UNIVERSITY OF GLASGOW

Optimal discrimination of quantum states

by

Graeme Weir

A thesis submitted in partial fulfillment for the
degree of Doctor of Philosophy

in the
College Of Science And Engineering
School Of Physics And Astronomy

June 2018

Declaration of Authorship

I, Graeme Weir, declare that this thesis titled, ‘Optimal discrimination of quantum states’ and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: ____

Date: ____

“It’s a magical world, Hobbes, ol’ buddy... let’s go exploring!” - Calvin

Bill Watterson

UNIVERSITY OF GLASGOW

Abstract

College Of Science And Engineering

School Of Physics And Astronomy

Doctor of Philosophy

by Graeme Weir

Quantum state discrimination is a fundamental task in the field of quantum communication and quantum information theory. Unless the states to be discriminated are mutually orthogonal, there will be some error in any attempt to determine which state was sent. Several strategies to optimally discriminate between quantum states exist, each maximising some figure of merit. In this thesis we mainly investigate the minimum-error strategy, in which the probability of correctly guessing the signal state is maximised. We introduce a method for constructing the optimal Positive-Operator Valued Measure (POVM) for this figure of merit, which is applicable for arbitrary states and arbitrary prior probabilities. We then use this method to solve minimum-error state discrimination for the so-called trine states with arbitrary prior probabilities - the first such general solution for a set of quantum states since the two-state case was solved when the problem of state discrimination was first introduced. We also investigate the difference between local and global measurements for a bipartite ensemble of states, and find that in certain circumstances the local measurement is superior. We conclude by finding a bipartite analogue to the Helstrom conditions, which indicate when a POVM satisfies the minimum-error criteria.

Acknowledgements

First and foremost, thank you to my supervisors, Sarah Croke and Stephen Barnett, for their patient guidance and support, and for creating a wonderful atmosphere in which to work; it has been a pleasure and a privilege.

Especial thanks also to Matthias Sonnleitner, for the daily help and chit-chat - I couldn't have asked for a better office-mate.

Thank you to my co-author, Catherine Hughes, for productive discussions, for teaching me several useful tips and tricks, and for having exquisitely-formatted Mathematica files.

Thank you to the rest of the QT group members, past and present: Gergely Ferenczi, Fiona Speirits, Thomas Brougham, Jim Cresser, Neel MacKinnon, Rob Cameron, Scarlett Gao, Vašek Potoček, Anette Messinger, Andrew Colin, Can Ritboon, Kieran Flatt, Jörg Götte, and Kevin Piché. Each of you has helped me in some way, whether through talking about physics or simply being good company (or both).

Thank you to the secretaries, Kirsteen Faulds and Rachael McLauchlan, for excelling in their duties.

I would like to acknowledge the financial support I received from the University of Glasgow College of Science And Engineering.

Thank you to my friends Stuart Gray, Craig McAnally, Alex "Totti" Morton, Andy Isacsson, and Greg Law, simply for being my friends.

I would also like to thank my parents for their endless support, which manifested itself in countless ways. Lastly, I would like to thank my three furry pillars of emotional support: Jack, Morgan, and finally Harry, who is dearly missed every day.

Contents

List of Figures

Chapter 1

Introduction

1.1 Overview

When communicating using non-orthogonal quantum states, there is an intrinsic probabilistic error associated with any measurement process. This error is the result of the superposition principle, and lies at the heart of the difference between classical and quantum communication. The security of quantum communication is simply another manifestation of this effect, as any eavesdropper is unlikely to be able to both measure the message accurately and resend the message to the intended recipient without revealing his or her presence (as seen in, e.g., the BB84 protocol [3]). This intrinsic error also means that determining which sequence of quantum states – and thus, which message – was sent is a non-trivial problem. One can define a number of different figures of merit, and then attempt to find a measurement which maximises (or minimises) this for an arbitrary set of quantum states. In this thesis we focus on the minimum-error and maximum confidence figures of merit, both already well-studied at this point [4–12]. We will begin by introducing the tools of quantum mechanics and quantum information in this chapter, including a description of the most general form of measurement on a quantum state. The next chapter discusses quantum state discrimination in depth, and summarises what is currently known in this area. In Chapter 3, we discuss a new solution to the problem of minimum-error quantum state discrimination for arbitrary single-qubit signal states with arbitrary prior probabilities. In Chapter 4, we investigate the so-called trine states and how to maximise various figures of merit for the discrimination of these states in the single-qubit regime. This work is continued in Chapter 5, where we extend to the multiple-photon case and discuss the effect of non-ideal detectors on this measurement process. In Chapter 6 we introduce an analogue to the well-known Helstrom conditions for minimum-error discrimination, adapted to the problem of bipartite state discrimination. The work contained herein has resulted in the following publications (presented in the same order as above):

- G. Weir, S. M. Barnett and S. Croke, "Optimal discrimination of single-qubit mixed states", *Physical Review A* **96**, 022312 (2017);
- G. Weir, C. Hughes, S. M. Barnett and S. Croke, "Optimal measurement strategies for the trine states with arbitrary prior probabilities," *Quantum Sci. Technol.* 3(3):035003 (2018);
- G. Weir, C. Hughes, S. M. Barnett and S. Croke, "Optimal measurement strategies for multiple copies of the trine states," *In preparation*;
- S. Croke, S.M. Barnett and G. Weir, "Optimal sequential measurements for bipartite state discrimination," *Physical Review A* **95**, 052308 (2017).

1.2 Observables In Quantum Mechanics

Quantum Information Theory is concerned with storing, transmitting, manipulating and extracting information in a quantum system. Extracting information involves measuring the system, preferably in such a way that we maximise confidence that we are correctly identifying which state the system was prepared in. That is, given an ensemble of quantum states, how might one construct a physically-realisable measurement such that we may optimally determine which state the system was prepared in, and therefore extract information about our ensemble?

The vague language of "optimal" determination is a deliberate choice, as there are various metrics by which one can gauge the efficacy of a measurement. The simplest and most intuitive metric is to minimise the probability of error, P_{err} [5]. That is, given only one opportunity to determine which state the system was prepared in, this measurement will maximise the probability that we may answer correctly. Another measurement strategy is unambiguous discrimination, which will tell the observer with certainty which state the system was prepared in - at the cost of occasionally producing an inconclusive outcome [13–16]. A third strategy we might employ is that of maximum confidence - this is a generalisation of unambiguous discrimination, in which we tailor each measurement outcome in such a way that we maximise the probability that it is correctly identifying the corresponding signal state [7]. This strategy also includes an inconclusive outcome in certain circumstances. Each of these will be explained in more detail in what follows.

The form of the measurement itself is also of interest, as there are different classes of measurements. We will focus on only two measurement classes: LOCC (Local Operations with Classical Communication) and global measurements. These are explained further in §2.1.

1.3 Quantum States

1.3.1 Basic definitions

In order to start our discussion of quantum states, we must first define a *Hilbert space*: a Hilbert space is a complex vector space with an inner product defined on it. Hilbert spaces may have any number of dimensions - including infinitely many - but for our purposes we will only consider them to be finite-dimensional.

It is also useful to define several other terms. A set of basis vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ for a vector space is called *linearly dependent* if it is possible to define the origin in terms of these vectors with non-zero coefficients. That is,

$$0 = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n, \tag{1.1}$$

where at least some of $\{a_1, a_2, \ldots, a_n\}$ are not zero. A set of basis vectors is said to be *linearly independent* if it is not linearly dependent - or, equivalently, if each point in the vector space has one and only one decomposition in terms of the space's basis vectors.

A set of basis vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is described as *orthogonal* if

$$\mathbf{v}_i \cdot \mathbf{v}_j = 0 \quad \forall i \neq j, \tag{1.2}$$

where we use $\cdot$ to describe the inner product operation. A set of basis vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is described as *orthonormal* if it is both orthogonal and each basis vector has unit inner product with itself. That is,

$$\mathbf{v}_i \cdot \mathbf{v}_j = \delta_{ij}, \tag{1.3}$$

where δ_{ij} is the Kronecker delta. By definition, any orthogonal or orthonormal set of basis vectors is linearly independent.

Unless explicitly stated otherwise, all vector spaces in this work shall have an orthonormal basis.

1.3.2 State vectors

We may describe the state of a quantum system with a vector in a Hilbert space. Such a vector is called a state vector, and may be presented in two equivalent ways. The first, and more common, is the ket vector:

$$|\psi\rangle, \tag{1.4}$$

which represents column vectors. The second way to present a state vector is the bra vector:

$$\langle\phi|, \tag{1.5}$$

which represents row vectors, or one-forms. As is usual in a vector space, these may be multiplied by scalars. If we multiply the previous ket vector by the complex scalar λ, this yields

$$\lambda|\psi\rangle, \tag{1.6}$$

which has the corresponding bra vector:

$$\langle\psi|\lambda^*, \tag{1.7}$$

where λ^* denotes the complex conjugate of λ.

We may also define addition of two vectors. If $|\psi_1\rangle$ and $|\psi_2\rangle$ are valid state vectors, then so is any vector of the form

$$|\Psi\rangle = a|\psi_1\rangle + b|\psi_2\rangle, \tag{1.8}$$

where a and b are again complex numbers. The bra form of this vector is given by:

$$\langle\Psi| = \langle\psi_1|a^* + \langle\psi_2|b^*. \tag{1.9}$$

This is the superposition principle, and is one of the fundamental characteristics of quantum mechanics. It is due to this that it is impossible for two non-orthogonal quantum states (i.e. two states with non-zero mutual inner product) to be discriminated perfectly - we will see this in more detail in §2.2. It is this effect that gives security to quantum communications such as the famous BB84 protocol [3], which we will discuss in §1.5.3.

The inner product (or scalar product) of two vectors $|\psi\rangle$ and $|\phi\rangle$ is denoted as $\langle\psi|\phi\rangle$ and has the standard properties of being conjugate-symmetric, linear, and positive-definite. That is,

$$\langle\psi|\phi\rangle = \langle\phi|\psi\rangle^*, \tag{1.10}$$

$$\langle\psi|(a|\phi_1\rangle + b|\phi_2\rangle) = a\langle\psi|\phi_1\rangle + b\langle\psi|\phi_2\rangle, \tag{1.11}$$

$$\langle\psi|\psi\rangle \geq 0, \tag{1.12}$$

with equality in the last line if and only if $|\psi\rangle = 0$.

1.3.2.1 Composite systems

We may also denote the presence of several quantum states at once by using the tensor product, $\otimes$. A state consisting of two copies of the same state $|\psi\rangle$ may be written:

$$|\Psi\rangle = |\psi\rangle \otimes |\psi\rangle. \tag{1.13}$$

Three copies of the state may be written:

$$|\Psi'\rangle = |\psi\rangle \otimes |\psi\rangle \otimes |\psi\rangle, \tag{1.14}$$

and so on. If we wish to, say, perform the unitary operation U on the middle qubit in the ensemble described in equation (1.14), we simply perform the operation $\mathbb{1} \otimes U \otimes \mathbb{1}$ on the state vector $|\Psi'\rangle$.

1.4 Measurements

John von Neumann is responsible for the standard textbook formalism of measurements; we will begin with a description of this before moving on to the more general POVM paradigm.

1.4.1 von Neumann Measurements

It is conventional to represent some observable A by the Hermitian operator $\hat{A}$; we let this operator have n eigenstates $\{|\lambda_0\rangle, |\lambda_1\rangle, \ldots |\lambda_n\rangle\}$ with real eigenvalues $\{\lambda_0, \lambda_1, \ldots, \lambda_n\}$. That is,

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

for some $j \in \{0, \ldots, n\}$. The eigenvalues are the possible outcomes of a measurement of A, hence the necessity for $\hat{A}$ to be Hermitian - it is the most general form of operator which still must have real eigenvalues in orthogonal eigenspaces. The eigenstates are orthonormal and form a complete set - that is, any arbitrary ket $|\alpha\rangle$ in the ket-space of the observable A may be written

$$|\alpha\rangle = \sum_i c_i|\lambda_i\rangle.$$

We may therefore consider these eigenstates to be a basis in which to view our observable A. Indeed, we may express the operator $\hat{A}$ purely in terms of its eigenstates and eigenvalues:

$$\hat{A} = \sum_i \lambda_i|\lambda_i\rangle\langle\lambda_i|.$$

Now suppose we have an ensemble in a mixed state where each pure state $|\psi_j\rangle$ occurs with a probability p_j. Then the expectation value of A is given by

$$\begin{aligned}\langle\hat{A}\rangle &= \sum_j p_j\langle\psi_j|\hat{A}|\psi_j\rangle \\ &= \sum_j p_j\,\mathrm{Tr}(|\psi_j\rangle\langle\psi_j|\hat{A}) \\ &= \sum_j \mathrm{Tr}(p_j|\psi_j\rangle\langle\psi_j|\hat{A}) \\ &= \mathrm{Tr}(\sum_j p_j|\psi_j\rangle\langle\psi_j|\hat{A}) \\ &= \mathrm{Tr}(\hat{\rho}\hat{A}),\end{aligned}$$

where we have introduced the density operator $\hat{\rho} = \sum_j p_j|\psi_j\rangle\langle\psi_j|$, which fully describes our knowledge of the quantum ensemble.

