frac{1}{12} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.25}$$

Finally, for any value of the overlap β, we have:

$$\bar{\rho}_1 = X \cos^2(\beta) + Y \sin^2(\beta) \tag{3.26}$$
$$\bar{\rho}_2 = X \cos^2(\beta) + Z \sin^2(\beta).$$

3.3.2 Optimum discrimination with a known overlap

Having obtained the density matrices, we may now use Helstrom's rule to find optimum discrimination rates. The rule is summarised in equations 3.4 and 3.5.

Figure 3.3 presents a set of results for the error rate expected from Triads with various degrees of overlap between the program qbits and a range of values for the *a priori* probabilities. The left-hand border describes the system when the two program qbits are orthogonal (zero overlap) and shows that the worst discrimination, for all overlaps, occurs when $\eta_1 = 1/2$. For values of overlap close to 1, a nearly optimal strategy is to choose the most likely value of the data qbit, depending on η_1. The error probability for $\eta_1 = 1/2$ and zero overlap $(\beta = \pi)$ coincides with the analytic solution presented at the end of the last section. Figure 3.4 gives a different view of some of the same data. With each value of overlap, the probability of error increases almost linearly to a maximum when the *a priori* expectations of the two states are equal.

3.3.3 Unambiguous discrimination with a known overlap

Finding the failure rate for unambiguous discrimination uses the same analysis as was presented in [40] and summarised in the previous chapter. The only difference is that whereas the Hunter Group found an average overlap of the program qbits of $\sqrt{\frac{1}{2}}$ we can use specific values. Figure 3.5 shows how the success rate depends on η_1 for various fixed overlaps, using the formula

$$P = \frac{2}{3}(1 - \sqrt{(\eta_1 \eta_2)})(1 - \overline{|\langle \psi_1 | \psi_3 \rangle|^2}). \tag{3.27}$$
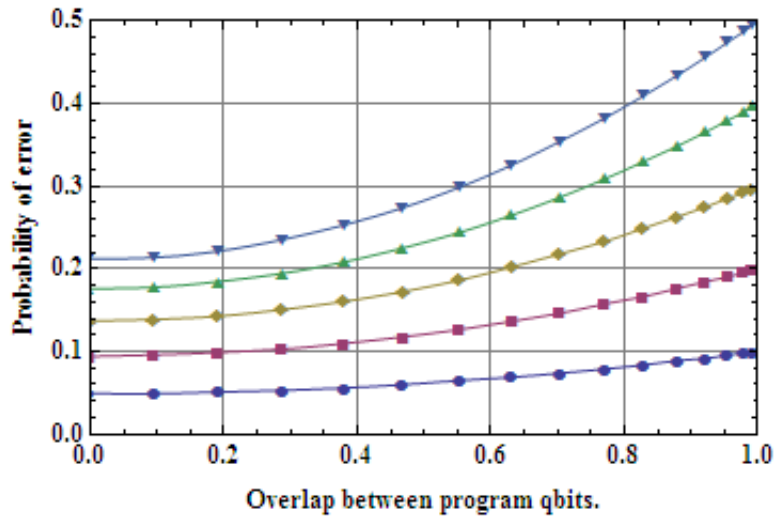
Figure 3.3: Probable errors for known overlap between program qbits. $\eta_1{=}0.1$ (lower curve) to $\eta_1{=}0.5$ (upper curve) in steps of 0.1
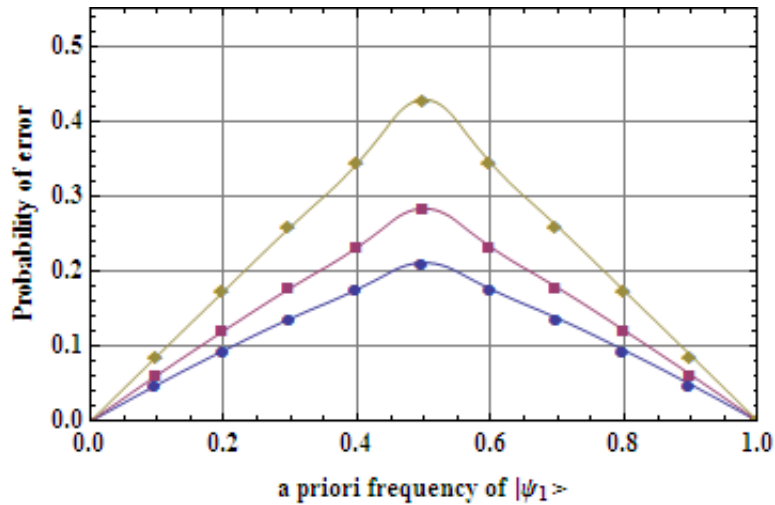

Figure 3.4: Probability of error against *a priori* frequency for overlap $= 0$ (upper curve), 0.4 (centre curve), and 0.8 (lower curve)
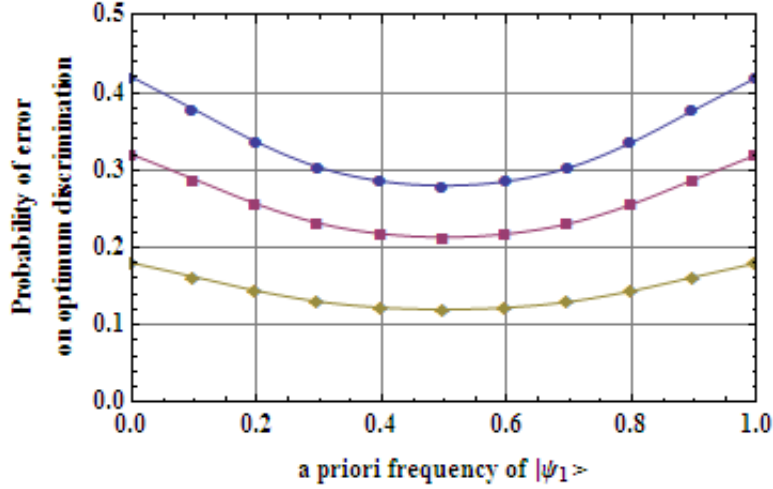
Figure 3.5: Probability of success in unambiguous discrimination against *a priori* frequency for overlap $= 0$ (upper curve), 0.4 (centre curve), and 0.8 (lower curve)

3.4 Both program qbits on a known great circle

3.4.1 Finding the density matrices

In the great circle case, we have *a priori* knowledge that both the program qbits lie on a *given* great circle of the Bloch sphere. This condition is important, as any two randomly selected different points on the Bloch sphere define *some* great circle, but we can exploit knowledge of the specific great circle in the choice of our measurement.

Another point to note is that any given great circle can be transformed to coincide with the equator, or with what we might call the extended Greenwich Meridian, just by rotating it about the sphere. It follows that the error rate for *all* given great circles will be the same.

Taking the great circle to be the extended Greenwich Meridian for which the azimuthal angle is either 0 or π, we can incorporate both possibilities by writing

$$
\begin{aligned}
|\psi_1\rangle &= \cos(\alpha/2)|0\rangle + \sin(\alpha/2)|1\rangle \\
|\psi_3\rangle &= \cos(\gamma/2)|0\rangle + \sin(\gamma/2)|1\rangle
\end{aligned}
\tag{3.28}
$$

and allowing α and γ to take values in the range 0 to 2π. Following the now accustomed pattern, we write

$$
\begin{aligned}
|\Psi_1\rangle &= |\psi_1\rangle|\psi_1\rangle|\psi_3\rangle \\
\hat{\rho}_1 &= |\psi_1\rangle|\psi_1\rangle\langle\psi_1|\langle\psi_1| \otimes |\psi_3\rangle\langle\psi_3| \\
&= \hat{\lambda}_1 \otimes \hat{\mu}_1.
\end{aligned}
\tag{3.29}
$$

Both $\hat{\lambda}_1$ and $\hat{\mu}_1$ must be averaged on our great circle (rather than the surface of the full Bloch sphere). Averaging $\hat{\lambda}_1$ over a circle defined by $(0 \leq \alpha \leq 2\pi)$ gives the density matrix:

$$\overline{\hat{\lambda}}_1 = \frac{1}{8} \begin{pmatrix} 3 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 3 \end{pmatrix} \tag{3.30}$$

Similarly, averaging $\hat{\mu}_1$ over a circle defined by $(0 \leq \gamma \leq 2\pi)$ leads to:

$$\overline{\hat{\mu}}_1 = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.31}$$

It then follows that

$$\overline{\hat{\rho}}_1 = \overline{\hat{\lambda}} \otimes \overline{\hat{\mu}}. \tag{3.32}$$

The density operator $\overline{\hat{\rho}}_2$ is worked out by a similar method:

$$\overline{\hat{\rho}}_2 = \overline{\hat{\mu}} \otimes \overline{\hat{\lambda}}. \tag{3.33}$$

The explicit forms of these two density operators are:

$$\overline{\hat{\rho}}_1 = \overline{\hat{\lambda}}_1 \otimes \overline{\hat{\mu}}_1 = \frac{1}{16} \begin{pmatrix} 3 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}. \tag{3.34}$$

$$\overline{\hat{\rho}}_2 = \overline{\hat{\mu}}_1 \otimes \overline{\hat{\lambda}}_1 = \frac{1}{16} \begin{pmatrix} 3 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 3 \end{pmatrix}. \tag{3.35}$$

3.4.2   Minimum error for a triad confined to a known great circle

The minimum-error strategy for distinguishing between these states requires us to find the eigenvalues of $\hat{\Lambda}$, as defined in equation 1.32. For equal prior probabilities $(\eta_1 = 1/2)$ we find the minimum error probability P.E. ≈ 0.323 .

3.4.3 Unambiguous discrimination for a triad confined to a known great circle

When the two program qbits are confined to a great circle, their average overlap is $\sqrt{\frac{1}{2}}$ We conclude that the probability of failure Q_F is

$$Q_F = \frac{2 + \sqrt{\eta_1 \eta_2}}{3}. \tag{3.36}$$

as in the general case described by the Hunter Group.

For unambiguous discrimination, knowledge that both program qbits are on a known great circle makes no difference to the probability of failure, as the overlap of the two program qbits is the same as when their positions on the Bloch sphere are unrestricted.

3.5 Program qbits confined to polar caps

The polar cap configuration posits a Triad in which one program qbit is located in a cap centred on the north pole of the Bloch sphere, and the other in a cap centred on the south pole. Both caps are the same size, bounded by lines of latitude defined by the angle θ, as shown in figure 3.1c.

3.5.1 Finding the density matrices

The northern program qbit may be written as

$$|\psi_N\rangle = \cos(\beta/2)|0\rangle + e^{i\alpha}\sin(\beta/2)|1\rangle \tag{3.37}$$

where $(0 \leq \beta \leq \theta)$. Similarly, the southern program qbit is

$$|\psi_S\rangle = \cos(\delta/2)|0\rangle + e^{i\gamma}\sin(\delta/2)|1\rangle \tag{3.38}$$

where $(\pi - \theta \leq \delta \leq \pi)$. The two possible states of the discriminator are:

$$
\begin{aligned}
|\Psi_1\rangle &= |\psi_N\rangle|\psi_N\rangle|\psi_S\rangle \\
|\Psi_2\rangle &= |\psi_N\rangle|\psi_S\rangle|\psi_S\rangle.
\end{aligned}
\tag{3.39}
$$

The density matrix for the first of these two states is obtained as before, by averaging the states over the relevant cap. If we define

$$
\begin{aligned}
\hat{\mu} &= |\psi_N\rangle|\psi_N\rangle\langle\psi_N|\langle\psi_N| \\
\hat{\lambda} &= |\psi_S\rangle\langle\psi_S|
\end{aligned}
\tag{3.40}
$$

then averaging gives

$$
\begin{aligned}
\bar{\hat{\mu}}_1 &= \frac{1}{\kappa}\int_0^{2\pi} d\psi \int_0^{\theta} d\beta \sin\beta\,\hat{\mu} \\
&= \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & b & 0 \\ 0 & b & b & 0 \\ 0 & 0 & 0 & c \end{pmatrix}
\end{aligned}
\tag{3.41}
$$

where $\kappa = 2\pi(1 - \cos\theta)$ is the surface area of the cap and

$$
\begin{aligned}
a &= \frac{1}{\kappa} \int_0^{2\pi} d\alpha \int_0^\theta d\beta \cos^4(\beta/2) \sin(\beta) && (3.42)\\
&= \frac{2}{3} \left(\frac{1 - \cos^6(\theta/2)}{1 - \cos\theta} \right)\\
b &= \frac{1}{\kappa} \int_0^{2\pi} d\alpha \int_0^\theta d\beta \cos^2(\beta/2) \sin^2(\beta/2) \sin(\beta)\\
&= 4 \left(\frac{\frac{1}{4}\sin^4(\theta/2) - \frac{1}{6}\sin^6(\theta/2)}{1 - \cos\theta} \right)\\
c &= \frac{1}{\kappa} \int_0^{2\pi} d\alpha \int_0^\theta d\beta \sin^4(\beta/2) \sin(\beta)\\
&= \frac{2}{3} \left(\frac{\sin^6(\theta/2)}{1 - \cos\theta} \right).
\end{aligned}
$$

$$(3.43)$$

Similarly, averaging $\hat{\lambda}$ over the southern cap gives

$$
\begin{aligned}
\overline{\hat{\lambda}}_1 &= \frac{1}{\kappa} \int_0^{2\pi} d\psi \int_{\pi-\theta}^\pi d\delta \sin\delta\, \hat{\lambda} && (3.44)\\
&= \begin{pmatrix} d & 0 \\ 0 & e \end{pmatrix} && (3.45)
\end{aligned}
$$

where

$$
\begin{aligned}
d &= \frac{1}{\kappa} \int_0^{2\pi} d\gamma \int_{\pi-\theta}^\pi d\delta \cos^2(\delta/2) \sin(\delta) && (3.46)\\
&= \left(\frac{\sin^4(\theta/2)}{1 - \cos\theta} \right)\\
e &= \frac{1}{\kappa} \int_0^{2\pi} d\gamma \int_{\pi-\theta}^\pi d\delta \sin^2(\delta/2) \sin(\delta) && (3.47)\\
&= \left(\frac{1 - \cos^4(\theta/2)}{1 - \cos\theta} \right).
\end{aligned}
$$

Tensor multiplication of these two matrices gives the density operator:

$$
\overline{\rho}_1 = \overline{\hat{\mu}} \otimes \overline{\hat{\lambda}}. \tag{3.48}
$$

The second density operator is obtained in the same manner:

$$
\overline{\rho}_2 = \overline{\hat{\lambda}} \otimes \overline{\hat{\mu}}. \tag{3.49}
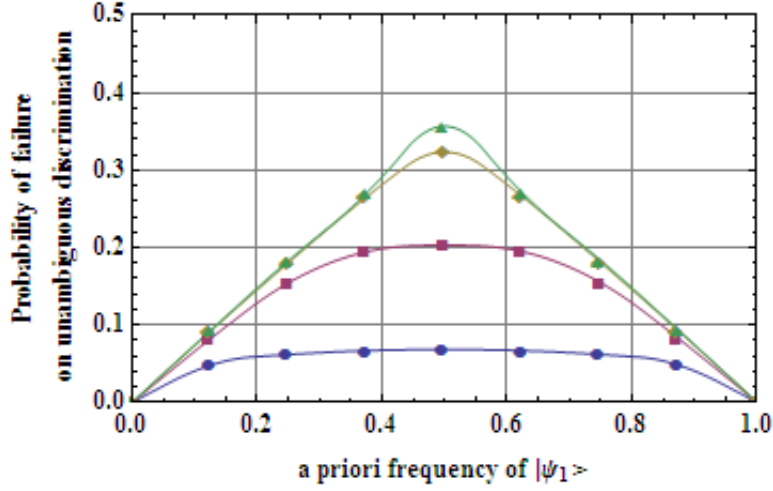$$

Figure 3.6: Probable error for $\theta = \pi/4$ (bottom curve) to $\theta = \pi/$ (top curve) in steps of $\pi/4$

Both matrices, expressed in terms of the variables a to e, are:

$$\bar{\rho}_1 = \bar{\bar{\mu}} \otimes \bar{\bar{\lambda}} = \begin{pmatrix} ad & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & ae & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & bd & 0 & bd & 0 & 0 & 0 \\ 0 & 0 & 0 & be & 0 & be & 0 & 0 \\ 0 & 0 & bd & 0 & bd & 0 & 0 & 0 \\ 0 & 0 & 0 & be & 0 & be & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & cd & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & ce \end{pmatrix}. \tag{3.50}$$

$$\bar{\rho}_2 = \bar{\bar{\lambda}} \otimes \bar{\bar{\mu}} = \begin{pmatrix} ce & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & be & be & 0 & 0 & 0 & 0 & 0 \\ 0 & be & be & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & ae & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & cd & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & bd & bd & 0 \\ 0 & 0 & 0 & 0 & 0 & bd & bd & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & ad \end{pmatrix}. \tag{3.51}$$

Figure 3.6 shows how the probable error varies with η_1, for a selection of values of θ. It is worth noting, as a check, that when $\theta = \pi$, and the positions of the program qbits are unrestricted, the results are identical with those of the Hunter Group for a similar configuration.

3.5.2 Unambiguous discrimination for a triad in the polar cap case

As before, the success rate of unambiguous discrimination is determined only by the ratio of η_1 and η_2, and by the overlap of the two program qbits. In the polar cap case the overlap changes
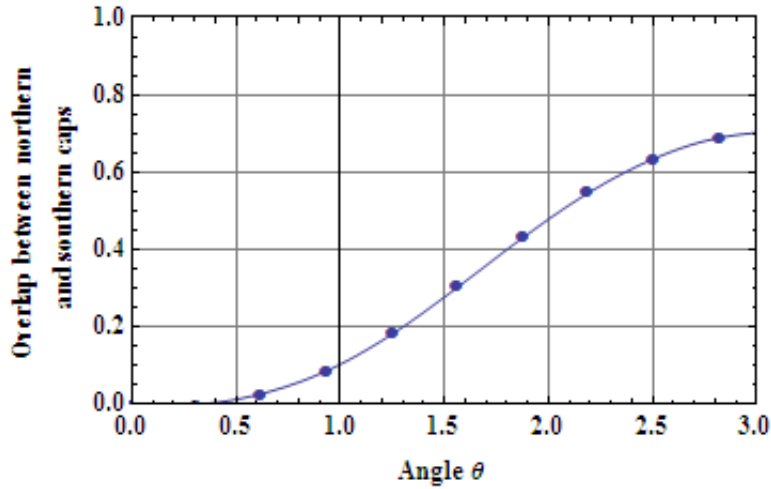
Figure 3.7: Overlap between polar caps for values of θ

with the angle θ in a non-linear manner, shown in figure 3.7. The corresponding success rates S are displayed in figure 3.8, using the formula

$$S = 1 - \frac{2}{3}(1 - \sqrt{(\eta_1 \eta_2)})(1 - \overline{|\langle \psi_1 | \psi_3 \rangle|^2}). \qquad (3.52)$$

3.6   Conclusions

We have investigated the error rates in discriminating the two possible states of a Triad, when certain types of additional information are available. In practice, this could often be the case. For example, if the two program qbits are coded as photons with different linear polarisations, they will lie on a great circle of the Bloch sphere.

We can report a number of findings, first for minimum error measurements:

- When the two program qbits of a Triad are orthogonal, and the two possible states are equally likely, the probability of error is less than if they were selected randomly (0.211 as opposed to 0.356). The minimum error rate rises monotonically as the overlap between the qbits increases.

- If both program qbits are on the same (given) great circle, the expected minimum error rate is 0.323.

- If the program qbits are confined to 'caps' centred in antipodal points on the Bloch sphere, the error rate grows as the areas increase in size. In particular, if each area covers just half the Bloch sphere, and $\eta_1 = \eta_2 = \frac{1}{2}$, the error rate in recognition is 0.203.
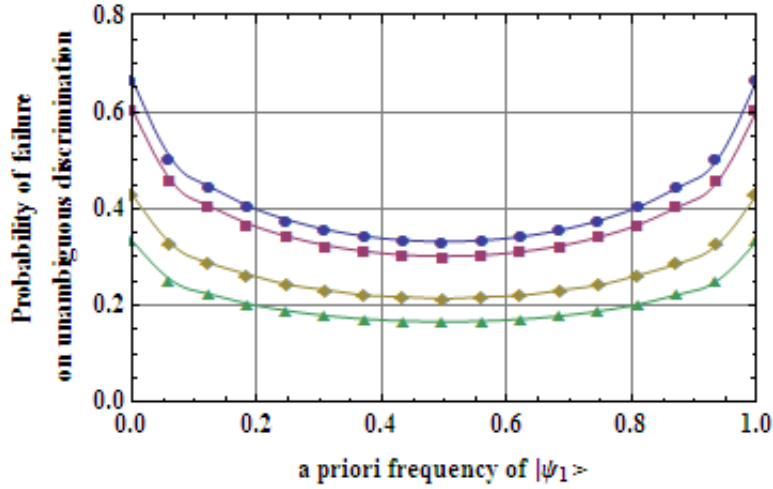
54

Figure 3.8: Probable error for $\theta = \pi/4$ (top curve) to $\theta = \pi$ (bottom curve) in steps of $\pi/4$

Our results on minimum error rates broadly confirm our expectations - that additional classical information about a Triad can improve the successful discrimination rate. An interesting result is that the error rate for the case when the two program qbits are in the northern and southern hemispheres of the Bloch sphere is actually slightly *lower* than the rate for orthogonal program qbits. It seems that absolute positional information carries more weight than relational data.

Second, for unambiguous measurements, the failure rate is disappointingly large. However, the rate is improved when the data qbits are known to have a small overlap, or are confined to small areas about antipodal poles. The knowledge that the program qbits are confined to a given great circle makes no difference to the error rate.

Chapter 4

Programmed discrimination of multiple sets of qbits with added classical information

"Controlling complexity is the essence of computer programming."
Brian Kernighan, Software Tools (1976)

4.1 Introduction

In this chapter we present more new results in the area of programmed quantum discrimination, which attempts to determine whether a data qbit is the same as the first or second of a given pair of program qbits. In Chapter 3 we examined the effect of additional classical information on this process. We now extend the investigation to consider the effects of replicating some or all of the qbits in the group. As the parameters to the calculations can take many values in several dimensions, we offer only a sample of results.

Turning to a possible application of the method, we investigate the efficacy of programmed discrimination in the context of data communication. The technique offers a key advantage over other methods of quantum-based data transmission; namely, that it is insensitive to any unitary transformation that may occur in transit, provided only that the same transformation applies equally to all the qbits in the group. However, the technique is costly in its use of resources. Using the best configuration we could find, with unambiguous discrimination, orthogonal program qbits, and a duplicated data qbit, the transmission of one binary digit reliably still needs 8 qbits.

A key aspect of any scientific paper is that the results should be independently verifiable and repeatable. To this end we describe our algorithms in some detail.

Chapter 4 is closely based on our paper "Programmed discrimination of multiple sets of qbits with added classical information", published in 2012 [48].

4.2 The problem

Our basic aim is to extend the analysis described in Chapter 3 to configurations where each of the qbits in the triad may be replicated two or more times. Chapter 2 quotes some analytic

solutions to this problem, but they only apply to symmetrical configurations where the program qbits are uncorrelated.

It is convenient to refer to a configuration which has x copies of one program qbit, y copies of the data qbit, and z copies of the other program qbit as $\{x, y, z\}$ (for example, $\{1,3,2\}$).

In the existing literature the state of a system might have been described as:

$$|\Psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes |\psi_3\rangle \tag{4.1}$$

From this point we drop the position indicators $_1$, $_2$ and $_3$ because these qbits may be repeated. Instead, we rely on the ordering of the qbits (or groups of qbits), which is always *first group of program qbits - group of data qbits - second group of program qbits*. . A useful notation for groups of identical qbits is $|\psi\rangle^{\otimes n}$, which signifies the n-fold tensor product $\frac{|\psi\rangle \otimes |\psi\rangle \otimes \cdots \otimes |\psi\rangle}{\longleftarrow n \longrightarrow}$.

The problem can be solved in a parameter space that spans many dimensions. Two primary dimensions are:

- a) The type of **relationship** between the program qbits (fixed overlap, same given great circle, and confined to caps)

- b) The type of **discrimination** (optimal or unambiguous)

Continuous dimensions include:

- c) The *a priori* probabilities of identity η_1 and η_2

- d) The degree of overlap (for a known overlap)

- e) The size of the caps (for confined qbits)

Within all these dimensions we also consider different patterns of replication.

The parameter space of the problem is massive; it would be impractical to explore it in full and present the results in this thesis. Instead, we display just a few sample results, and give the URL [1]of a website that holds our programs [?]. Readers can use this software to explore other parts of the space in which they have an interest.

4.3 Methods

This section presents describes the methods used to establish our results. They are derived from those described in Chapter 3, but there is a major difference. In chapter 3 we were dealing with systems consisting of just three qbits, and the analysis could largely be done by hand. Here we are concerned with qbits in far greater numbers. The analysis of any system with a total of n qbits involves matrices of dimension $2^n \times 2^n$. Increasing values of n results in an exponential growth in the size of the problem. For $n > 3$ it is no longer feasible to carry out the analysis by
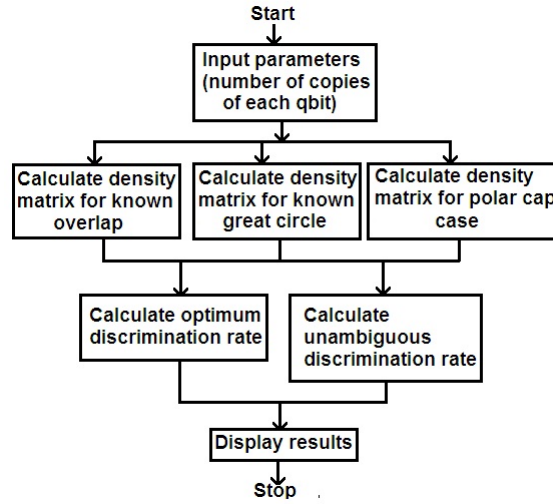
Figure 4.1: Program structure

hand except in special cases, so we rely on a computer to help solve the problem.

Figure 4.1 shows the structure of the programs we used for the problem. The computation of the density matrices depends on the type of constraint on the program qbits, and not at all on the mode of discrimination. Likewise, the discrimination process - whether optimal or unambiguous - is independent of the qbit constraint.

Figure 4.1 indicates that there are six routes through the programs, linking each type of constraint with each mode of discrimination. In practice each route has been made into a separate module, but this is only for convenience. The modules could easily have been combined, with a single extra parameter to indicate the routing to be followed.

Appendix 1 gives the "ReadMe" file [51] on the web site together with the modules.

4.4 Calculation of density matrices

The data qbit is always the same as one or other of the program qbits. It follows that for any configuration $\{x, y, z\}$, there are only two possible states of the system:

$$|\Psi_1\rangle = |\psi_1\rangle^{\otimes(x+y)}|\psi_3\rangle^{\otimes z} \tag{4.2}$$

$$|\Psi_2\rangle = |\psi_1\rangle^{\otimes x}|\psi_3\rangle^{\otimes(y+z)} \tag{4.3}$$

We need density matrices for $|\Psi_1\rangle$ and $|\Psi_2\rangle$ for each of the three types of constraint. These are illustrated in figure 4.2 (in a diagram similar to one in Chapter 3)

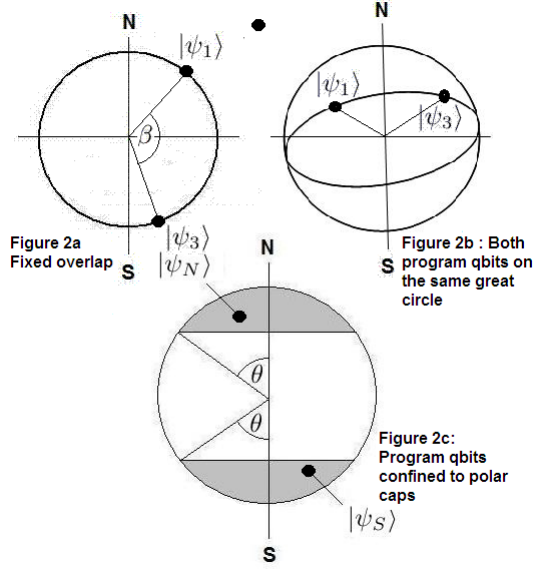[1] These programs are included in the CD-ROM which forms part of this thesis.

Figure 4.2: Three types of classical knowledge

4.4.1 Density matrices for fixed overlap

This case is illustrated in figure 4.2a. Following the approach used in Chapter 3, we provisionally place one of the program qbit groups at the north pole of the Bloch sphere.

$$|\psi_{1(prov)}\rangle^{\otimes(x+y)} = |0\rangle^{\otimes(x+y)}$$

The other provisional program qbit group can be written as:

$$|\psi_{3(prov)}\rangle^{\otimes z} = (\cos(\beta/2)|0\rangle + e^{i\alpha}\sin(\beta/2)|1\rangle)^{\otimes z}$$

where the known overlap is $\cos^2(\beta/2)$. This leads to a provisional density matrix:

$$\begin{aligned}\overline{\rho}_{1(prov)} &= \cos^2(\beta/2)|0\rangle^{\otimes(x+y+z)}\langle 0|^{\otimes(x+y+z)} \\ &+ \sin^2(\beta/2)|0\rangle^{\otimes(x+y)}|1\rangle^{\otimes z}\langle 0|^{\otimes(x+y)}|\langle 1|^{\otimes z}\end{aligned} \tag{4.4}$$

where x, y, and z form the pattern of replication.

The provisional density matrix for $|\Psi_{2(prov)}\rangle$ is found by placing the *other* program qbit group at the north pole. It is:

$$\begin{aligned}\overline{\rho}_{2(prov)} &= \cos^2(\beta/2)|0\rangle^{\otimes(x+y+z)}\langle 0|^{\otimes(x+y+z)} \\ &+ \sin^2(\beta/2)|1\rangle^{\otimes x}|0\rangle^{\otimes(y+z)}\langle 1|^{\otimes x}\langle 0|^{\otimes(y+z)}\end{aligned} \tag{4.5}$$

Next we apply a generalising transformation that lets both qbit groups be anywhere on the Bloch sphere. We replace $|0\rangle$ by the general form:

$$|0'\rangle = \cos(\theta/2)|0\rangle + e^{i\psi}\sin(\theta/2)|1\rangle \tag{4.6}$$

and $|1\rangle$ by the state orthogonal to $|0'\rangle$, namely,

$$|1'\rangle = \sin(\theta/2)|0\rangle - e^{i\psi}\cos(\theta/2)|1\rangle \tag{4.7}$$

The generalised forms of $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are now given by

$$
\begin{aligned}
|\Psi_1\rangle &= \cos(\beta/2)|0'0'0'\rangle^{\otimes(x+y+z)} \\
&+ e^{i\alpha}\sin(\beta/2)|0'0'\rangle^{\otimes(x+y)}|1'\rangle^{\otimes z}
\end{aligned} \tag{4.8}
$$

$$
\begin{aligned}
|\Psi_2\rangle &= \cos(\beta/2)|0'0'0'\rangle^{\otimes(x+y+z)} \\
&+ e^{i\alpha}\sin(\beta/2)|1'\rangle^{\otimes x}|0'0'\rangle^{\otimes(y+z)}.
\end{aligned} \tag{4.9}
$$

We can write $|\Psi_1\rangle$ as the weighted sum of two column vectors:

$$|\Psi_1\rangle = \cos(\beta/2)|p\rangle + e^{i\alpha}\sin(\beta/2)|q\rangle \tag{4.10}$$

where $|p\rangle$ is the expansion of the term $|0'0'0'\rangle^{\otimes(x+y+z)}$, in terms of the computational basis states $|0\rangle$ and $|1\rangle$, and $|q\rangle$ is the expansion of $\left(|0'0'\rangle^{\otimes(x+y)}|1'\rangle^{\otimes z}\right)$. Similarly,

$$|\Psi_2\rangle = \cos(\beta/2)|p\rangle + e^{i\alpha}\sin(\beta/2)|r\rangle \tag{4.11}$$

where $|p\rangle$ is the same as before and $|r\rangle$ is the expansion of the term $\left(|1'\rangle^{\otimes x}|0'0'\rangle^{\otimes(y+z)}\right)$.

Let $\mathbf{A}$ denote the outer product $|p\rangle\langle p|$ and $\mathbf{B}$ denote $|q\rangle\langle q|$. The density matrix $\bar{\rho}_1$ is the average, over the Bloch sphere, of the weighted sum of two matrices $\mathbf{A}$ and $\mathbf{B}$:

$$\bar{\rho}_1 = \frac{1}{4\pi}\left(\cos^2(\beta/2)\int_S \mathbf{A}\,dS + \sin^2(\beta/2)\int_S \mathbf{B}\,dS\right). \tag{4.12}$$

The integral $\int_S dS$ denotes integration over the surface of the Bloch sphere. Similarly, using the definintion of $\bar{\bar{\rho}}_2$ we find

$$\bar{\rho}_2 = \frac{1}{4\pi}\left(\cos^2(\beta/2)\int_S \mathbf{A}\,dS + \sin^2(\beta/2)\int_S \mathbf{C}\,dS\right) \tag{4.13}$$

where $\mathbf{C}$ is the outer product $|r\rangle\langle r|$.

We now turn to the calculation of the individual elements of the matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$. Let n be the total number of qbits in the configuration ($n = x + y + z$). Each of the matrices will be of order 2^n, and we can label the rows and columns by the binary numbers 0 to $(2^n - 1)$, so that - for example - $A[J,K]$ is the element of A at row J and column K. Each element of any of the three matrices is the average, taken over the surface of the Bloch sphere, of an expression that has two factors:

- a) a ket $|X_j\rangle$ that is the same for all the elements in column J

- b) a bra $\langle Y_k|$ that is the same for all the elements in row K.

We may put:

$$I[J,K] = \left(\frac{1}{4\pi}\right)\int_S |X_J\rangle\langle Y_K| \tag{4.14}$$

where the integration is over the whole sphere. **I** stands for **A**, **B** or **C** as appropriate. It is useful to code the current configuration as three binary numbers P, Q and R:

$$P = \underbrace{0000000000...0000000000}_{\longleftarrow n \longrightarrow} \tag{4.15}$$

$$Q = \underbrace{00000...00000}_{\longleftarrow x+y \longrightarrow} \underbrace{11...11}_{\longleftarrow z \longrightarrow} \tag{4.16}$$

$$R = \underbrace{111...111}_{\longleftarrow x \longrightarrow} \underbrace{00000...00000}_{\longleftarrow y+z \longrightarrow}. \tag{4.17}$$

The factor $|X_J\rangle$ for any of the matrices **A**, **B** or **C** can now be worked out by taking the appropriate number (P, Q or R) and matching it, bit by bit, against the binary version of the column label J. It is clear, from the definition of $|0'\rangle$ and $|1'\rangle$ that $|X_J\rangle$ will be the product of:

- a sign factor s_1, which may be $+1$ or -1

- a term $e^{t_1 i\phi}$ where t_1 is an integer such that $(t_1 \geq 0)$

- the cosine of the angle $(\theta/2)$, raised to the power u_1

- the sine of the angle $(\theta/2)$, raised to the power v_1. In all cases $(u_1 + v_1) = n$.

The generation of each term in $|X_J\rangle$ starts by initialising four integer variables which represent a partial product:

- $s_1 = +1$ (the sign of the term)

- $u_1 = 0$ (the power of the cosine term)

- $v_1 = 0$ (the power of the sine term)

- $t_1 = 0$ (the coeffiicent of ϕ).

The final product will be

$$|X_J\rangle = s_1 e^{t_1 i\phi} \cos^{u_1}(\theta/2) \sin^{v_1}(\theta/2) \tag{4.18}$$

Each step in generating this product consists of taking a digit from one of P, Q or R (say i) and the corresponding digit from the binary form of J, (say j) and applying the following rule. The order in which the bits are used is immaterial:

- $i = 0, j = 0$: Increment the cosine count u_1

- $i = 0, j = 1$: Increment the sine count v_1 and the exponent count t_1

- $i = 1, j = 0$: Increment the sine count v_1

- $i = 1, j = 1$: Increment the cosine count u_1, and the exponent count t_1. Change the sign of s_1.

The rule works because each of the four possible combinations of i and j identifies one of the terms in the expansion of $|0'\rangle$ and $|0'\rangle$ and multiplies it into the partial product.

The derivation of the factor $\langle Y_K|$ is the same, except that we use a row label K instead of a column label. As we are working on a ket, we decrement the exponent count instead of adding to it. We find:

$$\langle Y_K| = s_2 e^{t_2 i \phi} \cos^{u_2}(\theta/2) \sin^{v_2}(\theta/2) \tag{4.19}$$

and expect t_2 to be *negative*.

The expression to be integrated is the product of $|X_J\rangle$ and $\langle Y_K|$. As the integration is over the surface of a sphere we factor in $\sin(\theta)$, in the form $2\sin(\theta/2)\cos(\theta/2)$.

The average is:

$$\int_S dS = \frac{s}{4\pi} \left(\int_0^{2\pi} d\phi e^{ti\phi} \int_0^{\pi} d\theta \cos^{u+1}(\theta/2) \cos^{v+1}(\theta/2) \right) \tag{4.20}$$

where $s = s_1 \times s_2, t = t_1 + t_2, u = u_1 + u_2,$ and $v = v_1 + v_2$.

A useful feature of this expression is that it evaluates to zero for all values of t except 0. As the values of t_1 and t_2 for any column or row depend only on the number of ones in the binary expansions of J and K, it follows that the only non-zero terms in the matrix will be those where the intersecting row and column labels have the same number of ones. For these cells, the average can be worked out by a simple recursive method that uses a standard integral:

Define $I(m,n)$ as $\int_0^{\pi} \cos^n(\theta/2) \sin^m(\theta/2) d\theta$. Then

$$n = 1 \Rightarrow I(m,n) = \frac{1}{m+1} \tag{4.21}$$

otherwise,

$$I(m,n) = I(m, n-2) - I(m+2, n-2) \tag{4.22}$$

The C code that implements this module can be found at **fobm.c**.

4.4.2 Density matrices when both program qbits are on a given great circle

In this section, it is given *a priori* that the program qbits lie on a pre-defined great circle, as shown in Figure 4.2b. We note that any great circle on the Bloch sphere can be transformed into any other, just by rotating the sphere. The most convenient one to use is the 'Greenwich meridian', where the azimuth angle is always zero. We may put:

$$|\psi_1\rangle = \cos(\alpha/2)|0\rangle + \sin(\alpha/2)|1\rangle \tag{4.23}$$

$$|\psi_3\rangle = \cos(\gamma/2)|0\rangle + \sin(\gamma/2)|1\rangle \tag{4.24}$$

Following the established pattern,

$$|\Psi_1\rangle = |\psi_1\rangle^{\otimes(x+y)} |\psi_3\rangle^{\otimes z}$$

As $|\psi_1\rangle$ and $|\psi_3\rangle$ are independent, we can define the density matrix of the first state as:

$$\overline{\hat{\rho}_1} = \hat{\lambda}_1 \otimes \hat{\mu}_1$$

where

$$\hat{\lambda}_1 \quad = \quad |\psi_1\rangle^{\otimes(x+y)} \langle\psi_1|^{\otimes(x+y)} \tag{4.25}$$

$$\hat{\mu}_1 \quad = \quad |\psi_3\rangle^{\otimes z} \langle\psi_3|^{\otimes z}. \tag{4.26}$$

We compute $\hat{\lambda}_1$ and $\hat{\mu}_1$ separately, and then form their tensor product to find $\overline{\hat{\rho}_1}$.