If we also introduce the projection operator $\hat{P}_j = |\lambda_j\rangle\langle\lambda_j|$, the probability that a measurement of A will produce the eigenvalue λ_j is given by Born's rule:

$$\mathrm{P}(\lambda_j) = \langle\lambda_j|\hat{\rho}|\lambda_j\rangle = \mathrm{Tr}(\hat{\rho}|\lambda_j\rangle\langle\lambda_j|) = \mathrm{Tr}(\hat{\rho}\hat{P}_j). \tag{1.15}$$

For a pure state, $p_j = 1$ and so this is simply the expectation value of the state. A von Neumann measurement is defined to be one where the probability of some outcome λ_j is given by the above equation. It is important to note that this can easily be extended to the degenerate case where, for some degenerate orthonormal eigenstates $|\lambda_j^1\rangle, |\lambda_j^2\rangle$ which produce the eigenvalue λ_j, we simply define $\hat{P}_j = |\lambda_j^1\rangle\langle\lambda_j^1| + |\lambda_j^2\rangle\langle\lambda_j^2|$. These projectors $\hat{P}_j$ have the following properties:

(i) $\hat{P}_j = \hat{P}_j^\dagger$ (Hermitian operators)

(ii) $\langle\psi|\hat{P}_j|\psi\rangle \geq 0 \quad \forall \; |\psi\rangle$ (positive operators)

(iii) $\sum_j \hat{P}_j = \mathbb{1}$ (completeness)

(iv) $\hat{P}_i\hat{P}_j = \hat{P}_j\delta_{ij}$ (orthonormality)

We interpret these properties as follows: the projectors represent real quantities, and therefore must be Hermitian; the expectation values are probabilities, and so cannot be negative; the sum of all probabilities for all possible measurements on any single state must be unity. The fourth property, however, has no intuitive probabilistic interpretation, and this in fact does not hold for non-ideal von Neumann measurements [17, §4.2]. Removing this condition leads us to positive-operator valued measurements (POVMs). von Neumann measurements are restricted to allowing only N measurement outcomes for any N-dimensional space; POVMs remove this restriction and allow for arbitrarily many measurement outcomes.

1.4.2 Positive-operator valued measurements

We now introduce a set of probability operators $\hat{\pi}_j$, where, analogously to the above, the probability of outcome j on some system described by a density operator $\hat{\rho}$ is given by

$$\mathrm{P}_j = \mathrm{Tr}(\hat{\rho}\hat{\pi}_j).$$

The set of probability operators is collectively called a positive-operator valued measure, or POVM. The POVM elements have the below properties:

(i) $\hat{\pi}_j = \hat{\pi}_j^\dagger$ (Hermitian)

(ii) $\langle\psi|\hat{\pi}_j|\psi\rangle \geq 0 \quad \forall \quad |\psi\rangle$ (positive)

(iii) $\sum_j \hat{\pi}_j = \mathbb{1}$ (complete)

An important point to note is the fact that there is a bijection between measurements which fulfil these criteria and physically realisable measurements: that is, any POVM with the above properties may be experimentally realised, and any physical measurement apparatus may be expressed in POVM form [17, §4.3], [4].

Removing the condition for orthonormality affects the number of elements that the measurement may have: while a von Neumann measurement has the number of elements restricted to at most the number of eigenstates of our observable, the number of POVM elements may be smaller than or greater than the dimension of the ket-space of the observable.

To see why this can be advantageous over a simple von Neumann measurement, we consider unambiguous discrimination of non-orthogonal states.

We will discuss quantum state discrimination in more detail in the next chapter. For the purposes of this section, we will use the simplest definition possible: one communicating party, called Alice, is sending information in the form of quantum states to Bob. Bob may make any measurement he chooses in order to determine which quantum states were sent; there are a number of figures of merit he may wish to maximise with this measurement, but for now we will focus on a measurement which leads Bob to never misidentify the signal state.

1.4.2.1 Example: Unambiguous discrimination

Consider the two (generally) non-orthogonal states

$$|\psi_0\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle$$
$$|\psi_1\rangle = \cos\theta|0\rangle - \sin\theta|1\rangle,$$

(without loss of generality, $\theta \in (0, \frac{\pi}{2}])$ which, to simplify matters, we assume have each been prepared with equal probability; that is, $p_0 = p_1 = \frac{1}{2}$. It is known (and shown in §2.2) that a von Neumann measurement cannot unambiguously discriminate between these two states for arbitrary θ; if we wish unambiguous information, the best that we can do is form a measurement with two outcomes, one of which allows us to state with certainty that, say, the system was not prepared in state $|\psi_1\rangle$ (and therefore, by elimination, must have been prepared in state $|\psi_0\rangle$), while the other outcome will give us no unambiguous information [13–15]. This is realised by the measurement:

$$\begin{aligned}\hat{\pi}_0 &= (\sin\theta|0\rangle + \cos\theta|1\rangle)(\sin\theta\langle 0| + \cos\theta\langle 1|)\\ \hat{\pi}_? &= |\psi_1\rangle\langle\psi_1|,\end{aligned} \tag{1.16}$$

with the above probability interpretations supported by the fact that $\langle\psi_1|\hat{\pi}_0|\psi_1\rangle = 0$ and $\langle\psi_0|\hat{\pi}_?|\psi_0\rangle \neq 0 \neq \langle\psi_1|\hat{\pi}_?|\psi_1\rangle$. We find that the measurement described in equation (1.16) allows us to identify the state $|\psi_0\rangle$ with probability $\mathrm{P}_0 = \frac{1}{4}(1 - \cos 4\theta)$, while giving the inconclusive outcome with probability $\mathrm{P}_? = \frac{1}{4}(3 + \cos 4\theta)$.

However, by moving to the POVM paradigm, we find that we may unambiguously discriminate between $|\psi_0\rangle$ and $|\psi_1\rangle$ if we allow for an inconclusive POVM element which gives us no information. Consider

$$\begin{aligned}\hat{\pi}_0 &= a_0(\sin\theta|0\rangle + \cos\theta|1\rangle)(\sin\theta\langle 0| + \cos\theta\langle 1|)\\ \hat{\pi}_1 &= a_1(\sin\theta|0\rangle - \cos\theta|1\rangle)(\sin\theta\langle 0| - \cos\theta\langle 1|)\end{aligned}$$

for some $0 \leq a_0, a_1 \leq 1$. This has the desired property that $\langle\psi_0|\hat{\pi}_1|\psi_0\rangle = \langle\psi_1|\hat{\pi}_0|\psi_1\rangle = 0$ (i.e. given outcomes 0 or 1, we can tell with absolute certainty what state the system was prepared in), but unless our states are orthogonal, this does not form a complete measurement. We must allow for an inconclusive outcome,

$$\hat{\pi}_? = \mathbb{1} - (\hat{\pi}_0 + \hat{\pi}_1) \tag{1.17}$$

which occurs with probability

$$\mathrm{P}_? = \frac{1}{2}\langle\psi_0|\hat{\pi}_?|\psi_0\rangle + \frac{1}{2}\langle\psi_1|\hat{\pi}_?|\psi_1\rangle = 1 - \frac{1}{2}(a_0 + a_1)\sin^2 2\theta. \tag{1.18}$$

Clearly, our measurement is optimised when the probability of this outcome is minimised (for $a_0, a_1 \geq 0, \hat{\pi}_? \geq 0$). This occurs when $\mathrm{P}_? = \langle\psi_0|\psi_1\rangle = \cos 2\theta$ [4], corresponding to the measurement:

$$\begin{aligned}\hat{\pi}_0 &= \frac{1}{2\cos^2\theta}(\sin\theta|0\rangle + \cos\theta|1\rangle)(\sin\theta\langle 0| + \cos\theta\langle 1|)\\ \hat{\pi}_1 &= \frac{1}{2\cos^2\theta}(\sin\theta|0\rangle - \cos\theta|1\rangle)(\sin\theta\langle 0| - \cos\theta\langle 1|)\\ \hat{\pi}_? &= (1 - \tan^2\theta)|0\rangle\langle 0|.\end{aligned}$$

However, the minimum-error von Neumann measurement will give the incorrect answer with probability $\mathrm{P}_{\mathrm{Error}} = \frac{1}{2}(1 - \sqrt{1 - \cos^2 2\theta}) = \frac{1}{2}(1 - \sin 2\theta)$, as we shall see in §2.3.1.

Because unambiguous state discrimination requires more measurement outcomes than there are dimensions in the Hilbert space, the measurement must take the form of a POVM. This example, therefore, shows a clear advantage of the POVM approach over the “traditional” von Neumann approach.

1.5 Quantum Information

As we have already seen, it is possible - due to the superposition principle - for a quantum state with basis vectors $|0\rangle$ and $|1\rangle$ to take the general form

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{1.19}$$

where we normalise such that $|\alpha|^2 + |\beta|^2 = 1$. We may make an analogy from the $|0\rangle$ and $|1\rangle$ basis vectors to the binary 0 and 1 of classical computer bits. In this context, we refer to the state $|\psi\rangle$ as a quantum bit - or "qubit" from here on. Any quantum system with two quantum states may be a physical realisation of a qubit, from photon polarisation to electron energy levels in an atom or ion to the orientation of a spin-half particle.

As qubits allow the superposition of $|0\rangle$ and $|1\rangle$ states, it is possible to realise a provably secure communication channel from Alice to Bob using qubits, as we will see in §1.5.3. Part of the basis of this is the no-cloning theorem.

1.5.1 The No-Cloning Theorem

The no-cloning theorem states that, given an arbitrary unknown qubit state $|\psi\rangle$, it is impossible to take a second, blank qubit, $|B\rangle$, and copy the state of the first qubit onto the blank one. That is, the transformation

$$|\psi\rangle \otimes |B\rangle \rightarrow |\psi\rangle \otimes |\psi\rangle \tag{1.20}$$

is impossible [18, 19]. To see that this is the case, we first see what the desired outcome is. For an arbitrary qubit of the form $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$,

$$(\alpha|0\rangle + \beta|1\rangle) \otimes (\alpha|0\rangle + \beta|1\rangle) = \alpha^2|0\rangle \otimes |0\rangle + \alpha\beta|0\rangle \otimes |1\rangle + \alpha\beta|1\rangle \otimes |0\rangle + \beta^2|1\rangle \otimes |1\rangle. \tag{1.21}$$

Now suppose that our hypothetical method for cloning qubits works if the original qubit is prepared in the states $|0\rangle$ or $|1\rangle$:

$$\begin{aligned} |0\rangle \otimes |B\rangle &\rightarrow |0\rangle \otimes |0\rangle \\ |1\rangle \otimes |B\rangle &\rightarrow |1\rangle \otimes |1\rangle. \end{aligned} \tag{1.22}$$

From this – and the principle of linearity which governs quantum mechanics – it follows that our arbitrary qubit state will be transformed as

$$(\alpha|0\rangle + \beta|1\rangle) \otimes |B\rangle \rightarrow \alpha|0\rangle \otimes |0\rangle + \beta|1\rangle \otimes |1\rangle, \tag{1.23}$$

which does not match the desired output given in equation (1.21).

This is crucial to the security of quantum communications: if such an operation were possible, any eavesdropper would be able to intercept Alice's communications to Bob, make multiple copies of the state and perform state tomography to read the message with arbitrary accuracy, and then send an additional copy of the message to Bob without Alice or Bob's knowledge.

1.5.2 Bloch Sphere Representation

It is possible to represent any pure qubit state as:

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle. \tag{1.24}$$

This corresponds to a point on the unit sphere with spherical polar co-ordinates ϕ and θ, as seen in Figure 1.1. This sphere is called the Bloch Sphere, and is used to visually represent the state of a pure qubit. We may also place any number of single-qubit states on the Bloch Sphere to represent either signal states or measurement states - if Bob's measurement includes a POVM element of the form $\pi_0 = |0\rangle\langle 0|$, we will refer to this as him measuring along the state $|0\rangle$.

It is important to note that this is not restricted to pure states; we may place a mixed qubit state of the form, say, $|\psi_{\text{mixed}}\rangle = 0.75|0\rangle\langle 0| + 0.25|1\rangle\langle 1|$ *inside* the Bloch Sphere, with the distance from the centre determined by the difference between its two orthogonal components (for it is always possible to decompose any qubit mixed state in terms

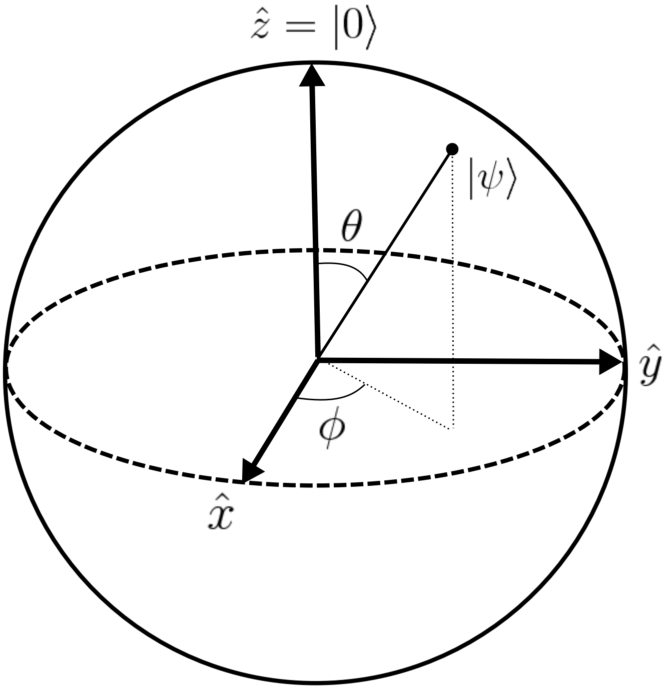

Figure 1.1: Visual representation of the pure state described in equation (1.24) as a point on the surface of the Bloch sphere, showing how θ and ϕ describe the state.

of two orthogonal components): in our example, the Bloch vector for $|\psi_{\text{mixed}}\rangle$ would be of length 0.5 and point in the $|0\rangle\langle 0|$ direction. The maximally mixed state is simply represented by a point in the centre of the Bloch sphere.

The Bloch sphere representation is useful for a number of reasons. Firstly, it gives us a very quick intuition of how easily-distinguishable two states are: the closer the two states appear on the Bloch sphere, the greater overlap they have. This also means that in a measurement situation, these states will be hard to distinguish between. Secondly, we can define a bijective map from the Bloch Sphere to the Poincaré Sphere, which depicts different polarisations of light as points on a sphere. This allows us to easily switch from thinking of qubit states to thinking of polarisations of light, and vice versa. Thirdly, we may depict the action of unitary operations simply as rotations on the Bloch sphere. For instance, the Pauli-X gate, whose operation is given by:

$$\hat{X} = \hat{\sigma}_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{1.25}$$

corresponds to a rotation around the X-axis of the Bloch Sphere of π radians (i.e. $|0\rangle$ flips to $|1\rangle$ and vice versa). It is clear that by applying this operation twice in a row, we simply rotate by 2π radians. That is, $\hat{X}^2 = \mathbb{1}$.

1.5.3 Quantum Key Distribution

Due to the superposition principle (and the no-cloning theorem), qubits naturally form a convenient starting point for realising a provably secure communication channel. This is the basis of quantum key distribution.

First, we discuss how the secret key which Alice and Bob share would be used. The simplest example is the *one-time pad*, which was proven to be information-theoretically secure (i.e., the encrypted message provides no information about the original message, apart from its maximum possible length) by Shannon [20]: in this, the message to be conveyed (called the plaintext) is added to the key, modulo two. That is,

$$\begin{aligned} 0+0=0 &\mod 2 \\ 0+1=1 &\mod 2 \\ 1+0=1 &\mod 2 \\ 1+1=0 &\mod 2. \end{aligned} \tag{1.26}$$

Alice adds the plaintext (P) to the key (K) in this way, bit by bit, to form the ciphertext (C):

$$P = 0101\ldots \tag{1.27}$$

$$K = 0100\ldots \tag{1.28}$$

$$C = P + K \mod 2 = 0001\ldots \tag{1.29}$$

On the other end, Bob then adds the key to the ciphertext to receive the plaintext:

$$C = 0001\ldots \tag{1.30}$$

$$K = 0100\ldots \tag{1.31}$$

$$P = C + K \mod 2 = 0101\ldots. \tag{1.32}$$

This works because the addition of the key with itself modulo two will always result in the identity: $K + K \mod 2 = 0000\ldots$.

Clearly for the purposes of security, it is optimal for the one-time pad to be as long as the plaintext message. However, if such a long key can be distributed securely, then Alice and Bob could simply use this channel to exchange messages. Alice and Bob require a way of exchanging the one-time pad over a channel which is not secure, while being able to identify if someone is eavesdropping on them.

This process involves using a quantum channel and classical communication to generate a secure private key, and is known as *quantum key distribution*, or QKD. QKD is perhaps the most important use of quantum state discrimination; note that while the basis used in our BB84 example below has useful properties, Alice could send any set of quantum states $\{|\psi_0\rangle, |\psi_1\rangle, \ldots, |\psi_n\rangle\}$ with any prior probabilities $\{p_0, p_1, \ldots, p_n\}$ to exchange a key with Bob.

The much-celebrated BB84 protocol [3] was the first quantum key distribution protocol to be developed, and is a useful example for illustrating the core concepts. Two parties, named Alice and Bob, are trying to communicate securely; an eavesdropper, Eve, is attempting to intercept their messages. The most important aspect of the BB84 protocol is that it yields a way for Alice to communicate with Bob with no fear that Eve might be able to glean any useful information - by measuring the states, Eve must introduce errors which reveal her presence.

Alice encodes qubits as follows: light polarisations $|H\rangle$ and $|D\rangle$ ($|0\rangle$ and $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, respectively) map to 0, while $|V\rangle$ and $|A\rangle$ ($|1\rangle$ and $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, respectively) map to 1. Alice then sends a string of 0s and 1s to Bob, randomly varying her basis from horizontal/vertical to diagonal/antidiagonal and back such that each basis is used 50% of the time. Her output resembles the following:

Photon number:	1	2	3	4	5	6	7	8
Output:	$\|H\rangle$	$\|V\rangle$	$\|D\rangle$	$\|A\rangle$	$\|A\rangle$	$\|A\rangle$	$\|D\rangle$	$\|H\rangle$
Bit value sent:	0	1	0	1	1	1	0	0

Bob then measures this string of bits, varying his measurement basis randomly from horizontal/vertical (H/V) to diagonal/antidiagonal (D/A). Bob announces publicly (i.e. assuming that Eve can intercept and read this communication) the measurement bases he used for each photon - but, crucially, not his measurement outcomes. Alice can then confirm to Bob on which occasions they used the same basis; the other bits are discarded. Neglecting the impact of Eve's actions at the moment, this looks like the following, where an "X" demonstrates that Bob's measurement basis did not match Alice's output basis:

Photon number:	1	2	3	4	5	6	7	8
Output:	$\|H\rangle$	$\|V\rangle$	$\|D\rangle$	$\|A\rangle$	$\|A\rangle$	$\|A\rangle$	$\|D\rangle$	$\|H\rangle$
Bit value sent:	0	1	0	1	1	1	0	0
Bob's measurement basis:	H/V	D/A	H/V	H/V	D/A	H/V	D/A	H/V
Match?		X	X	X		X		
Bit received:	0	1	1	0	1	0	0	0
Same-basis bits:	0				1		0	0

At this point, Alice and Bob announce several of the bits in which they used the same basis. In our unrealistically short example, they both share the bit string 0100; however, they both publicly announce that the first bit in this string is 0 - as this is public, they discard this bit. Their secret key is now 100.

Now suppose we have an eavesdropper, Eve, who is intercepting Alice's photons, making a measurement, and then sending on a photon matching her measurement outcome to Bob:

Photon number:	1	2	3	4	5	6	7	8
Output:	$\|H\rangle$	$\|V\rangle$	$\|D\rangle$	$\|A\rangle$	$\|A\rangle$	$\|A\rangle$	$\|D\rangle$	$\|H\rangle$
Bit value sent:	0	1	0	1	1	1	0	0
Measurement basis (E):	D/A	D/A	D/A	H/V	H/V	D/A	H/V	H/V
Eve's outcome:	0	1	0	1	1	1	0	0
Measurement basis (B):	H/V	D/A	H/V	H/V	D/A	H/V	D/A	H/V
Match?		X	X	X		X		
Bit received:	1	1	1	0	1	0	1	0
Same-basis bits (A):	0				1		0	0
Same-basis bits (B):	1				1		1	0

Again, Alice and Bob announce several of the bits in which they know they have used the same basis: however, Alice has the string 0100 but Bob has the string 1110. Alice publicly announces the first shared bit to be 0; Bob publicly announces the first shared bit to be 1. In principle, the only way for such a discrepancy to exist is that an eavesdropper intercepted Alice's horizontally-polarised photon (a 0 in this basis), measured it at be diagonally-polarised, and then sent another such photon to Bob: when Bob measured this $|D\rangle$ photon in the H/V basis, he received the outcome corresponding to 1. They now know that their quantum communication channel has been compromised.

Clearly a realistic scenario of this would involve the transfer of a far larger number of bits from Alice to Bob; however, this toy example illustrates the key concept. It is also important to note that in a real-world implementation of such a system, errors would be inevitable as a result of noise and other real-world sources of error. Alice and Bob would then need to decide what error-rate is acceptable for whichever quantum channel they are using: if, after some form of error correction has been performed [21], the error rate is below this, they may use the key to communicate; if it is above the agreed-upon level, they assume that the channel has been compromised.

Alice and Bob might also wish to estimate the amount of information that Eve may have obtained about each of the key bits they have shared. Therefore Alice and Bob wish to find the maximum probability

$$\mathrm{P}_{\text{Eve}} = \frac{1}{2}(1+\epsilon) \tag{1.33}$$

that Eve has correctly identified any individual bit, and preferably wish to put an upper bound on this. They do this by a process called privacy amplification, first mentioned by Brassard *et al.* [22], with various other forms existing [23–25]. We will discuss one method: first, Alice and Bob break their key up into m bits and use the parity of each "chunk" as a single bit in the final key – that is, an even number of bits with value 1 in their string of m bits will manifest as a 0 in the final key, while an odd number of bits with value 1 will manifest as a 1 in the final key. Eve can only identify this parity if she makes an even number of errors in identifying the m bits (including zero errors). This will occur with probability

$$\begin{aligned}\mathrm{P}_{\text{Eve}} &= (\frac{1}{2})^m(1+\epsilon)^m + (\frac{1}{2})^m\frac{m!}{2!(m-2)!}(1+\epsilon)^{m-2}(1-\epsilon)^2 + \dots \\ &= \frac{1}{2}(1+\epsilon^m), \end{aligned} \tag{1.34}$$

which is closer to the ideal of $\frac{1}{2}$ than the value given in equation (1.33). Alice and Bob must choose m in order to achieve their pre-decided maximum value for P_{Eve}.

It is worth stopping to reflect on what has been done here. We have shown that it is possible for Alice and Bob to exchange a key with which they can communicate securely, to an arbitrary degree of privacy. The key point is that Eve's attempts to measure the system will introduce errors which reveal her presence.

Note that we have not yet discussed the possible strategies Eve might take. Clearly, Eve wishes to avoid introducing errors into Bob's bit string, while still obtaining information herself. In the ideal scenario described here, the only way to achieve this is for her to measure fewer photons, which will also reduce the amount of information she receives. The optimal strategy Eve might employ varies depending on the specific key distribution system and Alice's choice of states; Eve's possible choices of strategies are also an important use of quantum state discrimination.

In what follows, we drop the "hat" for operators whenever it is not confusing to do so. That is, the POVM element $\hat{\pi}_i$ will now simply be represented by π_i.

Chapter 2

Quantum State Discrimination

2.1 Introduction

As mentioned in the previous chapter, quantum key distribution – and quantum communication in general – depends on the problem of quantum state discrimination [4, 26]. The standard formulation of this problem involves two communicating parties, Alice and Bob: Alice communicates with Bob by sending him a quantum state ρ_i which has been chosen from a set of possible states $\{\rho_j\}$, each with an *a priori* probability p_j. Bob knows these states and their probabilities, and his goal is to determine which state was sent, thereby decoding the message which Alice wishes to communicate. Clearly, Bob wishes to decode the message as best he can, and he may quantify this using any of a number of different figures of merit. The two most common figures of merit he might wish to maximise are mutual information in bits [27, 28], and the probability of correctly identifying the state [5, 9, 11, 29] (equivalent to minimising the error given by a POVM). He may also use the techniques of unambiguous discrimination [16, 30] – which either gives an inconclusive outcome or identifies the signal state with certainty – or maximum confidence, a generalisation of this which sometimes yields incorrect answers [7]. Note that we are assuming that all errors in state identification are equally bad - there is no merit in being "nearly right".

This has been a popular problem for a few decades, with theoretical solutions obtained and experiments performed for various sets of states and figures of merit [1, 6–9, 11, 13–16, 28, 30–39]. This popularity may be attributed partially to the fundamental nature of the problem, and also to its far-reaching consequences: in addition to being crucial for quantum key distribution, state discrimination has relevance in quantum information processing and quantum metrology [26], and also allows us to explore the constraints on different measurement classes such as global measurement or local measurement with classical feed-forward [40–42].

In our work, we will focus on LOCC and global measurements. An LOCC measurement involves Alice and Bob each being in possession of a single quantum state, with the knowledge that these states are the same (i.e. have been prepared identically). Alice performs a local (product) operation on her state, and classically communicates the result of this measurement to Bob. Using this information, Bob may alter his measurement; he then performs a local measurement on his copy of the state, and sends this result to Alice. This process may be repeated many times [33] if the measurements only disturb the state minimally [43]. In a global measurement, Alice and Bob's qubits are coupled in some way and then measured together in such a way that only one measurement outcome is produced (i.e., as opposed to one each for Alice and Bob).

Further applications of quantum state discrimination include bounding the dimension of a system's Hilbert space given incomplete information [44], and sharing information through imperfect cloning [45–47]. For minimum error and unambiguous discrimination, the problem of optimisation may be cast as a semi-definite programme, and particular instances may thus be solved efficiently numerically. However, despite recent progress in analytical techniques for minimum error discrimination, explicit analytic solutions are available only for the simplest cases.

2.2 Non-Orthogonal Signal States

One of the fundamental features of quantum mechanics is the superposition principle. When an N-level quantum system - which by definition exists in an N-dimensional vector space with orthogonal basis states $\{|i\rangle\}, i = 0, \ldots, N-1$ - exists in one of its basis states, a measurement will result in the corresponding eigenvalue with certainty. However, the superposition principle states that any linear superposition of the eigenstates of the form $|\Psi\rangle = \sum_i a_i|i\rangle$ is also an allowable state of the system, hence giving rise to the existence of non-orthogonal states. We show here that discrimination between non-orthogonal states will always have some associated intrinsic error.

Suppose we have an ensemble of states $\{|\psi_i\rangle\}$ from which Alice chooses one to send to Bob. Bob then makes a measurement to determine which state was sent. Let us also suppose that perfect discrimination is possible in this case: i.e., there exists some POVM element π_j such that we obtain a "click" at π_j if and only if the signal state is $|\psi_j\rangle$. That is, from the Born rule given in equation (1.15),

$$\mathrm{P}(\pi_j|\psi_i) = \mathrm{Tr}(\pi_j \rho_i) = \delta_{ij}, \tag{2.1}$$

where $\rho_i = |\psi_i\rangle\langle\psi_i|$. Given that $\pi_j \leq \mathbb{1}$, clearly $\mathrm{Tr}(\pi_j\rho_j) \leq 1$, with equality if and only if

$$\pi_j = |\psi_j\rangle\langle\psi_j| + \xi_j \tag{2.2}$$

where $\mathrm{Tr}(\xi_j\rho_j) = 0$ and, to satisfy positivity, $\xi_j \geq 0$. Therefore

$$\mathrm{P}(\pi_j|\psi_i) = \mathrm{Tr}(\pi_j\rho_i) = |\langle\psi_j|\psi_i\rangle|^2 + \mathrm{Tr}(\rho_i\xi_j). \tag{2.3}$$

If we wish for equation (2.1) to hold, we must have $\mathrm{Tr}(\rho_i\xi_j) = 0$ and $|\langle\psi_j|\psi_i\rangle|^2 = \delta_{ij}\forall i, j$. We therefore arrive at the conclusion that perfect discrimination is only possible if the states $\{|\psi_i\rangle\}$ are mutually orthogonal. The corollary to this is that, clearly, if the states $\{|\psi_i\rangle\}$ are *not* mutually orthogonal then there will be some intrinsic error in any attempt at state discrimination. In what follows, we attempt to mitigate this by trying to optimise a number of different figures of merit.