The derivation of the individual terms in either of the matrices follows the general lines described in the previous section.

Each term has the form:

$$\left(\frac{1}{2\pi} \int_0^{2\pi} d\omega \, \cos^v(\omega/2) \sin^u(\omega/2) \right)$$

where ω stands for α or γ, and $(v+u)$ is twice the number of bits in each row or column label. The integration rule is:

- define $J(m,n)$ as $\int_0^{2\pi} d\omega \, \cos^n(\omega/2) \sin^m(\omega/2)$
  Note that $(m+n)$ cannot be odd.

- m and n both odd $\Rightarrow J(m,n) = 0$

- $n = 0 \Rightarrow J(m,n) = \prod_{h=1}^{m/2} \left(\frac{2h-1}{2h} \right)$

- otherwise $J(m,n) = J(m, n-2) - J(m+2, n-2)$.

The density matrix of the second possible state $\overline{\hat{\rho}_2}$ is found by multiplying $\hat{\lambda}$ and $\hat{\mu}$ in the opposite order:

$$\overline{\hat{\rho}_2} = \hat{\mu} \otimes \hat{\lambda}. \tag{4.27}$$

The code for this module is in **gcbm.c**.

4.4.3 Density matrices when the program qbits are confined to polar caps

Figure 4.2c shows the assumed positions of the polar caps that hold the program qbits. As these qbits are independent, we can use the approach set out in the previous section, and define each density matrix as the tensor product of two smaller matrices, entirely based on one or other of the polar caps.

We can again define the density matrix of the first state as:

$$\overline{\hat{\rho}_1} = \hat{\lambda}_1 \otimes \hat{\mu}_1$$

where

$$\hat{\lambda}_1 \quad = \quad |\psi_1\rangle^{\otimes(x+y)} \langle\psi_1|^{\otimes(x+y)} \tag{4.28}$$

$$\hat{\mu}_1 \quad = \quad |\psi_3\rangle^{\otimes z} \langle\psi_3|^{\otimes z}. \tag{4.29}$$

The general form of $|\psi_1\rangle$ is $\cos(\alpha/2) + e^{i\beta}\sin(\alpha/2)$. Assuming that $|\psi_1\rangle$ is in the northern cap, bounded by the angle θ, each term in the density matrix $\hat{\lambda}$ will be of the form:

$$\frac{1}{\kappa}\int_0^{2\pi} d\beta e^{ni\beta} \int_0^\theta d\alpha \cos^u(\alpha/2) sin^v \sin(\alpha/2)$$

where κ is the area of the cap $(\kappa = 2\pi(1 - \cos(\theta)))$ and $(u+v) = $ the number of bits needed to label a row or column of the matrix.

The powers v and u are obtained in the same way as in the other cases, but the integration process is different:

- define $K(m, n, \theta)$ as $\int_0^\theta d\omega \cos^n(\omega/2)\sin^m(\omega/2)$

- m and n both odd $\Rightarrow K(m, n) = 0$

- $n = 0 \Rightarrow K(m, n) = \left(\frac{4}{m-1}\right)\left(\frac{1}{1-\cos\theta}\right)\left(1 - \cos^{m+1}(\theta/2)\right)$

- otherwise $K(m, n) = K(m, n-2, \theta) - K(m+2, n-2, \theta)$

pcbm.c is the code for this module.

4.5   Computing error and failure rates

4.5.1   Optimum discrimination error rates

Helstrom's rule [17] allows the **optimal discrimination error rate** to be calculated directly from the two density matrices. We substitute various values of η_1 and η_2 and take the following steps:

- 1. Compute a difference operator $\hat{\Lambda} = (\eta_1\bar{\hat{\rho}}_1 - \eta_2\bar{\hat{\rho}}_2)$

- 2. Find s, the sum of the magnitudes of the eigenvalues of $\hat{\Lambda}$

- 3. Calculate the error rate as $(1 - s)/2$.

4.5.2   Unambiguous discrimination failure rates

Finding the **unambiguous discrimination failure rate** also requires several steps and is an order of magnitude more laborious than calculating the optimum error rates.

Following the procedure described in [50], we establish a POM with three component operators:

- $\hat{\Pi}_1$, which unambiguously recognises state $|\Psi_1\rangle$

- $\hat{\Pi}_2$, which unambiguously recognises state $|\Psi_2\rangle$

- $\hat{\Pi}_0$, which returns the result **unknown**.
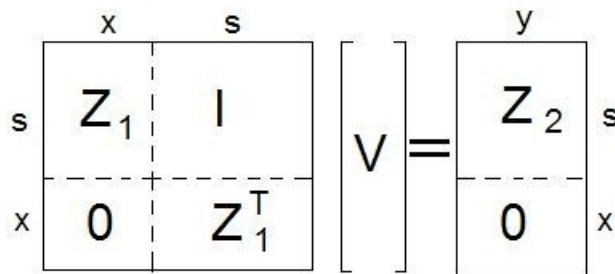
Figure 4.3: System of equations to calculate projectors

As one of these outcomes is certain to occur, we may write

$$\hat{\Pi}_0 + \hat{\Pi}_1 + \hat{\Pi}_2 = \mathbf{1} \tag{4.30}$$

$$\text{or } \hat{\Pi}_0 = \mathbf{1} - \hat{\Pi}_1 - \hat{\Pi}_2$$

where $\mathbf{1}$ is the unit matrix.

The steps in the computation are:

First we need to find two projectors $\hat{P}_1$ and $\hat{P}_2$, such that :

- a) $\text{Tr}(\hat{\rho}_1.\hat{P}_2) = \text{Tr}(\hat{\rho}_2.\hat{P}_1) = 0$

- b) $\hat{P}_1$ and $\hat{P}_2$ are completely orthogonal to one another

Using Jacobi's algorithm [?], we calculate the s eigenvectors of $\hat{\rho}_1$ and $\hat{\rho}_2$. and discard those with zero eigenvalues. We are left with a (rectangular) array Z_1 with s rows and x columns from $\hat{\rho}_1$, and an array Z_2 of dimension $s \times y$ from $\hat{\rho}_2$. Both x and y may be less than s (where s is the order of both $\hat{\rho}_1$ and $\hat{\rho}_2$). We suppose that each vector in Z_2 is composed of a linear combination of all the vectors in Z_1, plus a residual vector, which is what we need. This allows to write down x equations in s unknowns, which would not in general be enough to determine the residual vector, as $x \leq s$. We can find the extra equations by specifying that the residual vector is orthogonal to each of the vectors in Z_1.

The conditions can be arranged into a set of simultaneous linear equations, as shown in figure 4.3. Here all the vectors in Z_2 have been assembled into a matrix, padded out with zeros. When this system is solved using Gaussian elimination, the resultant array V holds the orthogonal vectors in its lower s columns. Naming them $|v_1\rangle$ to $|v_y\rangle$, the projector $\hat{P}_1$ can be constructed as:

$$\hat{P}_1 = \sum_{k=1}^{y} (|v_k\rangle\langle v_k|) \tag{4.31}$$

$\hat{P}_2$ is constructed in exactly the same way.

The probability for successful unambiguous discrimination is:

$$Q = \eta_1 \text{Tr}(c_1\hat{\rho}_1\hat{P}_1) + \eta_2 \text{Tr}(c_2\hat{\rho}_2\hat{P}_2)$$

where c_1 and c_2 are constants which must be selected so that $\hat{\Pi}_0$, defined in equation 4.31, has a smallest eigenvalue of zero, and Q has a minimum value.

It turns out that the relationship between the constants and the smallest eigenvalue of $\hat{\Pi}_0$ is **not** linear. We use a version of the *Golden Section Search* [2] to satisfy these conditions.

For certain small values of η_1 and η_2 the search does not converge, and there are no acceptable values of c_1 and c_2. This defines the range where the three-component POM is not valid.

The proportion of measurements, $(1 - Q)$, which return the result 'unknown' is the *failure rate*.

4.6   Results

The results in this section are generated by a suite of six programs that cover the three types of relationship (fixed overlap, known great circle, and polar caps) and both modes of discrimination (optimal and unambiguous) as applied to each relationship. We present the results of the fixed overlap programs in some detail, but give only brief summaries for the other two types.

4.6.1   Fixed overlap with optimal discrimination

First we examine multiple copies of the data qbit. Some results are given in table 4.1 for the configuration {1,1,1} and in table 4.2 for {1,6,1}. In these tables the vertical axis covers different values of η_1. The horizontal axis gives different values of β, where the overlap is $\cos^2(\beta/2)$. Figures in the body of the table represent **error rates**. Results for intermediate cases are not shown but can easily be calculated using the programs in [**?**]. In all cases the user can adjust the intervals between the values of β and η_1.

β	$\pi/6$	$\pi/3$	$\pi/2$	$2\pi/3$	$5\pi/6$	π
η_1						
0.25	0.241	0.217	0.183	0.150	0.125	0.116
0.50	0.481	0.428	0.356	0.283	0.231	0.211
0.75	0.241	0.217	0.183	0.150	0.125	0.116

Table 4.1: Fixed Overlap Optimal Discrimination {1,1,1}

β	$\pi/6$	$\pi/3$	$\pi/2$	$2\pi/3$	$5\pi/6$	π
η_1						
0.25	0.236	0.196	0.142	0.088	0.049	0.035
0.50	0.471	0.392	0.283	0.175	0.096	0.067
0.75	0.236	0.196	0.142	0.088	0.049	0.035

Table 4.2: Fixed Overlap Optimal Discrimination {1,6,1}
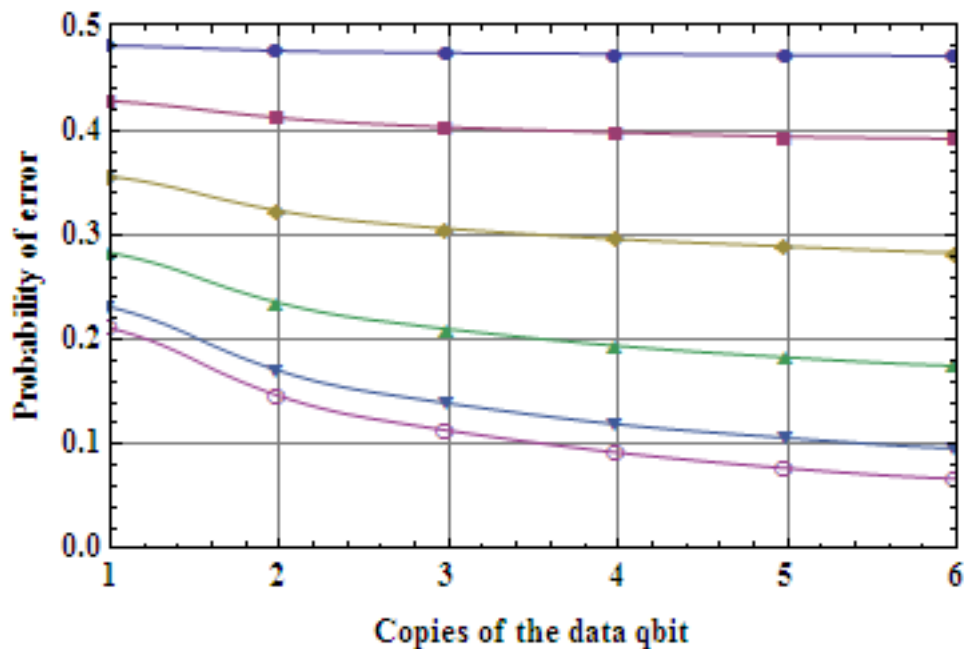
Figure 4.4: Configuration {1,n,1}. Optimal error for fixed overlap given by $\beta = \pi/6$ (top curve) to $\beta = \pi$ (bottom curve) in steps of $\pi/6$

Figure 4.4 is based on the configuration $\{1, n, 1\}(1 \leq n \leq 6)$. and shows how the error rate varies with the number of copies of the data qbit, for various degrees of overlap. The error rate generally falls when the number of copies is increased, but this effect is most marked when the overlap is small, and barely visible when the overlap is large.

Next we investigate the case where there are multiple copies of the program qbits, arranged symmetrically. We compare configurations of the forms $\{1, 2n - 1, 1\}$ with those of $\{n, 1, n\}$, which have the same number of qbits. Results show that the error rates are very similar.

Table 4.3, which is based on $\eta_1 = 0.5$ and includes columns for $\beta = \pi/2$ and $\beta = \pi$ makes this clear.

To end this section, we consider various non-symmetrical configurations. For $\beta = \pi/2$ the error rates for all configurations with the same number of qbits are close to one another, but for $\beta = \pi$ the relationship breaks down. Table 4.4 gives the figures.

The interesting conclusion is that in this mode (fixed overlap, optimum discrimination) it makes little difference how the qbits are used provided the arrangement is symmetrical: **any symmetric configuration with n qbits in total will result in substantially the same probable error rate**.

4.6.2 Fixed overlap with unambiguous discrimination

The results in this section are presented for a mid-range overlap ($\beta = \pi/2$). Figures for other degrees of overlap can be found by running the programs in [?].

Configuration	qbits	Error rate $(\beta = \pi/2)$	Error rate $(\beta = \pi)$
{1,1,1}	3	0.356	0211
{1,3,1}	5	0.306	0.113
{2,1,2}		0.308	0.115
{1,5,1}	7	0.289	0.077
{3,1,3}		0.288	0.076
{1,7,1}	9	0.280	0.059
{4,1,4}		0.279	0.057

Table 4.3: Fixed overlap optimal discrimination for symmetrical duplication of program qbits, compared to repetition of the data qbit.

Configuration	qbits	Error rate $(\beta = \pi/2)$	Error rate $(\beta = \pi)$
{1,1,1}	3	0.356	0211
{1,2,1}	4	0.323	0.146
{1,1,2}		0.319	0.138
{1,3,1}	5	0.306	0.113
{1,2,2}		0.287	0.073
{1,1,3}		0.302	0.105
{1,4,1}	6	0.296	0.092
{1,3,2}		0.272	0.045
{2,1,3}		0.294	0.089
{1,2,3}		0.272	0.045

Table 4.4: Fixed overlap optimal discrimination with asymmetrical replication of program qbits

As we mentioned briefly in the previous section, the graph of the relationship between the error rate and η_1, for any configuration, has three distinct sectors. In the centre of the range the best unambiguous discrimination is provided by a POM with three components. Outside this range, the 'best' discrimination rate is obtained by using one or other of the projectors $\hat{P}_1$ or $\hat{P}_2$. When η_1 is small, the error rate is given by:

$$\text{Failure rate} = \eta_1 + (1 - \eta_1)\text{Tr}(\hat{\rho}_2\hat{P}_2) \tag{4.32}$$

A similar equation holds when η_2 is small.

It is worth noting that when only one projector is used, one of the possible states of the system can sometimes be recognised unambiguously, but the other can **never** be detected. This suggests that in some circumstances, it might be better to keep using the three-component POM, even though it does not give the 'best' discrimination rates.

Typical curves are shown in figures 4.5, 4.6 and 4.7. Figures 4.5 and 4.6 use a fixed overlap
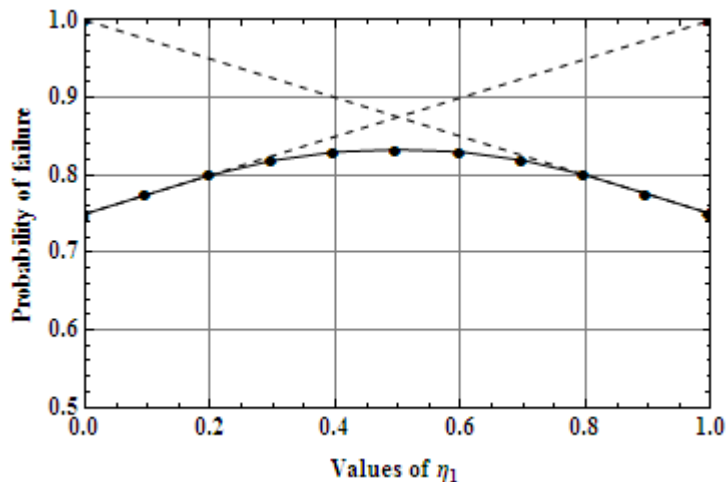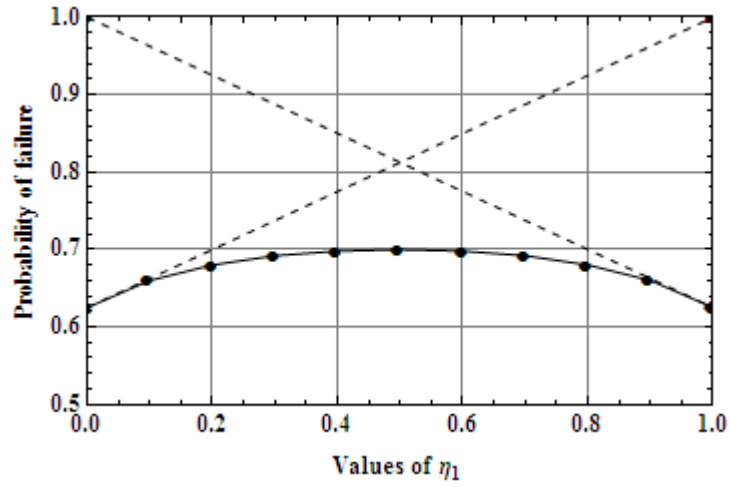
Figure 4.5: Fixed overlap $= \frac{1}{\sqrt{2}}$, failure rate in unambiguous discrimination for$\{1, 1, 1\}$

of $\frac{1}{\sqrt{2}}$. Corresponding figures drawn from analytic studies based on random program bits would look the same. We include our results here to give a more complete overview of the topic. The discrimination rates that correspond to the individual projectors are shown as dashed lines.

These curves illustrate two trends:

- 1) Symmetrical configurations, where the data qbit is repeated several times, improve the rate of successful discrimination. They also widen the useful range of the three-component POM.

- 2) Non-symmetrical configurations, in which one of the program qbits is duplicated, narrow the range of the POM, and give hardly any improvement in the rate of discrimination.

4.6.3   Known great circle with optimal discrimination

This set of results has one dimension fewer than the previous set, as we do not need to investigate different overlaps.

Figure 4.8 shows the error rates for three symmetrical configurations - $\{1, 1, 1\}$, $\{1, 2, 1\}$ and $\{1, 3, 1\}$. Increasing the number of data qbits gives a modest improvement of some 20% in the correct discrimination rate. Figure 4.9 illustrates the error rate for non-symmetrical configurations - $\{1, 1, 1\}$ to $\{1, 1, 3\}$. As the number of program qbits on the right increases, the discrimination rate improves slightly, and the optimum value of η_1 shifts away from 0.5.

4.6.4   Known great circle with unambiguous discrimination

Failure rates for $\{1, 1, 1\}$ to $\{1, 3, 1\}$ are shown in Figure 4.10. Finally, figure 4.11 shows the failure rates for non-symmetrical configurations - $\{1, 1, 1\}$ to $\{1, 1, 3\}$.