2.3 Minimum-Error Discrimination

Perhaps the most natural figure of merit that Bob might wish to maximise is the probability that any individual measurement will correctly identify the signal state. From Born's Rule, equation 1.15, this is given by:

$$\mathrm{P}_{\mathrm{Corr}} = \mathrm{Tr}(\sum_i p_i\rho_i\pi_i) \tag{2.4}$$

and it is this quantity which we wish to maximise. Necessary and sufficient conditions that a POVM must satisfy have been known since the inception of the problem [29, 48], and are named the Helstrom Conditions:

$$\pi_i(p_i\rho_i - p_j\rho_j)\pi_j = 0 \quad \forall i, j \tag{2.5}$$

$$\Gamma - p_j\rho_j \geq 0 \quad \forall j, \tag{2.6}$$

where $\Gamma = \sum_i p_i\rho_i\pi_i$. Note that we may also write $\Gamma = \sum_i p_i\pi_i\rho_i$ as a result of equation (2.5): by expanding the brackets and summing over all i and j, we see that $\sum_i p_i\pi_i\rho_i = \sum_j p_j\rho_j\pi_j$, i.e. $\sum_i p_i\pi_i\rho_i = \sum_i p_i\rho_i\pi_i$. Note that this also implies that Γ is Hermitian.

It is not too complicated to prove necessity and sufficiency of these conditions following the strategies used in [4] and [49], and we shall do so here. We begin with sufficiency. If $\{\pi_i\}$ corresponds to the minimum-error measurement strategy, then clearly:

$$\sum_i p_i \operatorname{Tr}(\rho_i\pi_i) \geq \sum_j p_j \operatorname{Tr}(\rho_j\pi'_j) \tag{2.7}$$

for all other possible POVMs $\{\pi'_j\}$. If we insert the identity $\sum_j \pi'_j = \mathbb{1}$, we obtain:

$$\sum_j \operatorname{Tr}((\sum_i p_i\rho_i\pi_i - p_j\rho_j)\pi'_j) \geq 0. \tag{2.8}$$

We know that $\pi'_j \geq 0$, and therefore the above is true if condition (2.6) holds. Therefore (2.6) is a sufficient condition.

Note that condition (2.5) is not sufficient, as any POVM of the form $\pi_i = \mathbb{1}$, $\pi_{i\neq j} = 0$ satisfies this condition for any choice of i.

To prove necessity, we introduce the Hermitian operators

$$G_j = \sum_i \frac{1}{2}p_i(\rho_i\pi_i + \pi_i\rho_i) - p_j\rho_j, \tag{2.9}$$

where the operators $\{\pi_i\}$ form a minimum-error POVM. We show that if this is the case, then each of these G_j operators must be positive. Suppose that, without loss of generality, for state ρ_0 the operator G_0 has a single negative eigenvalue $-\lambda$; that is,

$$G_0|\lambda\rangle = -\lambda|\lambda\rangle. \tag{2.10}$$

We show that if this is the case, then there exists another POVM which yields a greater probability for correctly guessing the signal state. Thus the positivity of G_0 is a necessary condition for a minimum-error POVM.

Now consider a new POVM with elements

$$\pi_i' = (\mathbb{1} - \epsilon|\lambda\rangle\langle\lambda|)\pi_i(\mathbb{1} - \epsilon|\lambda\rangle\langle\lambda|) + \epsilon(2-\epsilon)|\lambda\rangle\langle\lambda|\delta_{i0}, \tag{2.11}$$

where $0 < \epsilon \ll 1$. These clearly form a POVM, as $(\mathbb{1} - \epsilon|\lambda\rangle\langle\lambda|)\pi_i(\mathbb{1} - \epsilon|\lambda\rangle\langle\lambda|)$ and $|\lambda\rangle\langle\lambda|$ are clearly positive and $\sum_i \pi_i' = \mathbb{1}$. The probability that this POVM will correctly identify the signal state is given by

$$\begin{aligned} \mathrm{P}'_{\mathrm{Corr}} &= \sum_i p_i \operatorname{Tr}(\rho_i \pi_i') \\ &= \sum_i p_i \operatorname{Tr}[\rho_i(\mathbb{1} - \epsilon|\lambda\rangle\langle\lambda|)\pi_i(\mathbb{1} - \epsilon|\lambda\rangle\langle\lambda|)] + \epsilon(2-\epsilon)p_0\langle\lambda|\rho_0|\lambda\rangle \\ &= \mathrm{P}_{\mathrm{Corr}} - 2\epsilon\sum_i p_i\langle\lambda|\frac{1}{2}(\rho_i\pi_i + \pi_i\rho_i)|\lambda\rangle + 2\epsilon p_0\langle\lambda|\rho_0|\lambda\rangle + O(\epsilon^2) \\ &= \mathrm{P}_{\mathrm{Corr}} + 2\epsilon\lambda + O(\epsilon^2), \end{aligned} \tag{2.12}$$

which is greater than $\mathrm{P}_{\mathrm{Corr}}$, contradicting our assumption that $\{\pi_i\}$ is a minimum-error POVM. This is clearly true for any state ρ_j, so if any of the operators G_j has a negative eigenvalue then the corresponding POVM is not optimal. Therefore the positivity of each G_j is necessary for the POVM $\{\pi_i\}$ to be optimal. As we have seen, Γ is Hermitian, so

$$G_j = \Gamma - p_j\rho_j. \tag{2.13}$$

This shows the necessity of equation (2.6), but we also wish to show the necessity of equation (2.5). To do this, note that equation (2.6) combined with

$$\sum_i \operatorname{Tr}[(\Gamma - p_i\rho_i)\pi_i] = 0 \tag{2.14}$$

yields

$$\begin{aligned} (\Gamma - p_k\rho_k)\pi_k &= 0, \\ \pi_j(\Gamma - p_j\rho_j) &= 0. \end{aligned} \tag{2.15}$$

If we premultiply the first of these with π_j, postmultiply the second with π_k, and take the difference we obtain equation (2.6).

Clearly to distinguish between n states we need a measurement with at most n outcomes. The number of outcomes may be less than n (or equivalently, some of the operators π_i may be zero) if Bob's measurement procedure is such that some states are never identified.

Interestingly, it is sometimes optimal to avoid measurement and simply guess that the signal state is the *a priori* most likely state [50]. In this case, we use the measurement $\{\pi_j = \mathbb{1}, \pi_{k\neq j} = 0\}$ for some j where $p_j \geq p_k \quad \forall k$. In this case, condition (2.5) clearly holds. However, for condition (2.6) to hold we must have

$$p_j\rho_j - p_k\rho_k \geq 0 \quad \forall k. \tag{2.16}$$

This is never the case when one of the signal states is a pure state. The next most simple case to investigate is that of discriminating between two pure states.

2.3.1 Two States

The problem of minimum-error state discrimination between two states was solved by Helstrom [5], and before the present work (discussed in Chapter 4), was the only example of a set of states with a complete analytic, closed-form solution for the minimum probability of error for arbitrary prior probabilities. Here we have two states ρ_0, ρ_1, with respective probabilities p_0 and $p_1 = 1 - p_0$. The probability of correctly guessing the state based on the measurement outcome is given by:

$$\begin{aligned} \mathrm{P_{Corr}} &= p_0 P(\pi_0|\rho_0) + p_1 P(\pi_1|\rho_1) \\ &= p_0 \operatorname{Tr}(\rho_0\pi_0) + p_1 \operatorname{Tr}(\rho_1\pi_1). \end{aligned} \tag{2.17}$$

Substituting $\pi_0 = \mathbb{1} - \pi_1$ gives:

$$\mathrm{P_{Corr}} = p_0 - \operatorname{Tr}[(p_0\rho_0 - p_1\rho_1)\pi_1]. \tag{2.18}$$

This takes its maximum value when the latter term is minimised, i.e. when π_1 is a projector onto the negative eigenspace of the operator $p_0\rho_0 - p_1\rho_1$. We can make a similar argument with π_0 projecting onto the negative eigenspace of $p_1\rho_1 - p_0\rho_0$ (i.e. the positive eigenvalue of $p_0\rho_0 - p_1\rho_1$). As such, if we write $p_0\rho_0 - p_1\rho_1 = \lambda_+|\lambda_+\rangle\langle\lambda_+| + \lambda_-|\lambda_-\rangle\langle\lambda_-|$ we have:

$$\begin{aligned}
\mathrm{P}_{\mathrm{Corr}} &= p_0 + |\lambda_-| \\
\mathrm{P}_{\mathrm{Corr}} &= p_1 + |\lambda_+| \\
&= \frac{1}{2}(1 + \mathrm{Tr}\,|p_0\rho_0 - p_1\rho_1|),
\end{aligned} \tag{2.19}$$

where in the last line we have simply taken the average of the preceding two lines. As this is a two-state minimum-error problem, the optimal measurement is simply a von Neumann measurement of some description.

We can go further if we restrict to the pure-state case, and find the negative eigenvalue of $p_0|\psi_0\rangle\langle\psi_0| - p_1|\psi_1\rangle\langle\psi_1|$; if we now restrict to the qubit case and express (without loss of qubit generality) the states $|\psi_0\rangle$ and $|\psi_1\rangle$ as

$$\begin{aligned}
|\psi_0\rangle &= \cos\theta|0\rangle + \sin\theta|1\rangle \\
|\psi_1\rangle &= \cos\theta|0\rangle - \sin\theta|1\rangle,
\end{aligned} \tag{2.20}$$

then the eigenvalues of $p_0|\psi_0\rangle\langle\psi_0| - p_1|\psi_1\rangle\langle\psi_1|$ may be readily found:

$$\lambda_\pm = \frac{1}{2}(p_0 - p_1 \pm \sqrt{1 - 4p_0p_1\cos^2 2\theta}). \tag{2.21}$$

This gives us our optimal probability for guessing the signal state:

$$\begin{aligned}
\mathrm{P}_{\mathrm{Corr}} &= \frac{1}{2}(1 + \sqrt{1 - 4p_0p_1\cos^2 2\theta}) \\
&= \frac{1}{2}(1 + \sqrt{1 - 4p_0p_1|\langle\psi_0|\psi_1\rangle|^2}).
\end{aligned} \tag{2.22}$$

For $p_0 = p_1 = \frac{1}{2}$, the optimal measurement is to measure along the states

$$\begin{aligned}
|\phi_0\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\
|\phi_1\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle),
\end{aligned} \tag{2.23}$$

where $\pi_{0,1} = |\phi_{0,1}\rangle\langle\phi_{0,1}|$. However, as p_0 is increased, the measurement state $|\phi_0\rangle$ moves closer to the signal state $|\psi_0\rangle$ - see Figure 2.1 for an illustration of this.

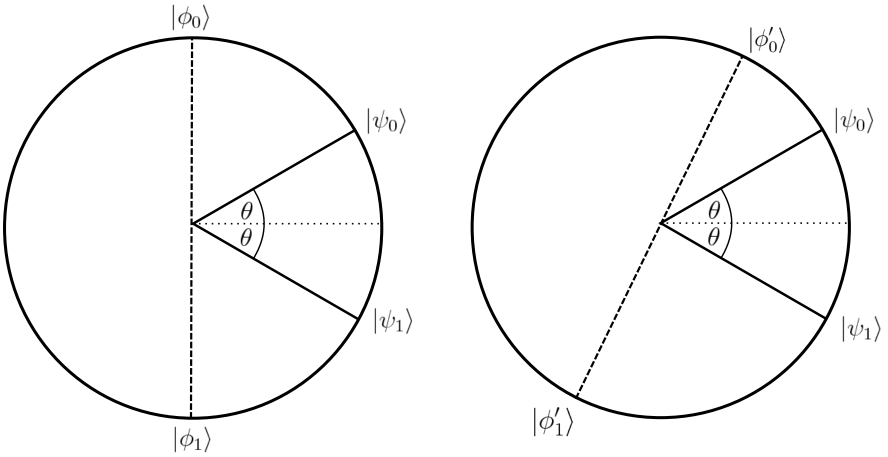

FIGURE 2.1: Two minimum-error measurements for the two states given in Equation (2.20), shown on the equator of the Bloch sphere. The image on the left shows the POVM given by $\pi_{0,1} = |\phi_{0,1}\rangle\langle\phi_{0,1}|$, which is optimal for $p_0 = p_1 = \frac{1}{2}$. As p_0 is increased, however, correctly identifying state ρ_0 becomes more important than identifying state ρ_1. In this situation, the optimal measurement more closely resembles that shown in the image on the right, given by $\pi'_{0,1} = |\phi'_{0,1}\rangle\langle\phi'_{0,1}|$. For some probability distribution $p_0 > p_1$, this POVM is optimal.