69

Figure 4.6: Fixed overlap $= \frac{1}{\sqrt{2}}$, failure rate in unambiguous discrimination for$\{1, 3, 1\}$
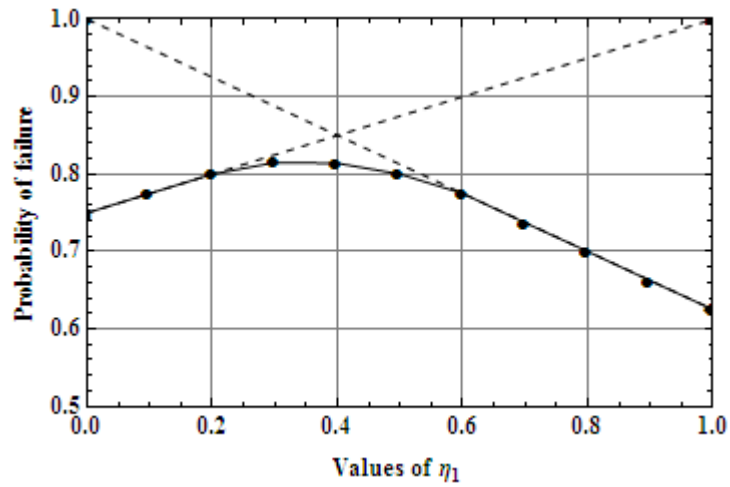
Figure 4.7: Fixed overlap $= \frac{1}{\sqrt{2}}$, probable error rate in unambiguous discrimination for$\{1, 1, 3\}$
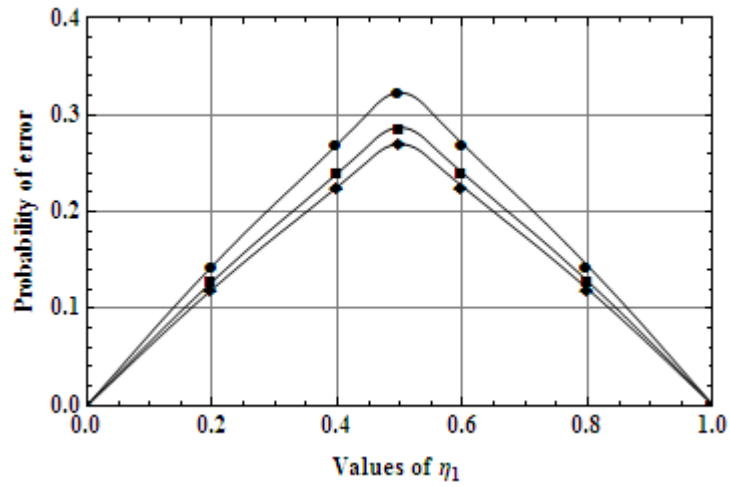
Figure 4.8: Known great circle, error rate in optimal discrimination for $\{1, 1, 1\}$ (upper curve) to $\{1, 3, 1\}$ (lower curve)
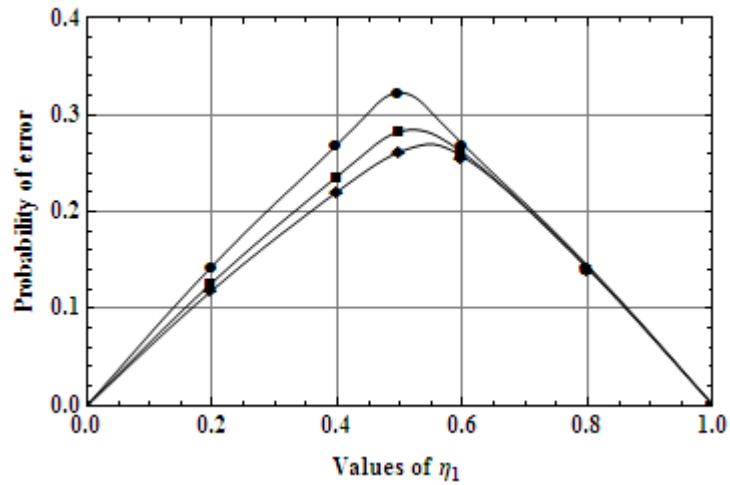
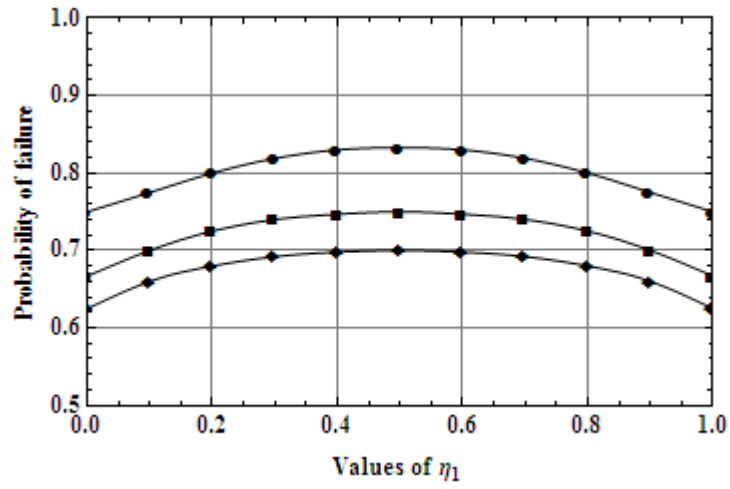Figure 4.9: Error rates (great circle, optimal discrimination) for $\{1, 1, 1\}$ (upper curve) to $\{1, 1, 3\}$ (lower curve)

Figure 4.10: Failure rates (great circle, unambiguous discrimination) for $\{1,1,1\}$ (upper curve) to $\{1,3,1\}$, (lower curve)
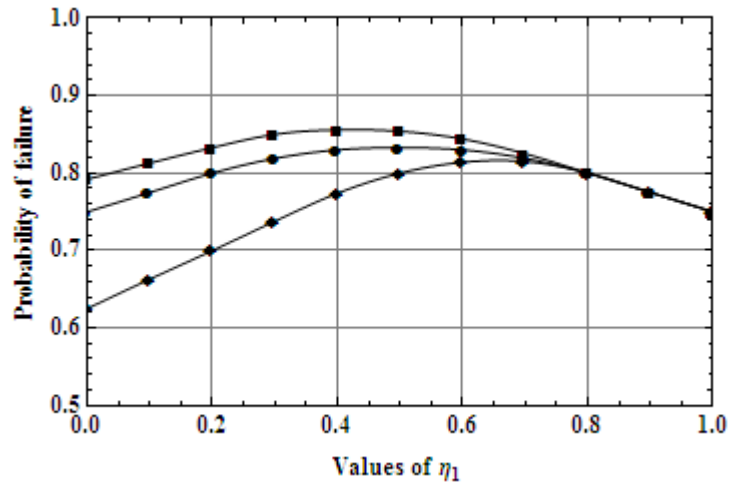
Figure 4.11: Failure rates (great circle, unambiguous discrimination) for $\{1,1,1\}$ (upper curve), $\{1,1,2\}$ (central curve) and $\{1,1,3\}$ (lower curve)
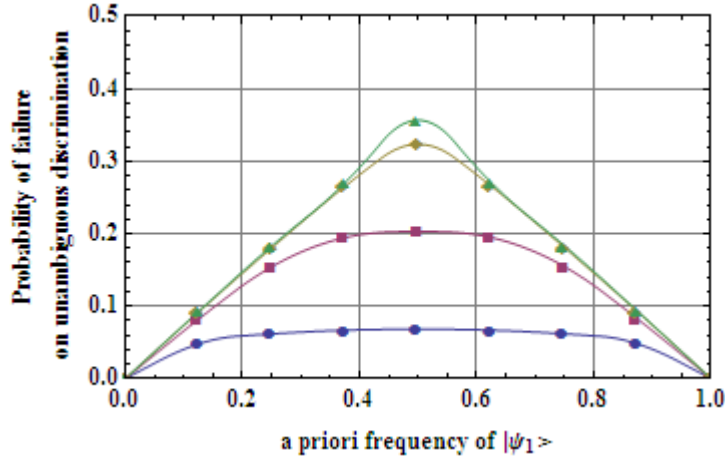
Figure 4.12: Optimal error rates (polar cap confinement) for $\{1, 1, 3\}$. The upper curve is for $\theta = \pi$, and the lower one for $\theta = \pi/6$.

4.6.5 Confinement to polar caps with optimal discrimination

Figure 4.1c shows how the angle θ is related to the size of the polar caps that hold the program qbits. Figure 4.12 plots the variation of the minimum error against η_1 for different values of θ. As the area of confinement decreases, so does the minimum error.

Doubling up the data qbit has little overall effect except to decrease the error rate slightly. Doubling up the right program qbit skews the diagram to the right.

4.6.6 Confinement to polar caps with unambiguous discrimination

Figure 4.13 shows how the unambiguous discrimination rate varies with the degree of confinement for various sizes of polar cap. This diagram is for the asymmetric configuration $\{1, 1, 3\}$, but apart from a small skew it is similar to other configurations in this context. In summary, it appears that discrimination rates in both the minimum error and unambiguous regimes depend strongly on the degree to which the program qbits are confined to the poles of the Bloch sphere. The number of qbits available, whether as data or program qbits, has only minor influence.

4.7 Transmitting data using programmed discrimination

The merits of using qbits to ensure data security are well established, and several appropriate protocols have been defined [31, 44, 52]. A potential application of programmed discrimination is in data transmission. The two possibilities $|\psi_2\rangle \equiv |\psi_1\rangle$ and $|\psi_2\rangle \equiv |\psi_3\rangle$ are used to encode the binary digits 0 and 1. In this section we examine the data-carrying capacity of a data transmission system that uses programmed discrimination. We use some of the results presented in previous sections of this chapter, as well as those derived from expressions in other published material [41, 42]. A useful measure will be the amount of information, in *bits*, reliably transmitted by a
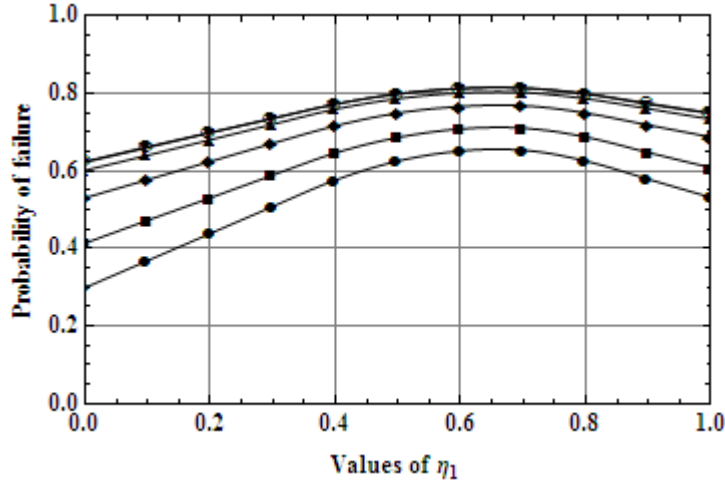
Figure 4.13: Failure rates (polar cap confinement) for$\{1,1,1\}$. The upper curve is for $\theta = \pi$, and the lower one for $\theta = \pi/6$.

single qbit.

The advantage of programmed discrimination over other methods is that the qbits can be allowed to undergo any unitary transformation during transmission without affecting the data they represent, as long as the same transformation applies to all the qbits in a group. In our study this advantage holds for groups in which the overlap of the program qbits is fixed, but not for the other two modes, as an arbitrary transformation could move the qbits away from a known great circle, or away from given polar caps on the Bloch sphere. Another restriction is that we only consider configurations in which the two possible signals are handled symmetrically.

It is well known that any digital data transmission system will work at its highest efficiency if the expectations of zeros and ones are statistically equal. There exist many algorithms to compress data to this standard. We would therefore expect that in any discrimination system, η_1 and η_2 (as previously defined in equations 1 and 2) should both be 0.5. If one used a configuration with unequal expectations where, for example, $\eta_1 = \frac{1}{3}$ and $\eta_2 = \frac{2}{3}$, the data would have to be recoded so that the ratio of zeros and ones was 2 to 1. This is not a realistic proposal.

Another form of asymmetry occurs when the numbers of program qbits of each type are not equal, as in the configuration $\{1, 3, 2\}$. The figures we have derived for these configurations are *averages*, but without symmetry the recognition rates will be different for zeros, as opposed to ones. The analysis of this situation is beyond the scope of this chapter.

4.7.1 Using optimal recognition

Consider a channel in which the probability of receiving a binary digit incorrectly is p. Shannon's Noisy Channel Theorem [53] states that provided that $p \neq 0.5$, standard error correction methods allow the channel to be used for error-free transmission, albeit at a reduced data rate. The

Configuration	{1,n,1}	{2,n,2}	{3,n,3}
Data qbits (n)			
1	0.02033	0.02192	0.01908
2	0.02304	0.02411	0.02064
3	0.02222	0.02292	0.01959
4	0.02063	0.02111	
5	0.01899		
6	0.01747		
7	0.01613		

Table 4.5: Bits per qbit, with random program qbits and optimal discrimination

effective data rate R, per bit transmitted, is:

$$R = 1 + p \times \log_2(p) + (1 - p) \times \log_2(1 - p) \text{ bits.} \qquad (4.33)$$

Suppose a group of n qbits is used to transmit a (classical) bit, with an expected error rate of p. The amount of information k sent by one qbit is:

$$k = \frac{1 + p \times \log_2(p) + (1 - p) \times \log_2(1 - p)}{n} \qquad (4.34)$$

Using error rates for optimal discrimination with no correlation between the program qbits, [42] the information transmitted per qbit, for various configurations, is shown in table 4.5. The efficiency of communication in this context, when compared to a possible standard of one qbit per bit, is barely 2% with the configuration $\{1, 1, 1\}$. This figure improves slightly as the data bit is duplicated, and then falls away again when further copies are brought into play. With multiple program qbits the best configuration, by a small margin, is $\{2, 2, 2\}$.

In previous sections of this thesis, we have calculated the expected error rates for various configurations, and for various degrees of fixed overlap. For orthogonal program qbits (where the overlap is 0) the number of bits per qbit is shown in table 4.6. Here the presence of classical information gives a four-fold improvement, and further illustrates the small, but real advantage of using multiple qbits.

4.7.2 Using unambiguous recognition

Shannon's theorems are not relevant to the protocol that might be used with unambiguous discrimination. Suppose Alice is sending data to Bob [54], using the basic $\{1, 1, 1\}$ configuration with no correlation between the data qbits. We know that Bob can only make an identification in one sixth of the triads, but in these cases there is no uncertainty about the outcome. For effective communication Bob must use a conventional data link to tell Alice when he receives a triad that he recognises. To send a bit reliably, Alice must keep re-transmitting the corresponding triad until she learns that Bob has received it. With the basic configuration Alice must send an average of six triads or 18 qbits for every bit, giving a rate of 0.0555 bits per qbit.

Configuration	{1,n,1}	{2,n,2}	{3,n,3}
Data qbits (n)			
1	0.08533	0.09690	0.08721
2	0.09978	0.11219	0.10013
3	0.09840	0.10996	0.09800
4	0.09295	0.10339	
5	0.08671		
6	0.08068		
7	0.07513		

Table 4.6: Bits per qbit, with orthogonal program qbits and optimal discrimination

Program qbits	1+1	2+2	3+3
Data qbits			
1	0.05556	0.04000	0.03061
2	0.06250	0.04630	0.03515
3	0.06000	0.04463	
4	0.05556	0.04375	
5	0.05102		
6	0.04688		
7	0.04321		

Table 4.7: Bits per qbit, with random program qbits and unambiguous discrimination

The effectiveness of different configurations without correlation is shown in table 4.7, and that for zero overlap in table 4.8. The best overall rate is achieved with the configuration $\{1, 2, 1\}$, using orthogonal program qbits and unambiguous recognition.

4.8 Conclusion

Our overall conclusion, which applies for all the cases we have considered, is that increasing the qbit count generally gives only modest increases in the rate of successful discrimination.

One exception to the rule, worth noting, is that confining the program qbits to small areas near the poles of the Bloch sphere does lead to a substantial improvement in the rate of correct discrimination. But this restriction violates the freedom to subject all the qbits in the triad to *any* unitary transformation, so the finding is of doubtful value, at least where data transmission is concerned.

Another curious result is that in the minimum error mode, the error rate depends mainly on the number of qbits used to transmit a bit, and much less on how they are distributed between program and data qbits.

For data transmission, the best performance we could find is given by a configuration with

Program qbits	1+1	2+2	3+3
Data qbits			
1	0.11111	0.08000	0.06123
2	0.12500	0.09259	0.07031
3	0.12000	0.09184	
4	0.11111	0.08750	
5	0.10204		
6	0.09375		
7	0.08642		

Table 4.8: Bits per qbit, with orthogonal program qbits and unambiguous discrimination

orthogonal qbits, unambiguous discrimination and some replication of the data qbit. We would need 8 qbits to transmit a single binary digit reliably. This figure is poor when compared to other communication methods. There would need to be good reasons to justify the use of this method in a practical situation.

Chapter 5

A summary of "Measurement Master Equation" [55]

"Sometimes the probabilities are very close to certainties, but they're never really certainties."
Murray Gell-Mann

For the next two chapters of this thesis, we consider measurement from a different point of view. We are no longer interested in discrimination between two states, but in the effect that measurement has on the evolution of an otherwise closed system.

This chapter summarises the paper entitled *Measurement master equation*, by Cresser, Barnett, Jeffers and Pegg [55]. This paper forms the background to our own investigation, which is described in Chapter 6 [56].

5.1   Summary of "Measurement master equation"

The authors used a Master equation [28, 29] to examine the behaviour of a quantum system under weak, but frequent measurement. The physical system they considered is an isolated two-state atom, with energy eigenstates $|1\rangle$ and $|0\rangle$, resonantly driven by an electromagnetic field. On its own, this forms a closed system. The probable state of the atom undergoes Rabi oscillations, at a rate that depends on the strength of the applied field.

It is helpful to treat the atom as effectively spinning, and in this context we introduce two Pauli operators:

$$\hat{\sigma}_1 = |0\rangle\langle 1| + |1\rangle\langle 1| \tag{5.1}$$
$$\hat{\sigma}_3 = |1\rangle\langle 1| - |0\rangle\langle 0|.$$

The hamiltonian that describes the interaction between the field and the atom is

$$\hat{H} = -\frac{\hbar\Omega}{2}\hat{\sigma}_1 \tag{5.2}$$

where Ω is the Rabi frequency.

The application of measurements implies that the system is no longer closed. Its trajectory is best described by a Master equation in the Lindblad form [28]:

$$\frac{d\hat{\rho}}{dt} = \frac{-i}{\hbar}[\hat{H}, \hat{\rho}] + \sum_j \gamma_j \left(\hat{K}_j \hat{\rho} \hat{K}_j^\dagger - \frac{1}{2}\hat{K}_j^\dagger \hat{K}_j \hat{\rho} - \frac{1}{2}\hat{\rho}\hat{K}_j^\dagger \hat{K}_j \right) \tag{5.3}$$

where $\hat{\rho}$ is the reduced density operator, $\hat{H}$ the system hamiltonian, and γ and K_j the Kraus operators which depend on the frequency and characteristics of the measurements.

Energy measurements would ideally correspond to the projectors $|1\rangle\langle 1|$ and $|0\rangle\langle 0|$. However, it is specified that the measurements may be weak, and have only a probability p $(p \leq 1)$ of returning a correct result. This leads to a POM with two elements:

$$\hat{\pi}_1 = p|1\rangle\langle 1| + (1-p)|0\rangle\langle 0| \tag{5.4}$$

$$\hat{\pi}_2 = p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|. \tag{5.5}$$

These operators satisfy the requirement that $\sum_j \hat{\pi}_j = \mathbf{1}$.

As the interaction with the environment takes the form of a measurement, each pair of Kraus operators [57] corresponds to one of the elements of the POM, according to the rule

$$\hat{\pi}_j = \hat{K}_j^\dagger \hat{K}_j. \tag{5.6}$$

These relationships do not specify a unique set of operators K_j. The simplest realisation of these Kraus operators are the square roots of the POM elements which, for our two probability operators, are:

$$\hat{K}_1 = \hat{K}_1^\dagger = \sqrt{p}|1\rangle\langle 1| + \sqrt{1-p}|0\rangle\langle 0| \tag{5.7}$$

$$\hat{K}_2 = \hat{K}_2^\dagger = \sqrt{p}|0\rangle\langle 0| + \sqrt{1-p}|1\rangle\langle 1| \tag{5.8}$$

Having defined the concept of a *weak* measurement, the paper now considers the aspect of measurement frequency. The authors use a parameter R, which indicates how often a measurement takes place, in terms of the average number of measurements per Rabi cycle. The presence of R modifies the basic Master equation, as follows:

$$\mathcal{T}(\hat{\rho}) = \frac{-i}{\hbar}[\hat{H}, \hat{\rho}] + R\left[\sum_j (\hat{K}_j \hat{\rho} \hat{K}_j^\dagger - \tfrac{1}{2}\hat{K}_j^\dagger \hat{K}_j \hat{\rho} - \tfrac{1}{2}\hat{\rho}\hat{K}_j^\dagger \hat{K}_j)\right]. \tag{5.9}$$

Using the fact that

$$\sum_j (\hat{K}_j^\dagger \hat{K}_j) = \mathbf{1}$$

this can be written as

$$\frac{d\hat{\rho}}{dt} = i\frac{\Omega}{2}[\hat{\sigma}_1, \hat{\rho}] + \frac{R}{2}\left(\sqrt{1-p} - \sqrt{p}\right)^2 (\hat{\sigma}_3 \hat{\rho} \hat{\sigma}_3 - \hat{\rho}) \tag{5.10}$$

or

$$\frac{d\hat{\rho}}{dt} = i\frac{\Omega}{2}[\hat{\sigma}_1, \hat{\rho}] + \gamma(\hat{\sigma}_3 \hat{\rho} \hat{\sigma}_3 - \hat{\rho}) \tag{5.11}$$

where

$$\gamma = \frac{R}{2}\left(\sqrt{1-p} - \sqrt{p}\right)^2. \tag{5.12}$$

The analytic solution of this equation shows that starting in a pure state $|\psi\rangle = |0\rangle$ or $\quad |\psi\rangle = |1\rangle$ a sufficiently large ensemble of atoms will decay into an average mixed state

$$\hat{\rho} = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|). \tag{5.13}$$

79

The decay is oscillatory if γ is small, but exponential for larger values of γ.

The Master equation can also be solved numerically to provide individual trajectories of the atom. The actual intervals between measurements depend on a random number generator. These trajectories are quite different from the analytic solution; they do not decay, and with high values of γ they exhibit *telegraphing*, where the Rabi cycle is suppressed and the atom spends long periods at or near the upper or lower energy level. Figures 5.1 to 5.4 show some simulated trajectories.
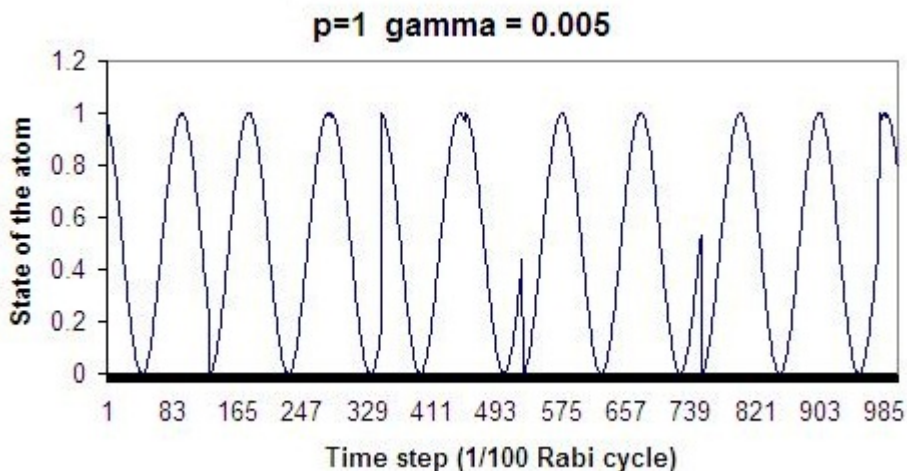
Figure 5.1: Trajectory with low gamma

Figure 5.1 uses a very low value of γ. The basic Rabi cycle is hardly disturbed. This continues indefinitely. Figure 5.2 is a trajectory produced by an intermediate value of γ. It is not obvious, from the graph, whether telegraphing is occurring or not. Figure 5.3 uses a much higher value of γ, and demonstrates the 'telegraphing' effect. The horizontal scale has been extended to 20 Rabi cycles. Figure 5.4 illustrates the details of a trajectory when the system is telegraphing, and switches from one state to the other. The 'spikes' can only occur when there is a possibility of incorrect measurement.
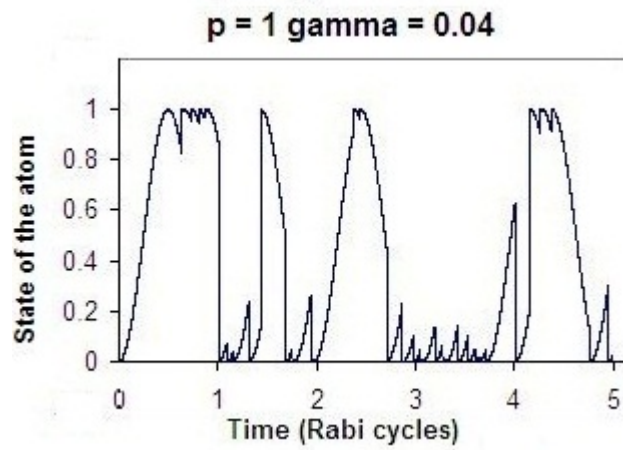
Figure 5.2: Trajectory with intermediate gamma
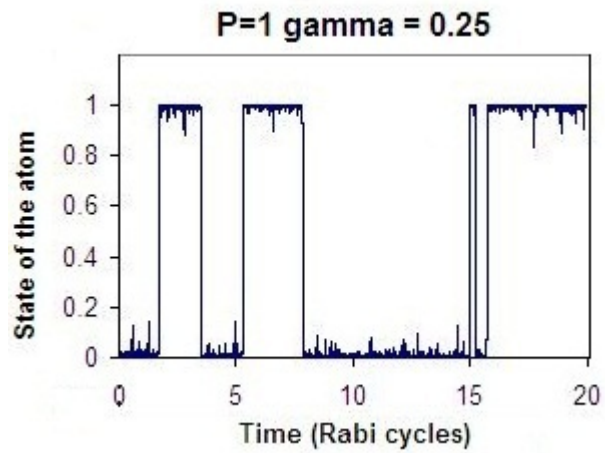
Figure 5.3: Trajectory with high gamma
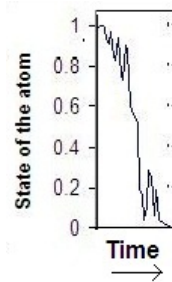
Figure 5.4: Switching trajectory when $p < 1$

In the final section of their paper, the authors provide a qualitative description of several phenomena which can be seen in these trajectories:

- Why very low values of γ have little effect on the continuous Rabi cycle

- Why telegraphing occurs with high values of γ

- Why the trace has a 'whiskery' appearance when telegraphing (as in figure 5.2)

- Why transitions between the two energy states may not be instantaneous but 'jagged' (as in figure 5.4).

Chapter 6

Measurement-driven dynamics for a coherently-excited atom

"You are never too old to set another goal or to dream a new dream."
C. S. Lewis

Chapter 6 offers a quantitative analysis of the phenomena described by Cresser, Barnett, Jeffers and Pegg [55] and summarised in Chapter 5. The predictions of our theoretical analysis, which takes into account two alternative definitions of telegraphing, are matched against practical observations derived from a simulator. Please note that there are several references to figures in the previous chapter.

6.1 The simulator

Our results are derived from a simulator, which is based on a 4th-order Runge Kutta integration procedure using a time-step, δt, of $\frac{1}{1000}$ of a Rabi cycle. Our reported statistics are based on runs of 1000 Rabi cycles, or 10^6 time-steps. The simulator takes two parameters:

- R, the average rate of measurements. In the simulator this is converted to a fraction q, the probability of a measurement in any one time-step.

$$q = \delta t \left(\frac{2\pi R}{\Omega} \right).$$

(6.1)

  where Ω is the angular frequency of the Rabi cycle.

- p, the probability that a measurement returns a result in keeping with the current value of the density matrix. There is a probability $(1 - p)$ that this result be inverted.

The Rabi cycle is simulated by solving Schrődinger's equation

$$\frac{d\hat{\rho}}{dt} = \frac{-i}{\hbar} \left[\hat{H}, \hat{\rho} \right]$$

(6.2)

where

$$\hat{H} = \begin{pmatrix} 0 & \frac{-\hbar\Omega}{2} \\ \frac{-\hbar\Omega}{2} & 0 \end{pmatrix}$$

(6.3)

and Ω is the angular frequency of the Rabi oscillation.

As a starting value we take

$$\hat{\rho} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad (6.4)$$

which corresponds to an initially excited atom. The occurrence of a measurement during any time cycle is determined by a rectangular random number generator with a range g, where $(0 \leq g \leq 1)$. A measurement takes place if $(g < q)$. It is assumed that measurements are so rare that there will never be more than one measurement in the same time-step. The validity of the model is bounded in two ways:

- The model is inaccurate for very high values of R, as there is then a significant chance that two or more measurements could fall in the same time-step. The simulator does not handle this situation, but its validity could be extended by using a shorter time-step.

- When R is very low, there are not enough events, even during 1000 Rabi cycles, to yield accurate statistics. A longer period of observation would solve this problem.

Notwithstanding these practical limitations, the model is accurate over a wide range of values for R : $(0.01 \leq R \leq 1000)$.

The overall effect of an imperfect measurement was computed as follows: suppose that the measurement has a probability p of being correct, and that when it is applied, the atom is part-way through its Rabi cycle, so that the probability of its being in state $|1\rangle$ is x. Then a correct measurement, using the projectors $|1\rangle\langle1|$ and $|0\rangle\langle0|$, would return $|1\rangle$ with probability x, and $|0\rangle$ with probability $(1 - x)$. But allowing for the weakness of the measurement, the actual probabilities are

$$(P|1\rangle) = px + (1 - p)(1 - x) \qquad (6.5)$$

$$(P|0\rangle) = p(1 - x) + (1 - p)x \qquad (6.6)$$

If the measurement is imperfect, the final state of the atom will not be the same as the result of the measurement. If the measurement returns $|j\rangle$, the change in the density matrix is

$$\hat{\rho} \rightarrow \frac{\hat{K}_j \hat{\rho} \hat{K}_j^\dagger}{\mathrm{Tr}(\hat{K}_j \hat{\rho} \hat{K}_j^\dagger)} \qquad (6.7)$$

where K_j and $K_j^\dagger$ are defined in equations 5.7 and 5.8. If $p = 1$ the effect is to shift the density matrix exactly to one of the states $|1\rangle$ or $|0\rangle$, but for any other value of p the density matrix is moved *towards* one of the states without actually reaching it [13].

6.2 Definitions of telegraphing

The atom under investigation alternates between two phases which we term 'up' dwells, associated with the eigenstate $|1\rangle$, and 'down' dwells, associated with $|0\rangle$. Measurements which move the state of the atom towards a given eigenstate are 'up' or 'down' measurements, respectively.

Telegraphing can be defined in two different ways:

- The *Measurement Centred* method relies entirely on the record of measurements. A dwell is defined as the interval between the first measurement to yield a given result (say $|1\rangle$) and the next following measurement that gives the opposite result. Intermediate measurements that give the same result are ignored. If measurements are widely spaced the atom may undergo several Rabi oscillations during the dwell. If there are no measurements, the dwell time is infinite. This method has the advantage that as the statistics rely on the outcomes of *measurements*, they could be obtained from real (not simulated) systems.
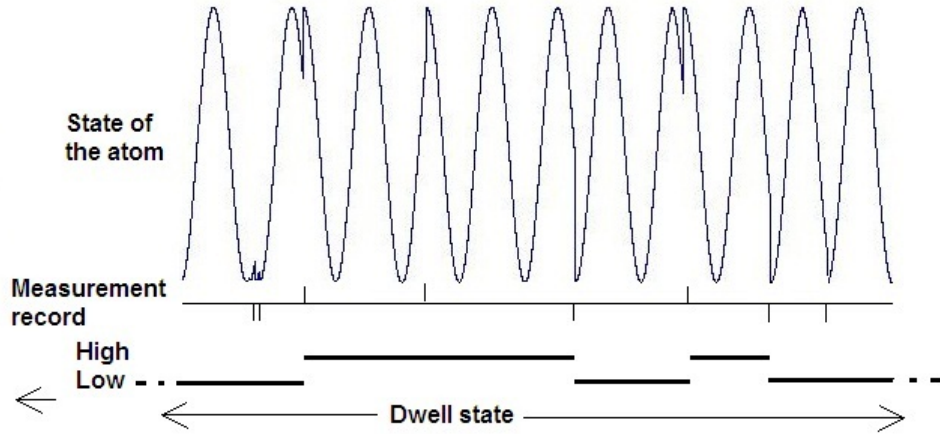
Figure 6.1: A typical measurement-centred dwell record with $p = 1$

- The *State Centred* method defines a dwell as the uninterrupted time period during which the probability of the atom being at or near one of its energy states $|1\rangle$ or $|0\rangle$ is greater than 0.5. This can only be found by a device or simulator that 'knows' or can estimate the state of the system between measurements. On the other hand, this definition is much closer to the intuitive understanding of telegraphing, specially for high values of γ.

  When an atom undergoes an undisturbed Rabi evolution without any measurements, then under the state-centred method, there will be two dwells for each Rabi cycle. The state centred method forms a basis for the Telegraph index, a numerical descriptor of the behaviour of an atom. Figure 5.3 illustrates a trajectory with state centred dwells, most of which last considerably longer than one Rabi cycle.

We present theoretical analyses for both methods with perfect measurement, and compare the results with readings taken from the simulator.

6.3 Telegraph index

A comparison of figures 5.3 and 5.1 clearly shows that telegraphing is present in one case, but not in the other. Figure 5.2 cannot readily be assigned to either of these classes; the presence

of telegraphing is not a binary quality. This suggests the definition of a 'telegraphing measure' or index, which would vary between 0 when a system is not telegraphing at all, and 1 when measurement is so frequent that the system almost never changes state. We define the telegraph index to be the time-averaged quantity:

$$T = \frac{1}{\tau} \int_0^\tau dt [2\langle \hat{\sigma}_3 \rangle^2 - 1], \tag{6.8}$$

where τ greatly exceeds the Rabi period. For pure Rabi oscillations this quantity is zero, but for a perfect telegraph evolution, in which at all times the atom is in state $|0\rangle$ or state $|1\rangle$, it will be unity. Applying this formula to the trajectories in figures 5.2 to 5.4 we find:

Figure	Telegraph Index
Figure 5.1	0.019
Figure 5.2	0.672
Figure 5.3	0.958

We would expect the Telegraph index to be asymptotic to 1 as γ increases indefinitely. Figure 6.2, which includes the points listed above and some others, shows that this is so. The fit is not exact, as the points on the graph are derived from a procedure that includes random elements.
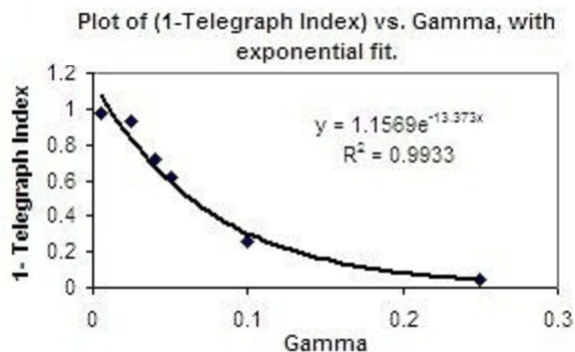
Figure 6.2: Relationship of Telegraph Index and γ

6.4 Analysis of the measurement centred method, with perfect measurements

This analysis is based on [30]. The derivation assumes measurements that give correct results ($p = 1$), so that after a measurement the state of the atom is exactly either $|1\rangle$ or $|0\rangle$. A *jump* is a measurement that produces the opposite result from the previous measurement. R is the rate of measurements per unit time.

Suppose that at $t = 0$ (just after a measurement) the state of the atom is $|1\rangle$. (It could equally well be $|0\rangle$, as the argument is symmetrical about these two states.)

If Ω is the Rabi frequency, the Rabi evolution gives

$$|\psi\rangle = \cos\left(\frac{\Omega t}{2}\right)|1\rangle + i\sin\left(\frac{\Omega t}{2}\right)|0\rangle. \tag{6.9}$$

The first measurement after $t = 0$ occurs between t and $t + dt$. The probability P_1 of a jump is:

$$P_1 = e^{-Rt}Rdt\sin^2\left(\frac{\Omega t}{2}\right). \tag{6.10}$$

The components of this expression are

- e^{-Rt} is the probability of reaching time t without a measurement

- Rdt is the probability of a measurement during this period

- $\sin^2\left(\frac{\Omega t}{2}\right)$ is the probability that a measurement causes a jump down to state $|0\rangle$.