2.3.2 Square-Root Measurement

Another way to investigate the problem of minimum-error measurements is to start with a "standard" POVM and find a set of states which will combine with this POVM to satisfy the Helstrom conditions. To this end, we define the square-root measurement, sometimes also called the "pretty-good" measurement [51–55]. In this case, we have

$$\pi_j = p_j \rho^{-\frac{1}{2}} \rho_j \rho^{-\frac{1}{2}}, \tag{2.24}$$

where $\rho = \sum_i p_i \rho_i$. It is clear that these operators are positive and collectively sum to the identity, thus forming a POVM.

The square-root measurement is known to be the optimal minimum-error measurement for a number of cases: e.g., a set of symmetric pure states such that $|\psi_i\rangle = U^i|\psi_0\rangle$ with $U^n = \mathbb{1}$ [34] - this was later generalised to mixed states [56] and sets of states with larger symmetries [8] (this is further explained in §2.6).

A necessary and sufficient condition for the square-root measurement to be optimal for linearly independent pure states was derived by Sasaki *et al.* [57]; this condition states that the probability of correctly identifying the signal state must be independent of which state was sent.

Mochon generalised the concept of the square-root measurement to allow different weightings of POVM elements [55], which may be tweaked in order to find optimal solutions different discrimination problems involving a consistent set of states with varying prior probabilities.

2.3.3 Other cases

The problem of minimum-error quantum state discrimination has been the subject of research for over forty years; as such, an overview of the current state of affairs is appropriate. As we saw in §2.3.1, the problem of discriminating between two states (whether pure or mixed) was solved by Helstrom over 40 years ago [5]. For a set of equiprobable pure states in which a subset sums to the identity, Yuen *et al.* [29] found the optimal measurement - if, for some $a_i \geq 0$, we have $\sum_i a_i|\psi_i\rangle\langle\psi_i| = \mathbb{1}$, then $\pi_i = a_i|\psi_i\rangle\langle\psi_i|$ is an optimal measurement. Hunter later gave a solution for any set of equiprobable pure qubit states [58]. Andersson *et al.* [6] found the minimum-error measurement for the mirror-symmetric set of states $|\psi_0\rangle = |0\rangle$, $|\psi_{1,2}\rangle = \cos\theta|0\rangle \pm \sin\theta|1\rangle$ with probabilities $p_0 = 1 - 2p, p_{1,2} = p$ - this is explained further in §2.6.

Hunter showed [50] that sometimes it is optimal to not even measure, and to simply guess the *a priori* most likely state; this only occurs when the signal states are highly mixed.

A general solution for minimum-error quantum state discrimination - of some form - was found by applying the theory of semidefinite programming in order to find an efficient algorithm for solving such problems [59]. However, such an approach only yields numerical results, which is not ideal when - for instance - a quantum state discrimination problem arises as part of a larger problem, as seen in Chapters 5 & 6.

The first truly general solution for minimum-error qubit state discrimination (i.e. for arbitrary states – which need not be pure – with arbitrary prior probabilities) was given by Deconinck and Terhal [12], in which the dual problem is used to find a geometric solution. This was later generalised to qudit states by Tyson [60].

Later, Bae rewrote the Helstrom conditions in the form of the so-called KKT (Karush-Kuhn-Tucker) conditions [10, 61, 62], which also pointed towards a geometric solution for quantum state discrimination. Ha and Kwon later used this method to give a general qubit solution [11], which was also later extended to the qudit case [63]. Both of these had geometric components, and were fairly computationally complex.

As we will show in Chapter 3, it is possible to construct a relatively simple analytic solution for arbitrary qubit states with arbitrary prior probabilities; in this case, the problem simplifies to that of solving a series of linear equations.

2.4 Unambiguous Discrimination

We have already discussed unambiguous discrimination for qubits in §1.4.2.1, where we examined the advantages of POVM measurements over simple von Neumann measurements.

In this section, we simply note that this can be extended beyond the qubit case. However, as was shown by Chefles [16], the signal states must be linearly independent for this to be a viable option: the key idea here is that, for N signal states, each POVM element which identifies a state must make a measurement orthogonal to $N-1$ of the signal states, with the remaining one being the signal state which is identified. This means that, for instance, in the case where there are three qubit signal states, unambiguous discrimination is impossible.

It is also possible to perform unambiguous discrimination on mixed states, provided that the set of states to be discriminated have non-overlapping supports (i.e. the space spanned by the eigenvectors with non-zero eigenvalues for each state must not overlap with that of any other state in the ensemble). This was first proposed by Rudolph *et al.* [64], and performed experimentally soon after [37].

2.5 Maximum Confidence Measurements

While unambiguous discrimination is only viable for sets of states which are linearly independent, we may define an analogue for linearly dependent sets of states. Instead of identifying states with 100% certainty, as is the case with unambiguous discrimination, we can instead try to maximise the confidence that the measurement outcome we received correctly corresponds to the signal state [7]. That is, we wish to maximise

$$\text{Confidence} = \mathrm{P}(\rho_j|\pi_j) = \frac{\mathrm{P}(\rho_j)\mathrm{P}(\pi_j|\rho_j)}{\mathrm{P}(\pi_j)} = \frac{p_j\,\mathrm{Tr}(\rho_j\pi_j)}{\mathrm{Tr}(\rho\pi_j)} \tag{2.25}$$

for each j, where we have used Bayes' theorem and the Born Rule – equation (1.15) – to simplify, and $\rho = \sum_i p_i\rho_i$ as before. Note that this, like unambiguous discrimination, may necessitate the use of an inconclusive measurement outcome. This is in opposition to the minimum-error measurement strategy, where the goal is to correctly identify the signal state as often as possible; in this case an inconclusive outcome is clearly detrimental, as the probability of correctly identifying the signal state could always be increased by simply guessing state at random.

Note that, as the operator π_j appears in both the numerator and denominator of equation (2.25), we can only determine the POVM elements up to a constant of proportionality. We define $\pi_j = c_jM_j$ with $c_j \geq 0$ such that

$$\sum \pi_j \leq \mathbb{1}, \tag{2.26}$$

and define $\pi_? = \mathbb{1} - \sum_i \pi_i$ to form a complete measurement. When formulating the POVM elements for this measurement, therefore, we do not concern ourselves with the completeness condition, and instead consider each POVM element separately. This simplifies the problem to that of maximising equation (2.25) for each state in the set $\{\rho_i\}$; the inconclusive outcome described above then completes the set of POVM elements, if necessary.

We maximise the confidence $\mathrm{P}(\rho_j|\pi_j)$ by defining $\pi_j = c_j\rho^{-\frac{1}{2}}Q_j\rho^{-\frac{1}{2}}$, where Q_j is a positive, trace-one operator, the exact form of which we will discuss soon; this means that the probability of obtaining outcome π_j is simply given by c_j. Therefore

$$\begin{aligned}\mathrm{P}(\rho_j|\pi_j) &= p_j\,\mathrm{Tr}(\rho^{-\frac{1}{2}}\rho_j\rho^{-\frac{1}{2}}Q_j)\\ &= p_j\,\mathrm{Tr}(\rho_j\rho^{-1})\,\mathrm{Tr}(\rho_j'Q_j),\end{aligned} \tag{2.27}$$

where $\rho_j' = \rho^{-\frac{1}{2}}\rho_j\rho^{-\frac{1}{2}}/\,\mathrm{Tr}(\rho_j\rho^{-1})$. As Q_j and ρ_j' are both trace-one positive operators, they can be thought of as density operators. Thus, we can maximise $\mathrm{P}(\psi_j|\pi_j)$ by defining Q_j to be a projector onto the pure state which has the largest overlap with ρ_j'. That is, if ρ_j' has λ'_{max} as its largest eigenvalue, corresponding to the eigenket $|\lambda'_{max}\rangle$, then we have

$$Q_j = |\lambda'_{max}\rangle\langle\lambda'_{max}|, \tag{2.28}$$

and the maximum possible value for $\mathrm{P}(\psi_j|\pi_j)$ is $p_j \, \mathrm{Tr}(\rho_j \rho^{-1})$, given by the POVM element

$$\pi_j = c_j \rho^{-\frac{1}{2}} |\lambda'_{max}\rangle\langle\lambda'_{max}| \rho^{-\frac{1}{2}}, \tag{2.29}$$

which simplifies to

$$\pi_j \propto \rho^{-1} \rho_j \rho^{-1} \tag{2.30}$$

if the state ρ_j is pure.

The precise measurement which will take place is dependent on the choice of constants c_j - depending on the set of signal states, it is sometimes possible to choose our constants in such a way that no inconclusive outcome is necessary. In cases where the inclusion of an inconclusive outcome is inevitable, we may choose to minimise the probability of obtaining such an outcome.

This has been realised experimentally for a set of three symmetric states, with the results matching what was predicted [32]; these states and the associated maximum confidence measurement are shown in Figure 2.2. An overview of maximum confidence measurements may be found in [65].

2.6 Symmetries

In certain state discrimination problems, symmetry proves to be useful: if the signal states have a certain symmetry, the optimal POVM may also have the same symmetry. This is best shown with an example.

Consider the trine states, which may be represented by three equidistant points on any great circle of the Bloch sphere. We will place them on the equator, like so:

$$\begin{aligned}
|\psi_0\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\
|\psi_1\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + e^{i\frac{2\pi}{3}}|1\rangle) \\
|\psi_2\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + e^{i\frac{4\pi}{3}}|1\rangle).
\end{aligned}$$

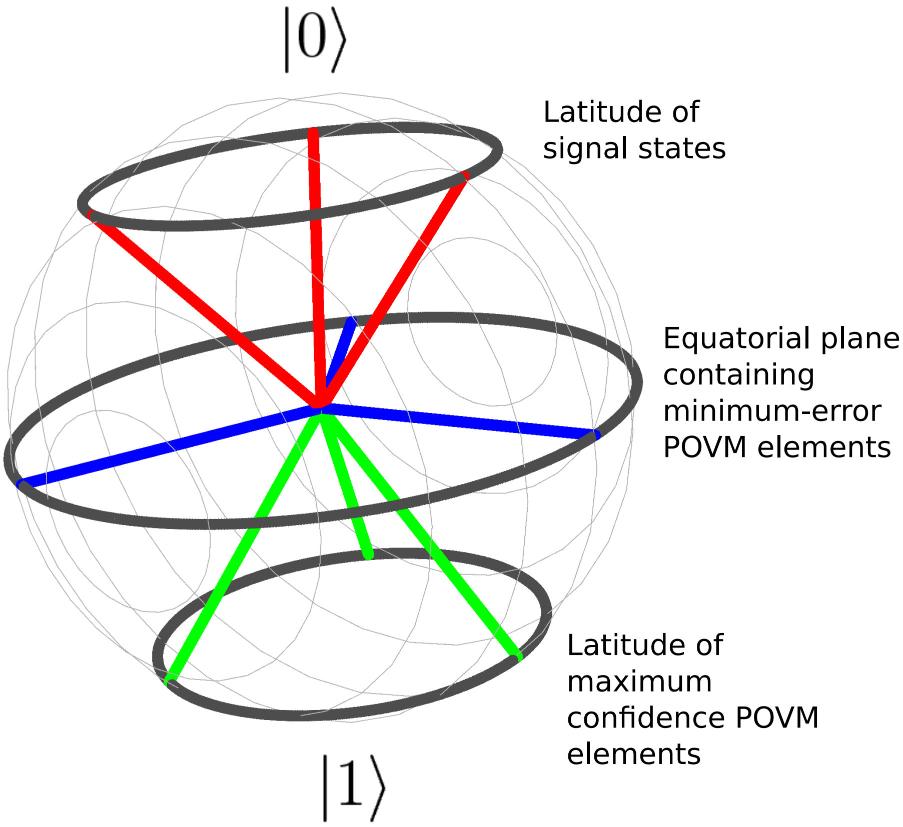
$|0\rangle$
Latitude of signal states
Equatorial plane containing minimum-error POVM elements
Latitude of maximum confidence POVM elements
$|1\rangle$

FIGURE 2.2: The Bloch sphere representation of states used in the first experiment demonstrating maximum confidence measurements. The red lines in the upper hemisphere show the three signal states. The blue lines on the equator show the minimum error POVM elements for this set of states, while the green lines in the lower hemisphere show the maximum confidence POVM elements for these signal states.

These will feature prominently in the work that follows. Note that these states are highly symmetrical: the unitary matrix $U = |0\rangle\langle 0| + e^{i\frac{2\pi}{3}}|1\rangle\langle 1|$ has the property that $U|\psi_i\rangle = |\psi_{i+1}\rangle$ for all $i \mod 3$. If these states are sent, each with probability $\frac{1}{3}$, then our probability of making an error in identifying the signal state is given by:

$$\begin{aligned} \mathrm{P_{Error}} &= \sum_i \sum_{j\neq i} p_i \langle\psi_i|\pi_j|\psi_i\rangle \\ &= 1 - \frac{1}{3}\sum_i \langle\psi_i|\pi_i|\psi_i\rangle, \end{aligned} \tag{2.31}$$

where we have used the completeness of POVM elements to simplify the final line. Now note that we may define two new sets of POVM elements, $\{\pi_i'\}$ and $\{\pi_i''\}$, with $\pi_i' = U^\dagger \pi_{i+1} U$ and $\pi_i'' = (U^\dagger)^2 \pi_{i+2} U^2$ ($i \mod 3$ in both cases). Note that these still satisfy the POVM condition because U is unitary. Clearly, these two sets of probability operators $\{\pi_i'\}$ and $\{\pi_i''\}$ are at least as good at discriminating the signal states as the original set $\{\pi_i\}$, as they must have the same probability of error:

$$\mathrm{P}_{\mathrm{Error}} = 1 - \frac{1}{3}\sum_i \langle\psi_i|\pi_i|\psi_i\rangle = 1 - \frac{1}{3}\sum_i \langle\psi_i|\pi_i'|\psi_i\rangle = 1 - \frac{1}{3}\sum_i \langle\psi_i|\pi_i''|\psi_i\rangle. \qquad (2.32)$$

Hence any linear combination of these will produce another measurement of the same efficacy. Our final set of probability operators will take this form: let $\tilde{\pi}_i = \frac{1}{3}(\pi_i + \pi_i' + \pi_i'')$, which has the property that $U\tilde{\pi}_i U^\dagger = \tilde{\pi}_{i+1}$ ($i \mod 3$), indicative of a cyclic symmetry of order 3 - similar to that of the trine states themselves (note that, as described in [33], we may always fine-grain this measurement so that it is rank one). Therefore the minimum-error measurement for such a system must measure three equidistant points on the equator of the Bloch sphere, i.e. $\tilde{\pi}_i = |\phi_i\rangle\langle\phi_i|$, where $|\phi_i\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i(\frac{2\pi}{3}+\phi)}|1\rangle)$ for some angle ϕ. Note that this only works in the case where $p_i = \frac{1}{3}$ - otherwise, equations (2.31) and (2.32) do not hold. This type of simplification of the problem only works, therefore, in cases where the states and their probabilities are in some way symmetric.

The symmetry also does not need to be cyclic in nature: a mirror-symmetric set of three signal states with probabilities $p_0 = p, p_1 = p, p_2 = 1 - 2p$ is investigated in a paper by Andersson, et. al. [6], and the minimum-error POVM is found to have a similar mirror symmetry. Such a collection of states $\mathcal{S}$ - where $\mathcal{S} = \{|\psi_i\rangle = U_i|\psi\rangle, U_i \in \mathcal{G}\}$, where $\mathcal{G}$ is a finite abelian group of unitary matrices U_i - are called geometrically uniform [66]. This can be extended to the case of multiply symmetric states, where two unitary matrices, U and V, are necessary to describe the symmetries of the signal states, for instance an entangled two-qubit state [8]. The difference between these and geometrically uniform state sets is that U and V need not commute.

2.7 Experiments

Quantum state discrimination has been performed in laboratories in a number of ways. Here we detail a few implementations of the strategies we have discussed. While it is possible for a qubit to be realised in a number of ways, in all of the examples discussed here the qubit will manifest itself in the polarisation of light. That is, e.g., using the transformation $|0\rangle \rightarrow |H\rangle, |1\rangle \rightarrow |V\rangle, |i\rangle \rightarrow |L\rangle, |-i\rangle \rightarrow |R\rangle, |+\rangle \rightarrow |D\rangle, |-\rangle \rightarrow |A\rangle$.

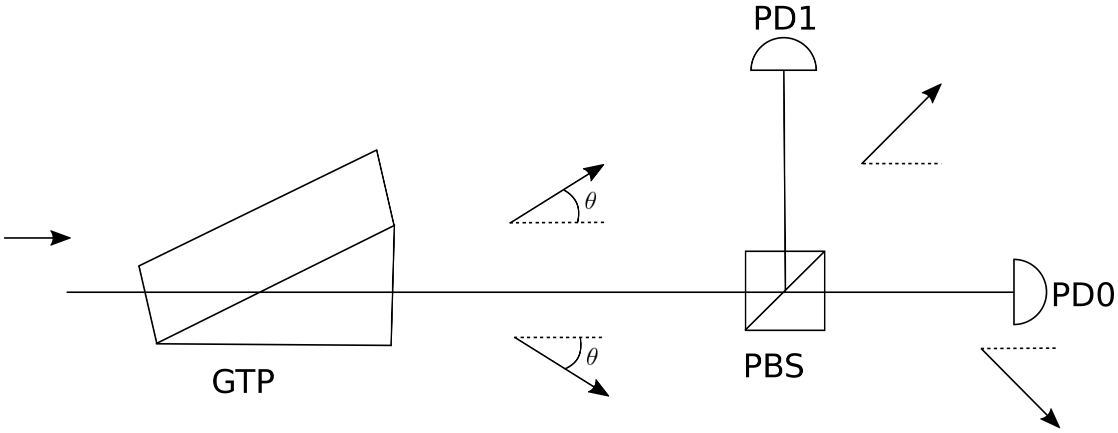

Figure 2.3: The set-up for the first experimental demonstration of the Helstrom bound. The Glan-Thompson polariser (GTP) selectively transmits/reflects the incoming beam to create the states described in equation (2.20). The polarising beam-splitter (PBS) then separates these states in the $\pm\frac{\pi}{4}$ basis, so a $|+\rangle$ photon is reflected and detected at photon detector 1 (PD1), while a $|-\rangle$ photon is detected at photon detector 0 (PD0).

2.7.1 Minimum-error discrimination

2.7.1.1 Helstrom bound

In [67], Barnett and Riis demonstrated the optimal measurement to discriminate between two equiprobable qubit states of the form shown in equation (2.20) for various values of θ.

This measurement was implemented using a polarising beam splitter at an angle of $\frac{\pi}{4}$ to the horizontal. This measurement is shown in Figure 2.3. The source was an attenuated pulsed laser which on average produced 0.1 photons/pulse. These photons then passed through a Glan-Thompson polariser to produce the input states. The polarising beam splitter then transmitted photons in the $|+\rangle$ basis, while reflecting those in the $|-\rangle$ basis, meaning photons of the form $|\psi_{0,1}\rangle$ had a probability $|\langle\psi_{0,1}|+\rangle|^2$ of being transmitted and hitting the photon detector which represented π_0. This set-up therefore realises the Helstrom measurement detailed in equation (2.23). If the input state $|\psi_0\rangle$ reached the photon detector representing π_0, then the state was identified correctly. If not, the measurement was incorrect. The reverse is true for input state $|\psi_1\rangle$. In the referenced experiment, the probability of correctly identifying the signal state given by equation (2.22) was verified to within a few percent.

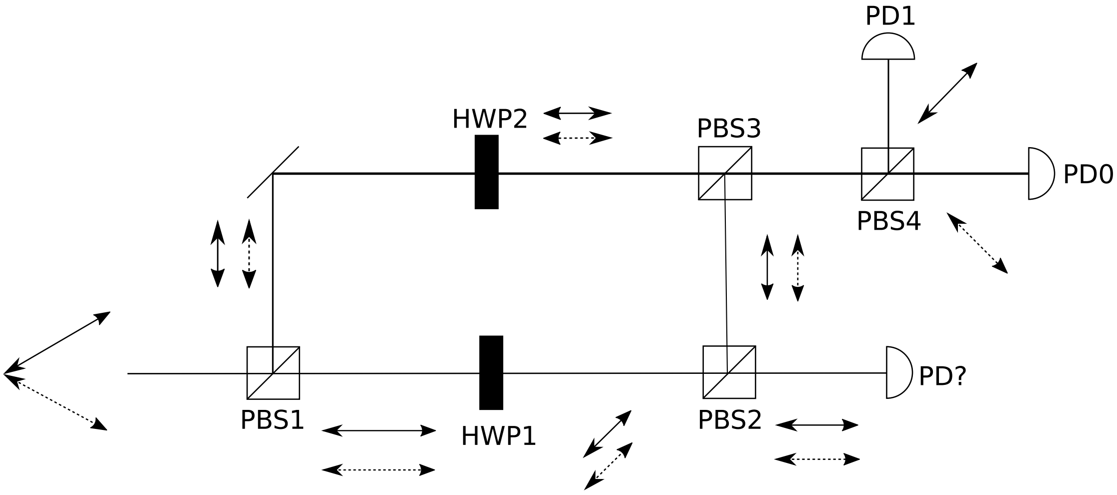

FIGURE 2.4: Schematic detailing the experimental set-up used by Clarke *et al.* [1] for unambiguous state discrimination of the equiprobable states $|\psi_0\rangle, |\psi_1\rangle$. Here, PD0 corresponds to the POVM element π_0, PD1 corresponds to π_1, and PD? corresponds to $\pi_?$ - however, in the analogous experiment where a similar set-up is used for minimum-error discrimination of the trine states, PD? instead corresponds to π_2.

2.7.1.2 Trine and tetrad states

Minimum-error discrimination of the trine states with equal prior probabilities has also been demonstrated, in [36]. In this example, the experimental set-up was identical to that described in Figure 2.4, but with the outcome corresponding to $\pi_?$ (PD? in Figure 2.4) instead corresponding to π_2. This measurement corresponds to the trine measurement, discussed more in §4.1.

In the same experiment (with some minor changes to the waveplates), the minimum-error discrimination strategy for the so-called tetrad states was demonstrated. This set of states was also prepared with equal prior probabilities. In this ensemble, the states are represented by

$$\begin{aligned}
|\psi_0\rangle &= |0\rangle \\
|\psi_1\rangle &= \frac{1}{\sqrt{3}}(-|0\rangle + \sqrt{2}e^{i\frac{2\pi}{3}}|1\rangle) \\
|\psi_2\rangle &= \frac{1}{\sqrt{3}}(-|0\rangle + \sqrt{2}e^{i\frac{4\pi}{3}}|1\rangle) \\
|\psi_3\rangle &= \frac{1}{\sqrt{3}}(-|0\rangle + \sqrt{2}|1\rangle).
\end{aligned} \tag{2.33}$$

2.7.2 Unambiguous state discrimination

The first experimental demonstration of unambiguous discrimination of two pure states was given by Huttner *et al.* [38], in which an optical fibre with polarisation-dependent absorption was used to selectively absorb the horizontal polarisation of input states. The length of fibre was chosen so that this horizontal component was reduced by a factor of $\tan\theta$. As a result, the state $|\psi_{0,1}\rangle$ will be absorbed with probability $(1 - \tan^2\theta|\langle H|\psi_{0,1}\rangle|^2)$; hence the input states are transformed such that $|\psi_{0,1}\rangle \rightarrow \sin\theta(|0\rangle \pm |1\rangle)$. These states are orthogonal and may be perfectly discriminated using a properly-oriented polarising beam splitter. The downside of this approach is that the lost photons – which correspond to the inconclusive result $\pi_?$ – are not registered, and so the number of inconclusive outcomes cannot be measured. However, the surviving photons were shown to give unambiguous results to within 1.7% accuracy.

Huttner *et al* also suggested an alternative implementation which was realised by Clarke *et al* [1]. In this implementation – shown in Figure 2.4 – a polarising beam splitter separates the input states into horizontal and vertical component. A half wave plate (HWP1) is then used to rotate the horizontal polarisation component in such a way that $(1 - \tan^2\theta)^{\frac{1}{2}}$ of the transmitted photons pass through PBS2 and end up at the photon detector representing $\pi_?$. The remaining horizontal component – which is now vertically polarised due to the actions of HWP1 and PBS2 – recombine with the photons which were transmitted at PBS1 in such a way that PBS4 may perfectly discriminate between them. These photons are therefore detected at the photon detectors corresponding to π_0 and π_1 with – in principle – perfect efficiency. In reality, the experiment performed to within 1% accuracy of the IDP limit for a variety of values of θ. The central machinations of this experiment were identical to those of Huttner *et al* in [38], but with the key difference that $\pi_?$ was physically realised, instead of simply corresponding to losses in optical fibre.

Unambiguous state discrimination has also been shown between three linearly independent pure states, as shown in [37]. This uses the set-up proposed by Sun *et al* [68], in which a multi-rail optical system is used. Note that one of the input states for this was also a mixed state, verifying the proposal put forward in [64].

2.7.3 Maximum confidence measurement

Maximum confidence measurement has been demonstrated for a set of three states which are a generalisation of the trine states, of the form

$$\begin{aligned}|\psi_0\rangle &= \cos\theta|0\rangle + \sin\theta|1\rangle \\ |\psi_1\rangle &= \cos\theta|0\rangle + e^{i\frac{2\pi}{3}}\sin\theta|1\rangle \\ |\psi_2\rangle &= \cos\theta|0\rangle + e^{i\frac{4\pi}{3}}\sin\theta|1\rangle. \qquad (2.34)\end{aligned}$$

The experiment, detailed in [32] and [65], was performed for ten values of θ, equally spaced between 0 and $\frac{\pi}{4}$. The results varied with θ, with the inconclusive outcome occurring with high frequency for $\theta < 10°$ due to the high overlap between the signal states. However, the experiment demonstrated a clear advantage over the minimum-error strategy in the confidence given by measurement outcomes for the range $10° \leq \theta \leq 30°$. For θ above this range, the two strategies offer very similar levels of confidence.

Chapter 3

Minimum-error discrimination of arbitrary single-qubit mixed states

3.1 Introduction

As we have seen in §2.3.3, geometric solutions to the problem of minimum-error discrimination for arbitrary qubit states with arbitrary prior probabilities have been found. However, a concise analytic solution is missing in the literature.

In this chapter, we give an alternative method of constructing optimal measurements from the minimum error conditions. Previous work [10] has demonstrated that finding a single operator Γ, sometimes referred to as the Lagrange operator, is equivalent to solving the minimum error discrimination problem: the trace of this operator gives the optimum probability of success, and optimal measurements may be readily constructed once it is known. We construct linear constraints on this operator and its inverse, which in the qubit case may be readily solved for Γ and thereby the optimal measurement [39]. Our algebraic approach is complementary to the geometric approach already discussed.

3.2 The minimum-error conditions

Recall that the Born rule, expressing the probability of obtaining outcome j in a measurement on a system prepared in state ρ is given by:

$$\mathrm{P}(j|\rho) = \mathrm{Tr}(\rho\pi_j). \tag{3.1}$$

For a minimum-error detection strategy, a "click" at the detector corresponding to element π_j is taken to indicate that the state ρ_j was sent. Bob's probability of correctly guessing the state Alice sent is then given by $\mathrm{P}_{Corr} = \sum_{i=0}^{n-1} p_i \, \mathrm{Tr}(\pi_i\rho_i)$, and it is this that we wish to maximise in the minimum error problem (Bob's probability of error, of course, is given by $1 - \mathrm{P}_{\mathrm{Corr}}$).