The probability $\sim P_1$ that the first measurement does *not* give a jump is

$$\sim P_1 = e^{-Rt}Rdt\cos^2\left(\frac{\Omega t}{2}\right). \tag{6.11}$$

We can now extend this to the case of several no-jumps, at times t', t'', $\cdots$ (working backwards) followed by a jump at time t. For example,

$$P_3 = e^{-Rt''}Rdt''\cos^2\left(\frac{\Omega t''}{2}\right) \times e^{-Rt'}Rdt'\cos^2\left(\frac{\Omega(t'-t'')}{2}\right) \times e^{-Rt}Rdt\sin^2\left(\frac{\Omega(t-t')}{2}\right). \tag{6.12}$$

This can be rewritten as

$$P_3 = e^{-Rt}R^3dt''\,dt'\,dt\sin^2\left(\frac{\Omega(t-t')}{2}\right) \times \cos^2\left(\frac{\Omega(t'-t'')}{2}\right)\cos^2\left(\frac{\Omega t''}{2}\right). \tag{6.13}$$

Now we find the *average* time $\overline{T}$ between changes of state:

$$\begin{aligned}
\overline{T} = {}& \int_0^\infty Rte^{-Rt}dt\left[\sin^2\left(\frac{\Omega t}{2}\right)\right. \\
& + \int_0^t Rdt'\sin^2\left(\frac{\Omega(t-t')}{2}\right)\cos^2\left(\frac{\Omega t'}{2}\right) \\
& + \int_0^t Rdt''\sin^2\left(\frac{\Omega(t-t')}{2}\right)\cos^2\left(\frac{\Omega(t'-t'')}{2}\right)\cos^2\left(\frac{\Omega t''}{2}\right) \\
& \left. + \dots\right].
\end{aligned} \tag{6.14}$$

Using the technique of *differentiation under the integral sign* we get

$$\overline{T} = \frac{-d}{dq}\int_0^\infty e^{-qRt}dt\left[\dots\right]\Bigg|_{q=1} \tag{6.15}$$

where the contents of $[\cdots]$ are the same as in the previous equation. When this expression is expanded, the first term

$$\int_0^\infty e^{-qRt}dt\,\sin^2\left(\frac{\Omega t}{2}\right)$$

has the form of a Laplace transform, and the second and subsequent terms, the form of Laplace transform convolutions such as:

$$\int_0^\infty e^{-qRt} dt\ \sin^2\left(\frac{\Omega t}{2}\right) \int_0^t Rdt'\sin^2\left(\frac{\Omega(t-t')}{2}\right) \times \cos^2\left(\frac{\Omega t'}{2}\right).$$

By evaluating the transforms and convolutions we can replace the integrals with algebraic expressions. This leads to the transforms

$$S = \mathcal{L}\left[\int_0^\infty e^{-qRt} dt\ \sin^2\left(\frac{\Omega t}{2}\right)\right] = \frac{1}{2}\left[\frac{1}{qR} - \frac{qR}{q^2R^2+\Omega^2}\right]$$
$$C = \mathcal{L}\left[\int_0^\infty e^{-qRt} dt\ \cos^2\left(\frac{\Omega t}{2}\right)\right] = \frac{1}{2}\left[\frac{1}{qR} + \frac{qR}{q^2R^2+\Omega^2}\right].$$

$$(6.16)$$

It follows that

$$\overline{T} = \left(\frac{-d}{dq}\right)\left[S(qR) + RC(qR)S(qr) + R^2C^2(qR)S(qR) + \cdots\right]\big|_{q=1} \qquad (6.17)$$

$$= \left(\frac{-d}{dq}\right)\frac{S(qr)}{1-RC(qR)}\bigg|_{q=1}$$

$$= \left(\frac{-d}{dq}\right)(\frac{1}{2})\frac{\left[\frac{1}{qR} - \frac{qR}{q^2R^2+\Omega^2}\right]}{1-\frac{R}{2}\left[\frac{1}{qR} - \frac{pR}{q^2R^2+\Omega^2}\right]}\bigg|_{q=1}$$

Simplifying, differentiating with regard to q and simplifying again, gives

$$\boxed{\overline{T} = \frac{2R}{\Omega^2} + \frac{2}{R}}$$

$$(6.18)$$

This is the *dwell time*, or average time between changes of state. It remains to be shown that the P_n form is a valid probability distribution. We define a *moment generation function $I(q)$* as:

$$I(q) = \int_0^\infty e^{-qRt} dt[\sin^2(\frac{\Omega t}{2}) + ...].$$

$$(6.19)$$

We expect that

$$R\ I(q)|_{q=1} = 1.$$

$$(6.20)$$

Writing $I(q)$ as a Laplace transform gives:

$$I(q) = \frac{\frac{1}{2}\left(\frac{1}{qR} - \frac{qr}{q^2r^2+\Omega^2}\right)}{1-\frac{R}{2}\left(\frac{qr}{q^2r^2+\Omega^2}\right)} = \frac{\Omega^2}{2}\left(q^3R^3 - q^2R^3 + qR\Omega^2 - \frac{R\Omega^2}{2}\right)^{-1}$$

$$(6.21)$$

It follows that

$$R\ I(q)|_{q=1}\ \frac{R\Omega^2}{2}\left(R^3 - R^3 + R\Omega^2 - \frac{R\Omega^2}{2}\right)^{-1} = 1\ \text{(as expected)}$$

$$(6.22)$$

We can also find the variance of the dwell times, using similar methods. The variance of the dwell time V is given by

$$V = \overline{T^2} - (\overline{T})^2$$

$$(6.23)$$

where

$$\overline{T^2} = \frac{1}{R}\left(\frac{d^2}{dq^2}\right)(I(q))\Big|_{q=1} = \frac{4R^2}{\Omega^4} + \frac{4}{R^2} + \frac{8}{\Omega^2} \qquad (6.24)$$

which results in

$$V = \overline{T^2} - (\overline{T})^2 = \left(\frac{4R^2}{\Omega^4} + \frac{4}{R^2}\right) \qquad (6.25)$$

and the standard deviation of each dwell

$$\boxed{\sigma} = \sqrt{\left(\frac{4R^2}{\Omega^4} + \frac{4}{R^2}\right)} \qquad (6.26)$$

6.4.1 Measurements of measurement-centred dwells

Measurements on the simulator are subject to the expected statistical variation, but the averages show excellent agreement with the dwell lengths predicted by this formula. We ran the simulator for some 60 values of R ranging from 0.01 to 1000 measurements per Rabi cycle, increasing exponentially in steps of 1.2 . Figure 6.3 is a plot of $\log(R)$ against $\log(\overline{T})$; we use logs as otherwise small values of R and T would disappear.
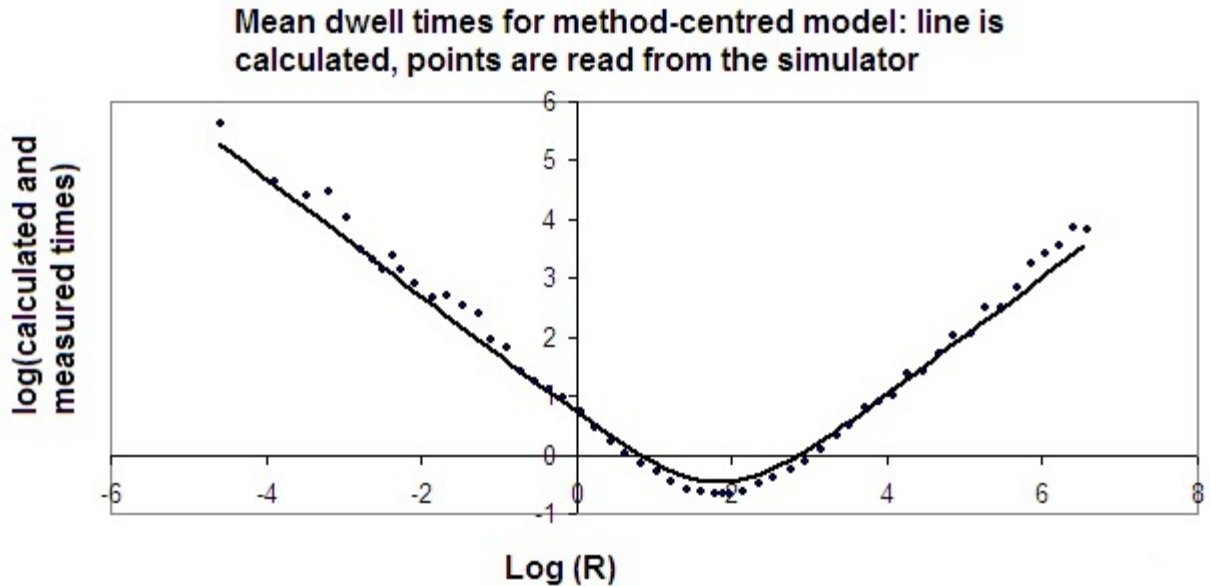
Figure 6.3: Agreement between calculated and simulated dwell times for Measurement-centred method with $p = 1$.

The calculated and measured standard deviations alse agree well, and yield a plot almost identical to figure 6.3 .

The corresponding analysis for imperfect measurements is beyond the scope of this chapter, but when measurements are sufficiently far apart to ensure that a random number of Rabi cycles elapse between one measurement and the next, both outcomes of the later measurement are equally likely. *This is not changed by a fractional success rate.* We conclude that for sparse

measurements the measurement-centred dwell time is independent of the accuracy of the measurements.

6.5 Telegraphing statistics for the state centered method with perfect measurement

Any perfect measurement of the atom will move its state to exactly one or other of the states $|0\rangle$ or $|1\rangle$. This fact simplifies the analysis.

When an atom is telegraphing, the state-centred method assumes that it dwells at or near one of its two possible energy eigenstates for extended periods before it switches to the other. This section calculates the statistics of dwells for a range of values of the parameter R, the mean measurement frequency, and compares the results with measurements made on actual (simulated) trajectories. We use the shorthand 'half-way' to label the state in which the eigenstates are equally probable.

The heart of the method consists of computing the average evolution of the system for one 'dwell' which consists of one or more 'episodes'. Our explanation is based on an 'up' dwell near $|1\rangle$, the high-energy state, but the problem is symmetrical about the two states, and we could equally well have chosen the low-energy state.

The analysis of the state-centred method is complicated because there are two different ways that an atom can change its state. On one hand, a measurement that returns the opposite value to the current dwell causes an immediate switch, but on the other hand, in the absence of measurements, the natural Rabi evolution will bring about a switch in a quarter of a Rabi cycle, as the atom passes through the half-way state - that is, the state in which a measurement would be equally likely to return either result.

In many cases, an 'up' dwell starts in the $|1\rangle$ state, directly following an 'up' measurement in the preceding 'down' dwell. However, the preceding dwell might have ended with a quarter Rabi cycle without measurements, in which case the dwell will start in the half-way state. A way to resolve this complexity is to consider the dwell as consisting of two components:

An initial segment that starts in the half-way state, followed by s final segment that starts in the eigenstate. In many cases the initial segment will be absent, and in some cases the final segment will vanish.

Figure 6.4 illustrates two types of final segment, and three types of initial segment.

1. Type 1 final segments start with an 'up' measurement from the preceding 'down' dwell. As time passes, the atom Rabi cycles, until another measurement occurs. As the state is still near $|1\rangle$, the probable result is another 'up' measurement which restores the atom to exactly the $|1\rangle$ state. The exact probability depends on how far the atom has progressed on its Rabi cycle since the last measurement. This continues until a measurement returns 'down' . This ends the dwell and starts the following 'down' dwell at its eigenvalue, without an initial segment

2. Type 2 final segments start in the same way, but ends when the gap between measurements
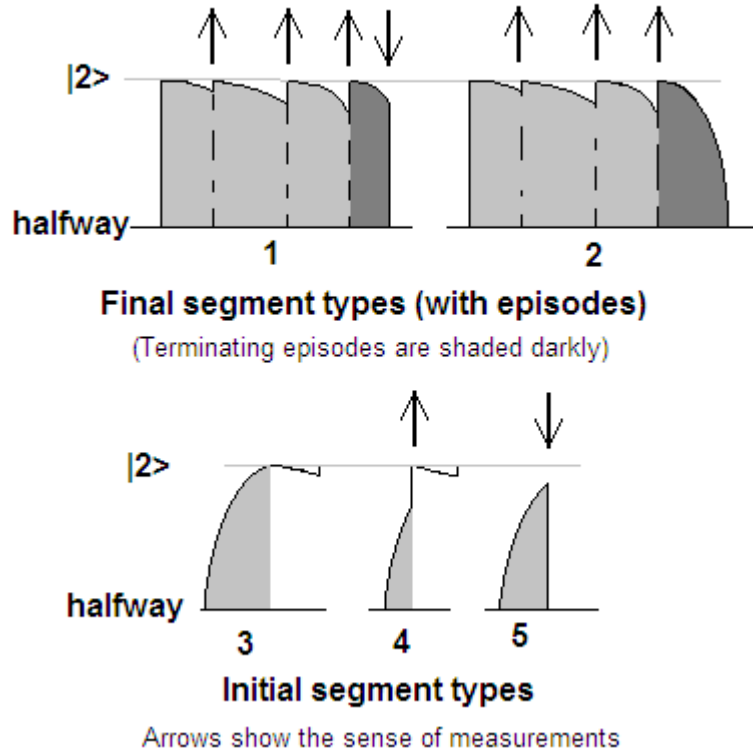
Figure 6.4: Types of segment

is quarter of a Rabi cycle, which is long enough for the atom to reach the state where the two eigenstates are equally probable. This ends the current dwell and starts the next one, with an initial segment that starts in the half-way state

3. Type 3 initial segments start in the half-way state, when the preceding 'down' dwell ends with a quarter Rabi cycle. The initial quarter Rabi cycle is undisturbed by a measurement

4. Type 4 initial segments start in the half-way state, but are interrupted by an 'up' measurement.

5. Type 5 initial segments start in the half-way state, but are interrupted by a 'down' measurement. This type of segment is *not* followed by a final segment, but leads directly to a dwell in the opposite state.

Now we introduce *episodes*, the individual components that constitute a segment. An episode starts when the system is exactly in the upper energy eigenstate, at time $t = 0$, and will evolve away from that state. The episode can end in any of three ways:

1. A measurement can return it to the initial eigenstate (the dwell continues)

91

2. A measurement can switch it to the other eigenstate (the dwell ends, and the next one starts at the eigenvalue in the opposite phase)

3. If no measurements occur, the atom will execute part of a Rabi cycle, and reach the half-way state after a quarter-cycle. (the dwell ends, and the next one starts in the opposite phase, at the half-way point).

As we shall see, the duration of an episode is variable. The length of a dwell is the sum of the durations of all the consecutive episodes that return to the initial eigenstate, *plus* the final episode.

Consider a system that is modelled by time-steps of duration δt. The probability of a measurement occurring at any step is $R\delta t$. Suppose the system starts in state $|1\rangle$, at the 'top' of the Rabi cycle.

In the absence of any measurement, the density matrix at time t is

$$\hat{\rho} = \cos^2\left(\frac{\Omega t}{2}\right)|1\rangle\langle 1| + \sin^2\left(\frac{\Omega t}{2}\right)|0\rangle\langle 0| + i\sin\left(\frac{\Omega t}{2}\right)\cos\left(\frac{\Omega t}{2}\right)(|0\rangle\langle 1| - |1\rangle\langle 0|). \qquad (6.27)$$

Each time-step can have three possible outcomes:

1. With probability $(1 - R\delta t)$, no measurement takes place. Time moves forward by one step

2. With probability $R\delta t\cos^2\left(\frac{\Omega t}{2}\right)$, a measurement takes place, and the atom returns to the state $|1\rangle$. This signals the end of this episode and the start of the next episode in the *same* dwell

3. With probability $R\delta t\sin^2\left(\frac{\Omega t}{2}\right)$ this measurement returns the value $|0\rangle$. This implies the end of this episode *and* of this dwell.

Figure 6.5 is a probability tree for such an episode. The ellipses are intermediate states reached by time-steps in which no measurements take place. Rectangles are terminal nodes . Each arc is marked with the probability that will be followed.

- The probability that the atom reaches time t without a measurement is $(1-R\delta t)^{t/\delta t} \approx e^{-Rt}$

- The probability that the atom returns to eigenstate $|1\rangle$ at time t is $e^{-Rt}\delta t\cos^2\left(\frac{\Omega t}{2}\right)$

- Similarly the probability that the atom switches to state $|0\rangle$ at time t is $e^{-Rt}\delta t\sin^2\left(\frac{\Omega t}{2}\right)$

- The duration t_n of a quarter Rabi cycle (when no measurements take place) is

$$t_n = \frac{1}{4} \times \frac{2\pi}{\Omega} = \frac{\pi}{2\Omega}. \qquad (6.28)$$

The probability p_0 that the atom reaches the end of chain, a quarter way round the Rabi cycle, without any measurements is

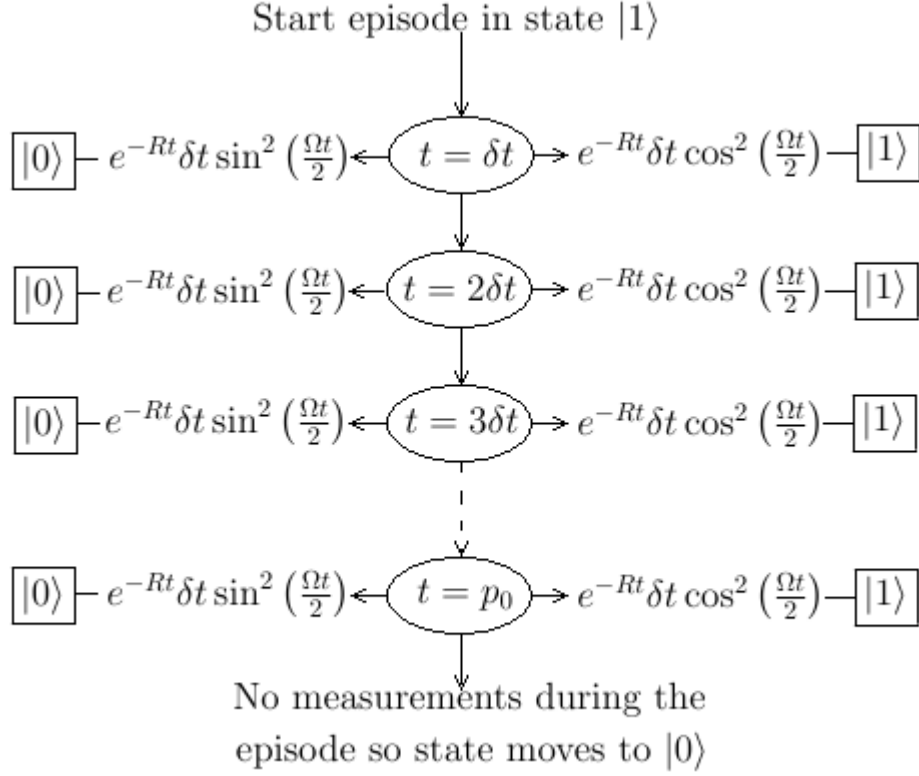$$p_0 = \exp\left(-\frac{R\pi}{2\Omega}\right). \qquad (6.29)$$

92

Start episode in state $|1\rangle$

Figure 6.5: Probability tree

The overall probability p_1 that the atom completes the episode and returns to the starting eigenstate is

$$p_1 = \sum_{t=0}^{t=t_n} \delta t\, R\cos^2(\frac{\Omega t}{2}). \tag{6.30}$$

Similarly, the overall probability p_2 that the atom completes the episode and switches to the opposite eigenstate is

$$p_2 = \sum_{t=0}^{t=t_n} \delta t\, R\sin^2(\frac{\Omega t}{2}). \tag{6.31}$$

As $\delta t \to 0$ these relations can be replaced by

$$p_1 = \int_0^{t_n} dt\, R\cos^2(\frac{\Omega t}{2})$$

$$p_2 = \int_0^{t_n} dt\, R\sin^2(\frac{\Omega t}{2}). \tag{6.32}$$

The *average length* A of each episode that ends with an 'up' measurement can be found by summing the probabilities of exit at each stage, multiplied by the time of exit. As this type of episode is not the only possible outcome of the evolution, we must divide by p_1, the probability

of the outcome. We may put:

$$A = \int_0^{t_n} dt\, tR \cos^2\left(\frac{\Omega t}{2}\right)/p_1. \qquad (6.33)$$

Likewise, the average length of an episode that ends with a 'down' measurement is

$$B = \int_0^{t_n} dt\, tR \sin^2\left(\frac{\Omega t}{2}\right)/p_2. \qquad (6.34)$$

Similar adjustments must be made to all calculations of average episode length.

As p_1 is a probability ($p_1 \le 1$), d_1, the total length of a type 1 segment, is:

$$d_1 = A(1 + p_1 + p_1^2 + p_1^3 + \ldots) + B = A \sum_{n=0}^{\infty} p_1^n + B = \frac{A}{1 - p_1} + B. \qquad (6.35)$$

Similarly, $d2$, the length of a type 2 segment is

$$d_2 = \frac{A}{1 - p_1} + t_n \qquad (6.36)$$

The statistics of initial segments can be calculated in a similar way to those of final segments, except that:

- Each initial segment consists of only one episode

- The segment starts at the half-way state.