As we have seen, the solution to the problem of minimum-error quantum state discrimination is equivalent to finding a POVM satisfying the conditions [5, 29, 69]:

$$\begin{aligned} \Gamma - p_j\rho_j &\geq 0 \quad \forall j, && (3.2) \\ \pi_i(p_i\rho_i - p_j\rho_j)\pi_j &= 0 \quad \forall i, j, && (3.3) \end{aligned}$$

where $\Gamma = \sum_i p_i\rho_i\pi_i$. The first condition is both necessary and sufficient for $\{\pi_i\}$ to describe an optimal measurement procedure, and we note that the conditions are not independent: the second, which is necessary but not sufficient, follows from the first. It is useful however to give both conditions, as often the second is more convenient to use in practice. Note that Γ is a Hermitian operator $\Gamma = \Gamma^\dagger = \sum_i p_i\pi_i\rho_i$, which follows from condition (3.2), and may be seen explicitly by summing over both i and j in condition (3.3). An alternative condition is obtained by summing over i in equation (3.3), giving:

$$(\Gamma - p_j\rho_j)\,\pi_j = 0. \tag{3.4}$$

This is a necessary (but not sufficient) condition on any optimal measurement $\{\pi_j\}$, and is central to our and other methods [9, 11], allowing us to construct operators π_j satisfying $\Gamma = \sum_i p_i\rho_i\pi_i$ once a candidate Γ is given. Indeed, both π_j and $\Gamma - p_j\rho_j$ (according to inequality (3.2)) are positive operators, and thus equation (3.4) can hold only if they are orthogonal, that is π_j is entirely within the kernel (or the eigensubspace corresponding to zero eigenvalue) of $\Gamma - p_j\rho_j$. It follows that π_j can be non-zero only if

$\Gamma - p_j\rho_j$ has at least one zero eigenvalue. We further note that

$$\mathrm{Tr}(\Gamma) = \sum_{i=0}^{n-1} p_i \, \mathrm{Tr}(\pi_i \rho_i) = \mathrm{P}_{\mathrm{Corr}}. \tag{3.5}$$

Therefore if we can find Γ, we find both the optimal probability of success, and a way of constructing the optimal measurement operators. The problem of finding the optimal measurement $\{\pi_i\}$, a set of n operators, is thus equivalent to finding a single positive operator Γ satisfying the condition (3.2) and from which operators $\{\pi_j\}$ satisfying (3.4) and forming a POVM can be constructed. Indeed the so-called dual problem in the semi-definite programming approach consists of finding the operator Γ with minimum trace that satisfies condition (3.2) for all j. Further, as is stressed by Bae [10], the operator Γ is unique for a given set of states, while the optimal measurement may not be - for example, in the case of $N \geq 4$ symmetric states [28, 55, 58].

3.3 Qubit state discrimination

There has recently been much progress in using the Helstrom conditions constructively, detailed in §2.3.3. Most pertinent to our work is recent work by Bae [10] which used the (so-called) Karush-Kuhn-Tucker, or KKT, conditions [61, 62] from semi-definite programming. These KKT conditions are necessary conditions which are used in optimisation problems, and may be used to define complementary states $\{\sigma_j\}$ with weights r_j such that

$$\Gamma = p_i\rho_i + r_i\sigma_i = p_j\rho_j + r_j\sigma_j. \tag{3.6}$$

This may be seen from equation (3.4), which must be true for all values of j. σ_j lies in the kernel of π_j, and so $\Gamma - p_j\rho_j = r_j\sigma_j$ for some weighting r_j. The geometric structure of the complementary states σ_j may be deduced from the conditions and the geometric structure of the signal states ρ_j, and in turn used to construct Γ. Bae discusses the qubit case, in which the Bloch sphere provides a convenient geometric picture, and the full details for three mixed qubit states were later calculated by Ha and Kwon [11].

We begin with some general considerations concerning the qubit state discrimination problem, and then discuss our method, which constructs Γ directly, without reference to complementary states [39]. Firstly, we note that for each j the operator $\Gamma - p_j\rho_j$ can have two, one, or no zero eigenvalues, corresponding to the zero operator, a rank-one operator, and a positive-definite operator respectively:

1. If $\Gamma - p_j\rho_j = 0$ for some j, then $\Gamma = p_j\rho_j$, which can only hold if $p_j\rho_j - p_k\rho_k \geq 0$ for all k. The no measurement strategy is then an optimal measurement, $\pi_k = I\delta_{jk}$ [50].

2. $\Gamma - p_j\rho_j$ has a single zero eigenvalue. If π_j is non-zero, it is a weighted projector onto the corresponding eigenstate.

3. If $\Gamma - p_j\rho_j$ is positive definite (all eigenvalues strictly greater than zero), then according to condition (3.4) it follows that $\pi_j = 0$ for every optimal measurement, and the corresponding state is never identified.

Given a set of qubit states $\{\rho_j\}$ with *a priori* probabilities p_j, it is easily checked whether for some j

$$p_j\rho_j - p_k\rho_k \geq 0, \quad \forall k. \tag{3.7}$$

If this does hold for some j, the optimal strategy is not to measure at all and simply guess ρ_j. For all other ensembles, it follows that the optimal measurement is made up of rank-one weighted projectors,

$$\pi_j = c_j|\phi_j\rangle\langle\phi_j| \tag{3.8}$$

for some c_j satisfying $0 \leq c_j \leq 1$, and where $|\phi_j\rangle$ is the eigenstate of $\Gamma - p_j\rho_j$ corresponding to the zero eigenvalue. Note that this is completely general for qubits, and holds whether ρ_j are pure or mixed states. Thus, for an optimal measurement each operator π_j is uniquely defined, up to a multiplying factor. There may however be more than one way of choosing the coefficients c_j such that the π_j thus found sum to the identity.

Secondly, we note that for minimum error discrimination of an arbitrary set of qubit states there always exists an optimal measurement with at most four outcomes. Intuitively, the constraint $\sum_i \pi_i = \mathbb{1}$ contains only d^2 independent linear constraints, where d is the dimension of our space: if a set of $N > d^2$ elements $\{\pi_j\}$ satisfies this, there is always a subset of these which, when appropriately weighted, also forms a resolution of the identity. A measurement with more than d^2 outcomes can always be decomposed as a probabilistic mixture of measurements with at most d^2 outcomes. If the mixture results in an optimal procedure, then any of the component measurements must also be optimal [12, 58].

Finally, note that the number of outcomes in our optimal measurement corresponds to the number of states that are identified with non-zero probability by the measurement: additional states are never identified. Denoting the number of outcomes k, the cases $k = 1$ and $k = 2$ are well-known, as these correspond to the no-measurement strategy and the Helstrom two-state discrimination measurement respectively [4, 5]. In the cases $k = 3$ and $k = 4$ it is more difficult to find optimal measurements although, as discussed above, strategies for these cases have been recently suggested.

For qubits, the Pauli operators together with the identity form a convenient basis in which to express any operator on the space. Thus, for example, we can write

$$\Gamma = \frac{1}{2}(a\mathbb{1} + \vec{b} \cdot \vec{\sigma}), \tag{3.9}$$

where $a > 0$, $\vec{b}$ is a real three-dimensional vector, and $\vec{\sigma}$ is the vector of Pauli operators: $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$. It will be convenient in what follows to also use such a representation for the inverse Γ^{-1}, and it is easily verified that:

$$\Gamma^{-1} = \frac{2}{a^2 - |b|^2}(a\mathbb{1} - \vec{b} \cdot \vec{\sigma}). \tag{3.10}$$

This may be seen with the aid of the identity

$$(\vec{x} \cdot \vec{\sigma})(\vec{y} \cdot \vec{\sigma}) = (\vec{x} \cdot \vec{y})\mathbb{1} + i(\vec{x} \times \vec{y}) \cdot \sigma, \tag{3.11}$$

as

$$\begin{aligned}
\Gamma \cdot \Gamma^{-1} &= \frac{1}{a^2 - |b|^2}(a\mathbb{1} + \vec{b} \cdot \vec{\sigma})(a\mathbb{1} - \vec{b} \cdot \vec{\sigma}) \\
&= \frac{1}{a^2 - |b|^2}[a^2\mathbb{1} + a\vec{b} \cdot \vec{\sigma} - a\vec{b} \cdot \vec{\sigma} - (\vec{b} \cdot \vec{\sigma})(\vec{b} \cdot \vec{\sigma})] \\
&= \frac{1}{a^2 - |b|^2}(a^2 - |b|^2)\mathbb{1} \\
&= \mathbb{1}.
\end{aligned} \tag{3.12}$$

Note that Γ is a strictly positive operator in the space spanned by the states to be discriminated, and so the inverse is always well-defined, as is its square-root, which we will use later. We thus need 4 parameters to completely specify Γ, and we discuss now how to construct 4 constraints, which are readily inverted to construct Γ.

3.4 Constructing Γ

We begin by proving a result found in the exercises of [70]. Suppose we have a positive Hermitian operator A of rank n with some decomposition $A = \sum_{i=1}^{n} a_i |\alpha_i\rangle\langle\alpha_i|$, where $\{|\alpha_i\rangle\}$ are linearly independent and $\{a_i\}$ are real. We may therefore write

$$\begin{aligned}\mathbb{1} &= A^{-\frac{1}{2}} A A^{-\frac{1}{2}} \\ &= \sum_{i=1}^{n} a_i A^{-\frac{1}{2}} |\alpha_i\rangle\langle\alpha_i| A^{-\frac{1}{2}} \\ &= \sum_{i=1}^{n} |\phi_i\rangle\langle\phi_i|\end{aligned}$$

which can only hold if $\{|\phi_i\rangle\}$ are an orthonormal set. Hence, for some choice of indices, $|\phi_i\rangle = \sqrt{a_i} A^{-\frac{1}{2}} |\alpha_i\rangle$. This gives the result

$$\frac{\langle\phi_i|\phi_i\rangle}{a_i} = \langle\alpha_i|A^{-1}|\alpha_i\rangle = \frac{1}{a_i}, \tag{3.13}$$

which we will use to impose constraints on Γ.

Suppose there is an optimal measurement which identifies $k > 2$ states: we will show that for each of these we may obtain one constraint on the parameters of Γ. It is of course not obvious *a priori* which states will be identified by an optimal measurement, however we can construct a candidate Γ, under the assumption that a particular subset of our states are identified in an optimal measurement, and then verify that this results in a physically allowed measurement procedure. We will return to this later. According to the discussion above therefore, for each of these k states the operator $\Gamma - p_j\rho_j$ has a single zero eigenvalue. Let us consider first the pure state case: $\rho_j = |\psi_j\rangle\langle\psi_j|$. The

KKT conditions (3.6) show that

$$\Gamma = p_j|\psi_j\rangle\langle\psi_j| + r_j|\phi_j^\perp\rangle\langle\phi_j^\perp|, \tag{3.14}$$

where $|\phi_j^\perp\rangle\langle\phi_j^\perp| = \sigma_j$, the complementary state from the KKT conditions. As $|\psi_j\rangle$ and $|\phi_j^\perp\rangle$ are linearly independent, equation (3.13) is applicable, giving:

$$p_j\langle\psi_j|\Gamma^{-1}|\psi_j\rangle = 1. \tag{3.15}$$

A similar relation was pointed out by Mochon [55], who discussed the inverse problem of characterising the sets of states and corresponding probabilities for which a given measurement procedure was optimal, although it does not seem to have been used constructively in the literature. Thus we find

$$2p_j\langle\psi_j|(a\mathbb{1} - \vec{b}\cdot\vec{\sigma})|\psi_j\rangle = a^2 - |b|^2. \tag{3.16}$$

Alternatively, if $\rho_j = |\psi_j\rangle\langle\psi_j|$ has Bloch vector $\hat{r}_j$ (a unit vector as ρ_j is a pure state): $\rho_j = \frac{1}{2}\left(\mathbb{1} + \hat{r}_j\cdot\vec{\sigma}\right)$, we may write:

$$2p_j\left(a - \hat{r}_j\cdot\vec{b}\right) = a^2 - |b|^2. \tag{3.17}$$

Each state gives rise to one such constraint, resulting in k independent constraints on Γ.

It is not obvious how to extend this result to the case of mixed qubit states; however, a little trick leads to a nice result which allows us to do so. Note that for qubit states, every mixed state can be written as a mixture of a pure state and the identity: $\rho_j = \alpha_j|\psi_j\rangle\langle\psi_j| + \beta_j\frac{1}{2}\mathbb{1}$, where $\alpha_j + \beta_j = 1$. Condition (3.14) then becomes:

$$\Gamma - \frac{1}{2}p_j\beta_j\mathbb{1} = p_j\alpha_j|\psi_j\rangle\langle\psi_j| + r_j|\phi_j^\perp\rangle\langle\phi_j^\perp|, \tag{3.18}$$

and using the same reasoning as previously, we obtain:

$$p_j\alpha_j\langle\psi_j|\left(\Gamma - \frac{1}{2}p_j\beta_j\mathbb{1}\right)^{-1}|\psi_j\rangle = 1. \tag{3.19}$$

Explicitly, this gives:

$$2p_j\alpha_j\langle\psi_j|[(a-p_j\beta_j)\mathbb{1}-\vec{b}\cdot\vec{\sigma}]|\psi_j\rangle=(a-p_j\beta_j)^2-|b|^2, \tag{3.20}$$

and after a litte rearranging, again writing $|\psi_j\rangle\langle\psi_j|=\frac{1}{2}\left(\mathbb{1}+\hat{r}_j\cdot\vec{\sigma}\right)$, we find

$$2p_j\left[a-\alpha_j\hat{r}_j\cdot\vec{b}-p_j\beta_j(\alpha_j+\frac{1}{2}\beta_j)\right]=a^2-|b|^2. \tag{3.21}$$

If there are k states identified by the optimal measurement this procedure, in both the pure state and mixed state case, gives k equations for the parameters of Γ. Clearly if $k=4$ this is enough to construct Γ. We further note that in equations (3.17) and (3.21) the non-linear right hand side is independent of j, thus we can easily take linear combinations to obtain $k-1$ linear equations. For $k=4$ these are readily solved to write all parameters in terms of a single one, e.g. a, which is finally determined by solving one quadratic equation.