The following equations serve. p_3 to p_5 are probabilities, and d_3 to d_5, expected durations.

p_3(segment reaches the eigenvalue state (type 3))	e^{-t_n}
p_4(segment ends with an 'up' measurement(type 4))	$\int_0^{t_n} dt\, e^{-Rt} \cos^2(\frac{\Omega t}{2} - \pi t_n)$
p_5(segment ends with a 'down' measurement (type 5))	$\int_0^{t_n} dt\, e^{-Rt} \sin^2(\frac{\Omega t}{2} - \pi t_n)$
d_3(duration of a type 3 segment)	0.25 Rabi cycle $= t_n$
d_4(duration of a type 4 segment)	$\int_0^{t_n} dt\, t\, e^{-Rt} \cos^2(\frac{\Omega t}{2} - \pi t_n)/p_4$
d_5(duration of a type 5 segment)	$\int_0^{t_n} dt\, t\, e^{-Rt} \sin^2(\frac{\Omega t}{2} - \pi t_n)/p_5$

Next we consider the average duration of *dwells* (as opposed to segments). There are two patterns of dwell that start at the eigenvalue, and no fewer than six that start with an initial segment. Figure 6.6 is a taxonomy of these different dwell types. Each arc of the graph is labelled with the probability that the arc is traversed, and each dwell type is marked with its duration.

The average dwell length for each of many different values of R was calculated by summing the durations of each of the dwell types, multiplied by the probability of their occurrence. Where the initial segment ends with a 'down' measurement, the calculation accounts for *two* dwells (one of each phase), so the contribution from this type of dwell is halved.

Statistics of measurement-centred dwells

For this part of the investigation the simulator was adjusted to detect steps in which the probability of one of the energy eigenstates switched from > 0.5 to < 0.5. A run with a given value
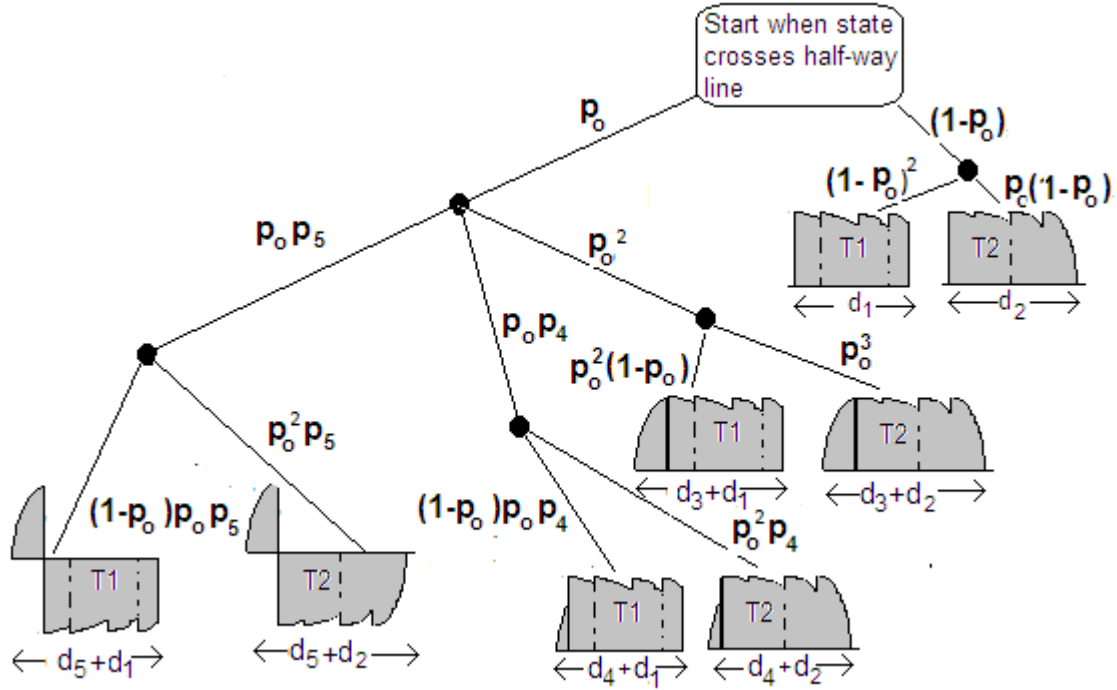
Figure 6.6: Taxonomy of dwell types

of R took 1000 Rabi cycles or 10^6 steps. This produced a number of dwells varying greatly in their length, but a roughly constant factor was the standard deviation of the lengths was close to their average. This is consistent with a negative exponential distribution, as might be expected and is typical of a random telegraph [71]. The *average* dwell times delivered by the separate runs were close to one another.

The graph in Figure 6.7 shows that for R values higher than 10 measurements per Rabi cycle, the relationship between R and the dwell time is almost linear. In this region, the end effects due to long gaps between measurements are negligible. When measurements are frequent, most of them occur when the state of the atom is close to its eigenstate. Here the density matrix is closely approximated by

$$\hat{\rho} = \left(1 - \left(\frac{\Omega t}{2}\right)^2\right)|1\rangle\langle 1| + \left(\frac{\Omega t}{2}\right)^2 |0\rangle\langle 0| + \left(\frac{i\Omega t}{2}\right)(|0\rangle\langle 1| - |1\rangle\langle 0|) \tag{6.37}$$

and the probability of a quarter Rabi cycle without any measurements is negligibly small. With these approximations, the average episode duration is *inversely* proportional to R, the frequency of measurement, but the probability of an episode ending with a 'down' measurement is inversely proportional to R^2. It follows that the dwell time increases linearly with R.
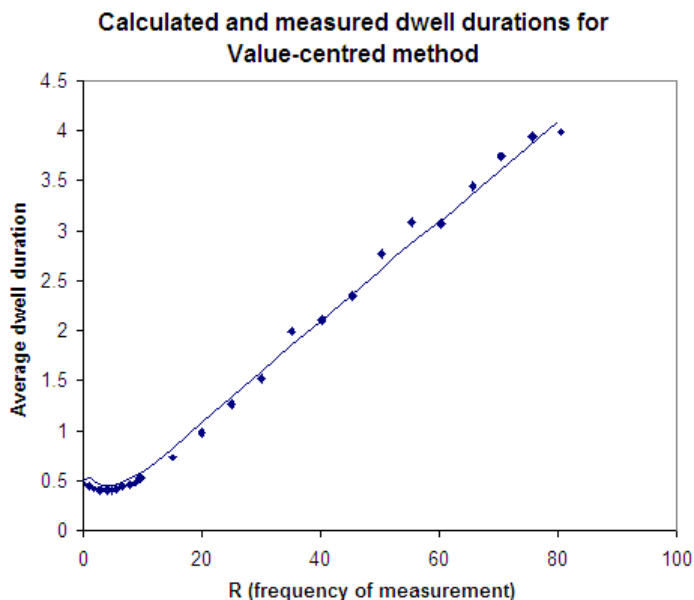
Figure 6.7: Measured and calculated dwell times for the Value-centred model

6.6 Observations of dynamics with error-prone measurements

Theoretical analysis of this regime is beyond the scope of this paper, but we offer some observations derived from the simulator.

6.6.1 Effect of imperfect measurements on dwell times, using the state-centred model

It is to be expected that when measurements are imperfect, any incorrect measurement is likely to force a premature change of the dwell state, and this is borne out by experiment. We collected statistics for three measurement frequencies (in terms of measurements per Rabi cycle): $R = 20$, $R = 50$, and $R = 80$.

The results are shown in figure 6.8, which illustrates the relationship between the error rate and the dwell time for each of the three R values. Error rates in measurement were varied between 0 and 0.5. Any error rate is likely to cause the dwell time to drop sharply, in an (approximately) exponential decay.

6.6.2 Effect of imperfect measurements on the telegraphing index.

The Telegraphing Index for runs with imperfect measurements are shown in figure 6.9, again using three values of R. For $R = 20$ the drop is almost linear, but for higher values of R, a high degree of telegraphing is maintained even for moderately high error rates.

96
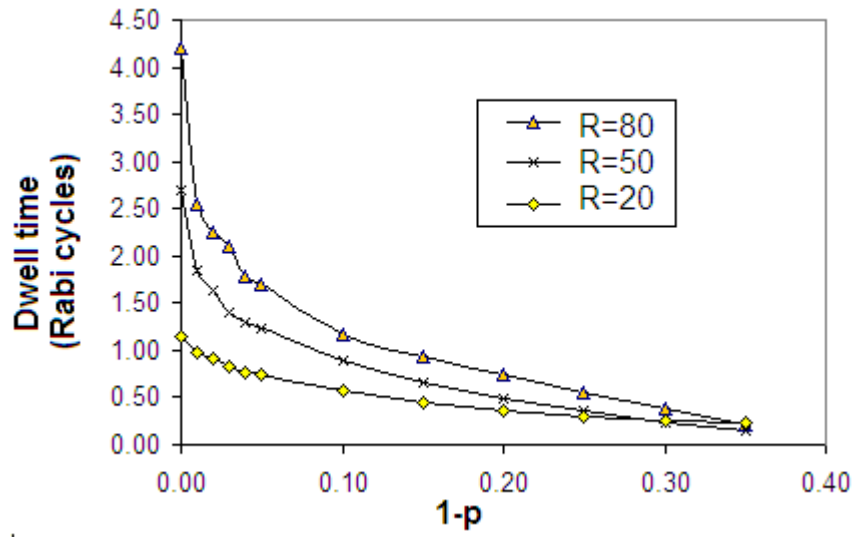
Figure 6.8: Relationship of dwell times and error rate in measurement
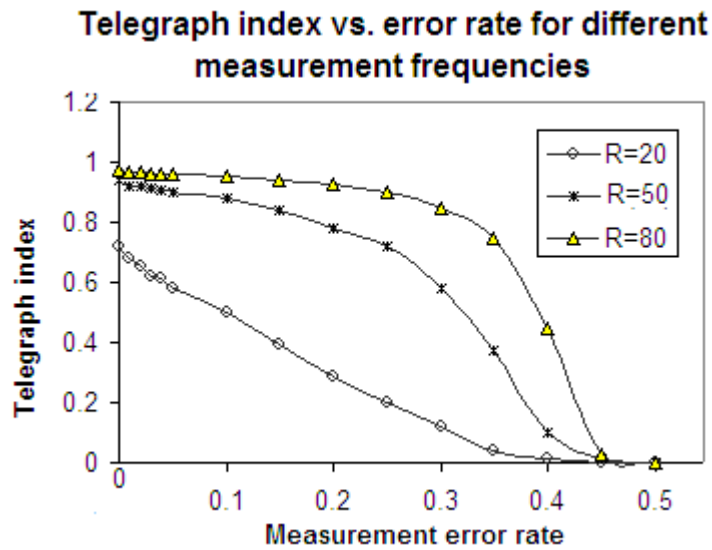
Figure 6.9: Relationship of Telegraph Index and R

6.6.3   Dwell state switching under imperfect measurement

Equation 6.7 shows that the state of an atom after an imperfect measurement is not in general either of the states $|0\rangle$ or $|1\rangle$, but rather a superposition of both. When $p = 1$ each measurement gives an abrupt change to an eigenvalue, but when $p < 1$ this is no longer the case, and switching between dwell states can take several measurements. As each measurement occurs when the

atom is in an intermediate state, the outcome is indeterminate, and the state can move in either direction. This is illustrated in figure 5.4.

As well as generating this trajectory, the simulator also produced the corresponding measurement record. Figure 6.10 is a much-expanded version of figure 5.4, which shows the relationship of measurements and the state of the atom. A copy of figure 5.4 is included for comparison.
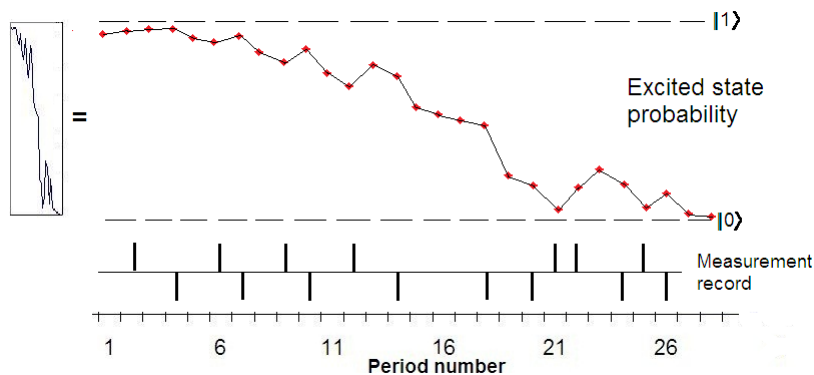
Figure 6.10: Imperfect measurements and the state of the atom during switching, using the state-centred model

The measurement record (alone) can also be used to find the dwell state of the atom according to the measurement-centred model. This is illustrated in figure 6.11.
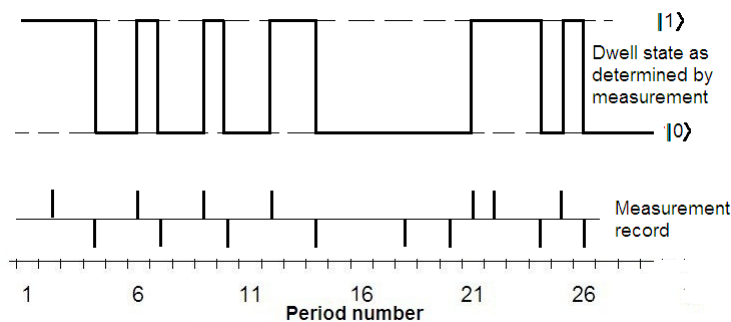
Figure 6.11: Dwell state of the atom during switching, using the measurement-centred model

6.7   Conclusion

The concept of frequent measurement is a useful tool for investigating the interaction of an atom with its environment. The general effect of the interaction is to exert a 'drag' on the closed evolution of the atom, so that it remains at or near one of its eigenstates longer than would otherwise be the case. We have suggested a 'telegraph index' as a measure for this drag, and offered two

possible definitions of telegraphing. Assuming perfect measurements, we have made theoretical analyses of both definitions of telegraphing. Measurements made with a simulator agree closely with those found by theoretical analyses. We found that when measurements are frequent the statistics associated with each definition coincide, but when they are sparse the statistics diverge.

Turning to imperfect measurements (such are defined in [55]) , we have no theoretical analysis, but we present graphs produced by the simulator. A general observation is that imperfect measurements make the atom much less stable; telegraphing times are reduced, the telegraph index generally drops, and switching between dwell states takes longer, with the state occupying an intermediate position in which the next measurement could easily be in either direction. During the switching process the measurement-centred model shows the dwell-state values oscillating, at a frequency similar to the frequency of measurements.

Chapter 7

Conclusion

"When I am in the company of scientists, I feel like a shabby curate who has strayed by mistake into a drawing room full of dukes."
W. H. Auden

With this thesis, my supervisors and I have made contributions to Physics in the general area of quantum measurement. Within this area, we have addressed two unrelated subjects:

- Programed discrimination with known classical relationships between the program qbits

- A quantitative analysis of the effect of frequent, but sometimes imperfect measurement of an otherwise closed quantum system

In both cases the research relied on certain key papers - [42, 55] as starting points. These papers were so important to our own work that we have included summaries of them in Chapters 2 and 5.

7.1   Programmed discrimination

In our first published paper we examined *triads* consisting of two program qbits $|\psi_1\rangle$ and $|\psi_3\rangle$, and one data qbit $|\psi_2\rangle$. It is given that the data qbit is identical to one or other of the program qbits, thus being a representation of a classical binary digit. There is a known *a priori* probability η_1 that $|\psi_2\rangle \equiv |\psi_1\rangle$. We investigated three kinds of classical knowledge, namely

- The program qbits have a known, fixed overlap

- The program qbits are confined to a known great circle on the Bloch spheres

- The program qbits are confined to polar caps of known size

In all three cases we looked at error rates in **optimum recognition**, and failure rates in **unambiguous discrimination**, and took into account varying values of the *a priori* probability η_1. The results confirmed the general principle that more information yields better recognition rates, but the effect was small. Another feature which emerged clearly was that unambiguous recognition can only be used in the central range of *a priori* probabilities. Outside this range the

best recognition rate is given by a Von Neuman measurement of the most likely configuration, but this means that the other, less probable configuration can *never* be detected. This could be a serious drawback in some circumstances; but in mitigation we note that in all conventional data storage and transmission techniques which handle binary digits (whatever their representation) the best efficiency is achieved when the numbers of zeros and ones are statistically equal.

In the second paper we extended the research to the case where the data qbit, and both the program qbits could be supplied in multiple copies. A configuration with n qbits in total (no matter how distributed) will lead to density matrices of order 2^{2n}. The procedure to set up the matrices requires a complicated algorithm in its own right, and the computer time needed to handle these matrices increases exponentially with n. The largest configuration we looked at had 13 qbits (this took an all-night run on a standard IBM PC).

With an arbitrary number of qbits, as well as different types of classical knowledge and recognition methods, the number of potential cases which can be investigated is very large. In our paper we only present a few of them, but we have published the relevant programs so that interested readers can pursue this subject on their own account [1].

Again, we found that the added information improved the recognition rate, but not a great deal. If costs are proportional to the number of qbits involved, then with certain exceptions, using multiple copies of qbits is not cost-effective.

The exceptions arise in the context of data transmission. The *triad* representation offers an advantage over other qbit transmission techniques, because the relationship between the three qbits is not destroyed when the qbits undergo unitary transformations (as long as they affect all three qbits equally).

This argument is only valid for the 'known overlap' case, as in the other two an arbitrary unitary transformation could move the qbits away from the known great circle, or out of the polar caps. Furthermore we take it that the *a priori* probability is exactly $\frac{1}{2}$, a universal rule for data transmission engineers.

With optimal recognition, the error rate in all cases is less than 0.5. Shannon's noisy channel theorem implies Alice can get information through to Bob reliably, although at a much lower rate. With orthogonal program qbits, the best possible rate is about 41 qbits per (reliable) bit. This is achieved with doubled program qbits and a doubled data qbit, but other configurations with small numbers of qbits yield similar figures.

With unambiguous discrimination (where there are failures but no errors) Shannon's theorem is not applicable. It is not a feasible way to transmit actual messages, but the method can be used for Alice and Bob to build up a key of random binary digits which they can then use to code a real message for Alice to transmit by classical means.

Alice records and sends Bob a long sequence of randomly coded triads. Bob only recognises some of them, but keeps a record of successful recognitions. When the sequence is ended, he contacts Alice over a classical channel and informs her which triads were recognised. The result is a long string of random digits that Alice can use as a key.

---

[1] These programs can be found on the CD which is part of this thesis.

The best configuration for such a transmission is just two program qbts and a doubled-up data qbit. This gives a rate of 8 qbits per (successful) bit transmitted; substantially better than the case for optimal recognition.

There is good reason to believe that in the future programmed discrimination will be a valuable technique with quantum-based logic devices (including computers).

7.2 Repeated measurement

Frequent measurement of closed quantum systems is interesting because it is a model for interaction with the outside world. As such, it has been used to investigate friction and could be used to examine many other types of interaction.

In our paper, which is presently in preparation and reported in Chapter 6, we examine the phenomenon of telegraphing. We propose a numerical measure for this phenomenon, and suggest two different mathematical models. In both cases, the predictions of the models coincide closely with observations made on a simulator of a closed system under frequent measurement.

An interesting feature of imperfect measurement is that unlike a good measurement, which always returns a system to one of its eigenvalues and destroys its previous state, a poor measurement, with a significant chance of yielding the wrong result, will only change the state of the system slightly.

7.3 Possible future work

In the area of programmed discrimination, a possible topic for further investigation is recognition when the basic constraint - that is, that the data qbit should be *exactly* the same as one of the program qbits - no longer applies. What if the condition specifies a minimum overlap instead?

Another potential investigation might be to examine the behaviour of groups made of qdits. For example, a group could consist of a program qdit with four possible states and four program qdits.

A topic that interests me greatly - although it has little to do with the topics of this thesis - is the automatic design of quantum discriminators working under various conditions. The problem could be addressed using various AI techniques such as genetic algorithms.

THE END

Appendix A

A.1   The Readme File

```
Read Me:
==== ==

Programs for
======== ===

Programmed discrimination of multiple sets of qbits with added
========== ============== == ======== ==== == ===== ==== =====
classical information
======== ===========


Introduction
============

This is not a commercial product, and we do not offer a "Wizard" to set the
system up for you. To use these programs effectively, you should have some
familiarity with

a) A Text editor such as EMACS or TextPad. Do not use Microsoft Word,
although Microsoft WordPad would serve.

b) Using a command line processor, which would be an emulated terminal
under UNIX and its descendants LINUX or UBUNTU, or the program "cmd.exe"
under Windows 7. On a PC running under Windows you will generally find this
program at c:\Windows\System32\cmd.exe.
```

c)   Compiling C programs with gcc (or another C/C++ compiler) and running them. We assume that the compiler has been downloaded to your computer.

To summarise, you should at least know how to enter, compile and run the traditional "Hello World!" program in C. If necessary, you should ask a colleague for help.

The programs have been tested on an IBM PC under Windows 7, using both the GNU and Borland compilers, and under UBUNTU using GNU.

The work reported in this paper was done with six C programs, whose functions are evident from their names:

fixedoverlapbest.c
fixedoverlapunam.c
greatcirclebest.c
greatcircleunam.c
polarcapbest.c
polarcapunam.c

The programs are provided in source form, so that they can be used on a variety of different computers and operating systems.

Setup
=====

1. Make a new directory (suggested name: "Discrimination"). Under UNIX and its variants LINUX and UBUNTU, the directory should be in your home directory. For Windows it can be in any convenient place,

2. You have already unzipped the file you have down-loaded (otherwise you would not be reading this). Copy all the files into the new directory.

3 Open the command line processor or a simulated terminal
(if you have not already done so),

4. Move to the new directory, using a 'cd' command.

5. To compile the first program, enter

```
gcc fixedoverlapbest.c -lm -o fixedoverlapbest.exe
```

This assumes that the GNU C/C++ compiler has been loaded. If necessary you can use another C/C++ compiler such as the one issued by Borland. In this case the command line would read

```
bcc32 fixedoverlapbest.c
```

6. Do the same for the other programs (now or later), using their names instead of "fixedoverlapbest.c".

Executing the programs:
========= === =========

To run the first program, enter the command line

```
   fixedoverlapbest.exe  a  b  c
```

where a,b and c stand for small whole numbers: For example

```
   fixedoverlapbest.exe  3 2 1
```

Here a is the number of program qbits on the "left";
     b is the number of data qbits;
     c is the number of data qbits on the 'right'.

The program displays a message like

```
    Output will be sent to:
         fixedoverlapbest(3,2,1).txt
```

This file will be in the same directory as the programs.

As the program runs, it displays a * from time to time to reassure you of progress.

When the program ends, it displays a message like

```
Elapsed time = 314 seconds
```

The other programs in suite are run in the same way,
but using their names instead of "fixedoverlapbest" .

Timing
======

The time taken to analyse any configuration increases more than
exponentially with the total number of qbits. It also varies
with the number of rows in your table, and doubles when the
input parameters are not symmetrical, so that (1,2,3) takes at
least twice as long as (2,2,2).

The programs for unambiguous discrimination take much longer
than the ones for best discrimination.

We strongly suggest that you start with small configurations
(such as (1,1,1)) and work up. with the total number of qbits.

Modification
============

You can make various changes by editing the configuration.c document
in your directory.

Formats
=======

The default format of the output is a table like this:

Fixed Overlap Optimum {1,1,3}

Beta	pi/6	pi/3	pi/2	2pi/3	5pi/6	pi

Eta

Eta						
0.0000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
0.2500	0.23995	0.21250	0.17500	0.13750	0.11005	0.10000
0.5000	0.47354	0.40126	0.30253	0.20379	0.13151	0.10505
0.7500	0.23719	0.20219	0.15439	0.10658	0.07158	0.05877
1.0000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

This can be loaded into a spreadsheet for further analysis and drawing graphs

You can alter the number of decimal places in the table by
changing the line

```
    #define DECIMALS   5
```

appropriately. For example,

```
    #define DECIMALS  8
```

will generate tables with 8 decimal places.

If you edit the source text by inserting the line

```
#define LATEX
```

the output will be in a format suitable for inclusion in a
Latex document:

```
\begin{table}[h]
  \begin{center}    \begin {tabular}{|l|cccccc|}\hline
$\beta$     & $\pi$/6 & $\pi$/3 & $\pi$/2 &2$\pi$/3 &5$\pi$/6 & $\pi$  \\
\hline
$\eta$&&&&&&\\
0.00    &0.000 &0.000 &0.000 &0.000 &0.000 &0.000\\
0.25    &0.240 &0.212 &0.175 &0.137 &0.110 &0.100\\
0.50    &0.474 &0.401 &0.303 &0.204 &0.132 &0.105\\
0.75    &0.237 &0.202 &0.154 &0.107 &0.072 &0.059\\
1.00    &0.000 &0.000 &0.000 &0.000 &0.000 &0.000\\
\hline
\end{tabular}
\caption{Fixed Overlap Optimum \{1,1,3\}}
\end{center}
\end{table}
```

The number of decimal places in the table is controlled by the line

```
#define LATEXDECIMALS 3
```

Changing the interval between values of eta
======== === ======== ======= ====== == ===

Suppose you want the interval between successive values of eta to be 0.05.
Find the line which says,

#define ETAINTERVAL 0.25

and change it to

#define ETAINTERVAL 0.05

It is important that the interval be a sub-multiple of 1

Changing the interval between values of beta in the two
======== === ======== ======= ====== == ==== == === ===
fixed-overlap programs
===== ======= ========

Suppose you want the interval between successive values of beta to be PI/8.
Find the line which says,

#define BETAINTERVAL (PI/6)

and change it to

#define BETAINTERVAL (PI/8)

It is important that the interval be a sub-multiple of PI. The round brackets
are essential.

Changing the interval between values of theta in the two
======== === ======== ======= ====== == ===== == === ===
polar caps programs
===== ==== ========

Suppose you want the interval between successive values of theta to be PI/4.
Find the line which says,

#define THETAINTERVAL (PI/6)

and change it to

```
#define THETAINTERVAL (PI/4)
```

It is important that the interval be a sub-multiple of PI. The round brackets are essential.

Important
=========

When you have changed the configuration file you MUST recompile the programs!

Help
====

If you have difficulties which cannot be solved, even with expert help, send an email to Andrew at andrew@crm.scotnet.co.uk

A.2   Contents of the CD ROM

- The complete text of this thesis (in PDF format)

- PDF version of *Programmed discrimination of qbits with added classical information*
  (A.J.T. Colin, S.M. Barnett and J. Jeffers), Eur.Phys J.D. **63**, 463-472 (2011)

- PDF version of *Programmed discrimination of multiple sets of qbits with added classical information*
  (A.J.T. Colin), Eur.Phys J.D.**66** 185 (2012)

- PDF version of *Measurement-driven dynamics for a coherently-excited atom*
  A.J.T. Colin, S.M. Barnett and J. Jeffers (in preparation)

- A zip file containing

  - All six programs mentioned in Chapter 4 (in C)

  - A library of useful functions (in C)

  - The 'ReadMe' file

Bibliography

[1] P.A.M. Dirac,
A new notation for quantum mechanics
Mathematical Proceedings of the Cambridge Philosophical Society 35 (3): pp. 416-418. (1939).

[2] W.H. Press, B.P. Flannery, S.A Teukolsky and W.T. Vetterling,
Numerical Recipes in C,
Cambridge University Press,(1990).

[3] J. Bolton and R. Lambourne (eds.)
The Quantum World (book 1): Wave Mechanics
The Open University (SM358).

[4] F. Bloch,
Nuclear induction
Phys. Rev. 70(7-8) (1946).

[5] J.S. Bell,
On the Einstein-Podolski-Rosen Paradox
Phys. Rev. A, 54: 3824 (1996).

[6] R.P. Feynman,
Electrons and Their Interactions (QED: The Strange Theory of Light and Matter).
Princeton University Press. p. 115 (1985)

[7] R. Loudon,
The Quantum Theory of Light
Oxford University Press, (2000).

[8] W. Gerlach and O. Stern,
Das magnetische Moment des Silberatoms
Zeitschrift fr Physik 9: 353355 (1922).

[9] J. Bolton and R. Lambourne (eds.)
The Quantum World (book 2): Quantum Mechanics and its interpretation
The Open University (SM358).

[10] R. Penrose,
   The Road to Reality (Chapter 29),
   Vintage Books, London (2004).

[11] J. von Neumann,
   Mathematical Foundations of Quantum Mechanics
   (Princeton University Press 1957).

[12] S.M. Barnett and S. Croke,
   Quantum state discrimination
   Advances in Optics and Photonics, Vol 1 238-278 (2009).

[13] S.M. Barnett,
   Quantum Information
   Oxford University Press (2009).

[14] H. Dehmelt,
   Laser Fluorescence Spectroscopy on Tl+ Mono-Ion Oscillator II (spontaneous quantum jumps),
   Bull. Am. Phys. Soc. 20, 60 (1975).

[15] I. M. Gelfand, M. A. Naimark,
   On the imbedding of normed rings into the ring of operators on a Hilbert space.
   Math. Sbornik 12 (2): 197217. (1943).

[16] B. Coecke and E.O. Paquette,
   POVMs and Naimark's theorem without sums
   Electronic Notes in Theoretical Computer Science, Vol 210 15-31 (2008).

[17] C.W. Helstrom
   Quantum detection and estimation theory
   Academic Press, New York, (1976).

[18] J. Bergou, U.Herzog and M. Hillery
   Discrimination of Quantum States
   Chapter 11 in "Quantum State Estimation" (eds. M.Paris and J. Rehajek)
   Springer, Berlin (2006).

[19] I.D. Ivanovic,
   How to differentiate between non-orthogonal states
   Phys. Lett. A 123, 257 (1987).

[20] D. Dieks,
   Overlap and distinguishability of quantum states
   Phys. Lett. A 126, 303 (1988).

[21] A. Peres,
   How to differentiate between non-orthogonal states
   Phys. Lett. A 128, 19 (1988).

[22] A. Chefles and S.M. Barnett,
Optimum unambiguous discrimination between linearly independent symmetric states.
Physics Letters A, 250 (4). pp. 223-229.(1998).

[23] S.M. Barnett and E. Riis,
Experimental demonstration of Polarisation Discrimination at the Helstrom Bound
Journal of Modern Optics 44, 1061-164 (1997).

[24] R.B.Clarke, A. Chefles, S.M. Barnett and E. Riis,
Experimental Demonstration of Optimal Unambiguous State Discrimination
Phys. Rev. A, 64,040305 (2001).

[25] W.K. Wootters and W.H. Zurek,
A Single Quantum Cannot be Cloned,
Nature 299,802 (1982).

[26] L. Mach and L. Zehnder,
Instrumentenkunde, 11-12 (1892).

[27] N.D. Mermin,
Quantum Computer Science
Cambridge University Press (2007).

[28] G. Lindblad,
Non-Equilibrium Entropy and Irreversibility
Delta Reidel. Dordrecht: . ISBN 1-4020-0320-X (1983).

[29] J. Salo, S. Stenholm, G. Kurizki and A.G.Kofman,
The varieties of master equations,
"Decoherence, Entanglement and Information Protection in Complex Quantum Systems".
NATP Science Series, Vol 189, Springer, (2005).

[30] M.S. Kim and P.L. Knight,
Quantum-jump telegraph noise and macroscopic intensity fluctuations
Physical Review A, 5265-5270 (1987).

[31] C. H. Bennett and G. Brassard,
Quantum Cryptography: Public key distribution and coin tossing
Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing
Bangalore, p. 175 (1984)

[32] L.K. Grover,
A fast quantum mechanical algorithm for database search
Proceedings, 28th Annual ACM Symposium on the Theory of Computing p. 212, (May 1996).

[33] P. W. Shor,
Algorithms for quantum computation: Discrete logarithms and factoring
Proc. 35nd Annual Symposium on Foundations of Computer Science,
IEEE Computer Society Press , 124-134 (1994).

[34] L. Rivest, A. Shamir, and L.M. Adleman,
Public Key Encryption U.S. Patent 4,405,829 (1977).

[35] R. Feynman,
Simulating Physics with Computers
International Journal of Theoretical Physics 21 (67): 467488, 1988

[36] E. Morrison and P. Morrison,
Charles Babbage's Papers
Dover Publications, New York (1962).

[37] D. Hanneke, J.P. Home, J.D. Jost, J.M. Amini, D. Leibfried and D.J. Wineland,
Realization of a two-qubit quantum processor
Nature Physics DOI: 10.1038/nphys1453 (2009).

[38] G. Jaeger and A. Shimony,
Optimal distinction between two non-orthogonal quantum states
Physics Letters A 197, 83-87 (1995).

[39] S.M. Barnett, A. Chefles and I. Jex
Comparison of two unknown pure quantum states
Physics Letters A 307 189-195 (2003).

[40] J. Bergou and M. Hillery
Universal Programmable Quantum State Discriminator that is Optimal for Unambiguously Distinguishing between Unknown States
Physical Review Letters 94, 160501 (2005).

[41] A. Hayashi, M. Horibe and T. Hashimoto
Unambiguous pure-state identification without classical knowledge
Physical Review A 73, 012328 (2006).

[42] J. Bergou, V. Bužek, E. Feldman, U. Herzog and M. Hillery
Programmable quantum-state discriminators with simple programs
Physical Review A 73, 062334 (2006).

[43] J.Bergou
Quantum state discrimination and selected applications
Journal of Physics: Conference Series 84 012001 (2007).

[44] C. Bennett,
Quantum cryptography using any two nonorthogaonal states,
Physical Review Letters f 68, 3121 (1992).

[45] U. Herzog and J. Bergou
Optimum unambiguous identification of unknown pure qudit states
Physical Review A 78, 032320 (2008).

[46] M. Hillery, E. Andersson, S. Barnett and D. Oi
Decision problems with black boxes
Journal of Modern Optics 53, 244-252 (2010).

[47] A.J.T. Colin, S. Barnett and J. Jeffers
Programmed discrimination of qbits with added classical information
European Physical Journal D. **63**, 463-472 (2011).

[48] A.J.T. Colin
Programmed discrimination of multiple sets of qbits with added classical information
European Physical Journal D.**66** 185 (2012).

[49] I. Jex, G. Alber, S.M. Barnett and A. Delgado,
Antisymmetric multi-partite quantum states and their applications,
Fortschr Phys.51 No 2-3,172-178/DOI 10.1002/prop.200310021

[50] U. Herzog and J.A. Bergou,
Optimum measurement for unambiguously discriminating two mixed states: General considerations and special cases,
Journal of Physics Conference Series **36** 49-54 ,(2006).

[51] A.J.T. Colin,
Multiple qbit discrimination programs,

   `www.phys.strath.ac.uk/cnqo/?page_id=450`

[52] A.K. Ekert, J.G. Rarity, P.R. Tapster, and G.M. Palma,
Physical Review Letters **84**,4729, (2000).

[53] C. E. Shannon and W. Weaver,
The Mathematical Theory of Communication. Urbana,
IL:University of Illinois Press, (1949, reprinted 1998).

[54] B. Schneier,
Applied Cryptography
John Wiley and Sons, (1994).

[55] J.D. Cresser, S.M. Barnett, J. Jeffers and D.T.Pegg
Measurement Master Equation
Optics Communications 264 352-361, (2006).

[56] A.J.T. Colin, S.M. Barnett and J. Jeffers
Measurement-driven dynamics for a coherently-excited atom
Journal of Modern Optics
Published on-line at
tandfonline.com/doi/abs/10.1080/09500340.2012.744479 (2012)

[57] T.M. Tong, J-L Chen, L.C.Kwek and C.H. Oh
Krauss representation of density operator of arbitrary open qubit system
arXiv:quant-ph/0311091v2 (2003).

At this point we list books and articles which are not referenced in the text but which supplied the general background to the research.

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

[59] S.M. Barnett and P.M.Radmore,
Methods in Theoretical Quantum Optics
Clarendon Press, Oxford, (1997).

[60] S. Croke, S.M.Barnett and S. Stenholm
Linear transformations of quantum states
Annals of Physics 323 893-906, (2008).

[61] S. Kryszewski and J. Czechowska-Kryszk
Master equation - tutorial approach
Lecture note published on line :arXiv:0801.1757v1 (2008).

[62] M. Sedlak, M. Ziman, V. Bužek and M. Hillery,
Unambiguous comparison of ensembles of quantum states,
Physical Review A **77**, 042304 (2008).

[63] A. Hayashi, T. Hashimoto and M. Horibe,
Reexamination of optimal quantum state estimation of pure states,
Physical Review A **72**, 032325 (2005).

[64] B. He and J.A. Bergou,
Programmable unknown quantum-state discriminators with multiple copies of program and data: A Jordan basis approach,
Physical Review A **75** 032316 (2007).

[65] G. Sentis, E. Bagan, J. Calsamiglia and R. Muñoz-Topia,
Multi-copy programmable discrimination of general qubit states,
Physical Review A **82** 042312 (2010).

117

[66] B.W.Kernighan and D.M.Ritchie,
The C programming language (2nd edition)
Prentice Hall, (1988).

[67] B. Stroustrup,
The C++ programming language (2nd edition)
Addison Wesley, (1991).

[68] R.P. Feynman, R.B. Leighton and M. Sands,
The Feynman Lectures on Physics (vols 1-3)
Addison Wesley, (1979).

[69] J.F. James,
A Student's Guide to Fourier Transforms
Cambridge University Press, (2004).

[70] C.G. Lambe,
Elements of Statistics
Longmans, Green and Co., (1952).

[71] D.K.C. macDonald,
Noise and Fluctuations; an introduction
John Wiley and Sons, (1962).

[72] A. Whitaker,
Bohr, Einstein and the Quantum Dilemma
Cambridge University Press, (2006).

[73] B. Bellomo, S.M.Barnett and J. Jeffers
Frictional quantum dependence,
J.Phys A: Math. Theory. 40 9437-9453 (2007).

[74] P. Cockshott, L.M. Mackenzie and G. Michaelson,
Computation and its limits
Oxford University Press, (2012).

[75] D.F. Griffiths and D.J. Higham,
Learning LATEX
Society for Industrial and Applied Mathematics (SIAM), (1997).

[76] H. Carmichael
An Open Approach to Quantum Optics
Springer-Verlag (1993).

[77] S. Croke,
Maximum Confidence Measurements
Ph. D. Thesis, Strathclyde University, (2007).

[78] S.M.Barnett and S. Croke
On the conditions for discrimination between quantum states with minimum error
Journal of Physics (A) **42** 062001 (2009).

[79] C. Cohen-Tannoudji and J. Dalibard
Single-Atom Laser Spectroscopy. Looking for Dark Periods in Fluorescence Light
Europhysics Letters **19** (1986).

[80] S.M. Barnett and J.D. Cresser
Quantum theory of friction
Physical Review A **72** 022107 (2005).

[81] R.J. Cook and H.J. Kimble
Possibility of Direct Observation of Quantum Jumps
Physical Review Letters Vol 54, no.10 (1985).

[82] J.A.Bergou, E. Feldman and M. Hillery
Optimal unambiguous discrimination of two subspaces as a case in mixed-state discrimination
Physical Review A **73**, 032107 (2006).

[83] M.J. Everitt, W.J. Munro and T.P. Spiller
Quantum measurement with chaotic apparatus
Physics Letters A 374 2809-2815 (2010).

[84] A. Chefles
Quantum State Discrimination
Lecture not published on-line: arXiv:quant-ph/0010114v1 31 (2000).