Thus we can construct optimal measurements with $k=1,2$, or 4 outcomes. For $k=3$ we don't yet have enough constraints to determine Γ; a further constraint however, is readily constructed, as we now discuss. We first note that for the special case in which three signal states lie in an equatorial plane of the Bloch sphere (as in [6]), we know from symmetry that the POVM elements, and therefore also Γ, must lie in the same plane as the signal states, thus giving us our final constraint. More generally, for the case of three equiprobable pure qubit states it is always possible to choose a representation in which the states sit at the same latitude of the Bloch sphere. The optimal measurement operators π_j then lie in the equator of the sphere, and Γ has the same latitude as the signal states [58].

We can generalize this idea to both pure and mixed states, and to non-equal prior probabilities. We first note that for a three outcome measurement, all three elements of the POVM must lie on a great circle of the Bloch sphere in order to form a resolution of the identity. Without loss of generality we choose our axes so that this is the $z=0$ plane. That is, we can always choose our axes so that $\pi_j=\frac{1}{2}\left(c_j\mathbb{1}+\vec{d_j}\cdot\vec{\sigma}\right)$, with $d_{jz}=0$, $\forall j$. Referring now to condition (3.4), it follows that

$$\Gamma-p_j\rho_j\propto\frac{1}{2}\left(c_j\mathbb{1}-\vec{d_j}\cdot\vec{\sigma}\right),$$

and thus $\langle \Gamma\sigma_z \rangle - p_j \langle \rho_j \sigma_z \rangle = 0$. Finally we therefore require

$$b_z = p_j \langle \rho_j \sigma_z \rangle = p_j \alpha_j \hat{r}_{jz}, \quad \forall \quad j,$$

where as before $\rho_j = \frac{1}{2}\left(\mathbb{1} + \alpha_j \hat{r}_j \cdot \vec{\sigma}\right)$. Thus if we define our z-axis to be such that the z-component of $p_j \rho_j$ is the same for each of the three signal states identified, then Γ also has the same z-component, and the optimal measurement operators lie in the equatorial plane. Note that a similar discussion may be found in [71].

Thus for a given set of qubit states $\{\rho_i\}$ with arbitrary priors $\{p_i\}$, if there exists an optimal minimum error measurement which identifies a subset of $k = 1, 2, 3, 4$ of these states, we have shown how to construct Γ, which in turn allows us to construct the optimal measurement. We illustrate below in an example how this may be employed in practice to find optimal measurements, and discuss later the problem of how we can know in general which states are identified by an optimal measurement.

3.5 Examples

3.5.1 Pure state example

To illustrate our method, we consider the problem of discriminating between three pure states which are mirror-symmetrically arranged on the equator of the Bloch sphere, previously investigated by Andersson, et. al. [6]. The states are:

$$\begin{aligned} |\psi_0\rangle = |+\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \\ |\psi_1\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + e^{i\theta}|1\rangle), \\ |\psi_2\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + e^{-i\theta}|1\rangle), \end{aligned}$$

and these occur with *a priori* probabilities $p_0 = 1 - 2p$, $p_1 = p_2 = p$, with $p \in [0, \frac{1}{2}]$. The so-called trine ensemble occurs at $\theta = \frac{2\pi}{3}$ [34, 72].

We begin by noting that as the states are all pure it is not possible to satisfy conditions (3.7) and the no-measurement solution is never optimal. We next check to see when a two-outcome measurement is optimal. Note that, due to the symmetry (as discussed in §2.6), the only sensible two-outcome measurement is one distinguishing $|\psi_1\rangle$ and $|\psi_2\rangle$: the optimal such measurement is a projective measurement in the eigenbasis of σ_y; $\pi_i = |\phi_i\rangle\langle\phi_i|$, where $|\phi_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$ and $|\phi_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)$. It is straightforward to calculate $\Gamma_{\text{ 2-element}}$, and we note that condition (3.2) is satisfied for $j = 1, 2$ by construction. In order to check this condition for $j = 0$, it is enough to verify that $\det(\Gamma_{\text{ 2-element}} - p_0\rho_0) \geq 0$, as demonstrated by the proof of necessity of the Helstrom conditions in §2.3. We find, as in [6], that this holds when

$$p \geq \frac{1}{2 + \cos(\frac{\theta}{2})[\cos(\frac{\theta}{2}) + \sin(\frac{\theta}{2})]}. \tag{3.22}$$

The corresponding optimal probability of correctly identifying the state is given by

$$\mathrm{Tr}(\Gamma_{\text{2-element}}) = p(1 + \sin\theta).$$

When condition (3.22) does not hold, we know that a three outcome measurement is optimal, and can use the method outlined above to find this. We first note that $\hat{r}_z = 0$ for each of our signal states. Thus, as discussed above, Γ must also have $b_z = 0$, and lies in the equatorial plane. Further, using equation (3.17), we obtain the following three constraints on a, b_x, and b_y:

$$a^2 - |b|^2 = 2(1 - 2p)(a + b_x)$$
$$a^2 - |b|^2 = 2p(a + b_x\cos\theta + b_y\sin\theta)$$
$$a^2 - |b|^2 = 2p(a + b_x\cos\theta - b_y\sin\theta)$$

It is clear from the latter two that $b_y = 0$. The remaining equations are readily solved for a and b_x, giving:

$$\begin{aligned} a &= b_x\frac{p\sin^2\frac{\theta}{2} + 1 - 2p - p\cos^2\frac{\theta}{2}}{3p - 1} \\ b_x &= \frac{(3p-1)(1-2p)}{1 - 2p - p\cos^2\frac{\theta}{2}} \end{aligned} \tag{3.23}$$

The corresponding probability of correctly identifying the state P_{Corr} is then given by:

$$\begin{aligned} \mathrm{P}_{\text{Corr}} &= \mathrm{Tr}(\Gamma_{\text{3-element}}) = a \\ &= \frac{(1-2p)(p\sin^2\frac{\theta}{2} + 1 - 2p - p\cos^2\frac{\theta}{2})}{1 - 2p - p\cos^2\frac{\theta}{2}} \end{aligned}$$

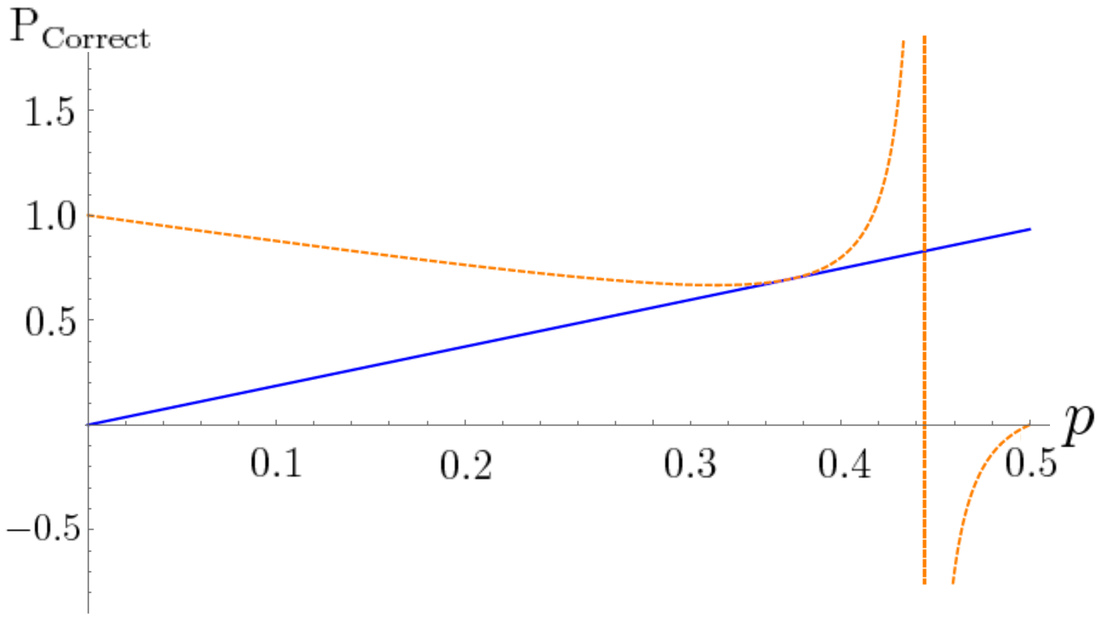

FIGURE 3.1: The two functions we obtained for P_{Corr} plotted against p for the optimal two-element (solid line) and three-element (dotted line) POVMs - we can see that for $p > 0.373$, the three-element POVM appears to be superior to the two-element POVM. However, this turns out to no longer be physically realisable, and fails to satisfy the condition in equation (3.2). Note that the function we find for $\text{Tr}(\Gamma_{\text{3-element}})$ is not strictly positive - at no point have we assumed that Γ must be positive. This plot is designed to illustrate the limitations of our method by showing that the functions we obtain for P_{Corr} do not always give sensible answers; if we are in a region of parameter space where a two-element POVM is optimal, $\text{Tr}(\Gamma_{\text{3-element}})$ may be greater than 1, or may even be negative. This is shown on the right-hand side of the graph, where $p > 0.373$.

which agrees with the solution provided in [6].

We finally note that we found the region in which a three outcome measurement was necessary by first finding the region in which a two-outcome measurement was optimal. If we use our method to find a candidate Γ in the region where in fact the optimal measurement has only two outcomes, we find that even though it is possible to construct Γ, it is not possible to construct a physically allowed measurement from the conditions (3.4), and the method fails. Further, it can sometimes return probabilities that are greater than 1, clearly indicating that something has gone wrong. This is illustrated in Figure 3.1. As we will see in §4.2.2, this is a result of the optimal three-element POVM including an element of the form $\pi_0 = [1 - (\frac{\sqrt{3}p}{4-9p})^2]|\psi_0\rangle\langle\psi_0|$, which is a negative operator for $p > \frac{4}{9+\sqrt{3}}$, breaking the condition of positivity we require for POVM elements.

3.5.2 Mixed state example

Consider an ensemble consisting of the trine states, each with the same probability and degree of "mixedness" α. That is,

$$\rho_0 = \frac{1}{2}\begin{bmatrix} 1 & \alpha \\ \alpha & 1 \end{bmatrix}$$
$$\rho_1 = \frac{1}{2}\begin{bmatrix} 1 & \alpha e^{i\frac{4\pi}{3}} \\ \alpha e^{i\frac{2\pi}{3}} & 1 \end{bmatrix}$$
$$\rho_2 = \frac{1}{2}\begin{bmatrix} 1 & \alpha e^{i\frac{2\pi}{3}} \\ \alpha e^{i\frac{4\pi}{3}} & 1 \end{bmatrix},$$

where we have used $\alpha + \beta = 1$ and we have $p_j = \frac{1}{3}$ for all j. If we insert these states into equation (3.21), we once again obtain three constraints on a, b_x, and b_y (note that, using the reasoning from the previous example, we already know that $b_z = 0$ for this ensemble):

$$a^2 - |b|^2 = \frac{2}{3}[a + \alpha b_x - \frac{1}{6}(1-\alpha^2)]$$
$$a^2 - |b|^2 = \frac{2}{3}[a + \frac{\alpha}{2} b_x + \frac{\alpha\sqrt{3}}{2} b_y - \frac{1}{6}(1-\alpha^2)]$$
$$a^2 - |b|^2 = \frac{2}{3}[a + \frac{\alpha}{2} b_x - \frac{\alpha\sqrt{3}}{2} b_y - \frac{1}{6}(1-\alpha^2)].$$

Once again, it is clear that $b_y = 0$. Furthermore, it is simple to show that $b_x = 0$ too - note that this is consistent with equation (3.23) for $p = \frac{1}{3}$. This gives us a quadratic equation in a, which has two roots:

$$a = \frac{1 \pm \alpha}{3}$$

We know that the optimal measurement will result in a probability of correctness which is at least as good as guessing. That is, $\mathrm{P}_{\mathrm{Corr}} \geq \frac{1}{3}$. We therefore end up with the result that

$$\mathrm{P}_{\mathrm{Corr}} = \mathrm{Tr}(\Gamma) = a \tag{3.24}$$
$$= \frac{1+\alpha}{3}. \tag{3.25}$$

We can compare this to known results. Firstly, we know that for pure states (i.e. $\alpha = 1$), we should obtain $\mathrm{P}_{\mathrm{Corr}} = \frac{2}{3}$; this agrees with our result. Furthermore, for $\alpha = 0$ (i.e. the maximally-mixed case), the states are indistinguishable, and therefore we will simply end up guessing, with the probability of guessing correctly being $\frac{1}{3}$. This also agrees with our result.

3.6 Discussion

We have presented a method to construct optimal minimum-error measurements from the known necessary and sufficient conditions. If we know which of a set of states are identified by an optimal measurement, the method presented here allows us to construct four linear conditions on either Γ or Γ^{-1}, from which we have enough information to reconstruct Γ. The remaining problem we have not addressed, and which is common to other methods in the literature [11, 71], is how to find which states our measurement should identify. We finish with some comments on this problem.

In the worst case, we can find the optimal measurement by exhaustive search: we first check if the no-measurement solution is optimal. If yes then we are done, and if not then we know that $k > 1$. We then check whether any measurement identifying just 2 of the states is optimal. This consists of constructing optimal measurements for each pair of states, and checking the condition (3.2) for the remaining $N - 2$ states in each case. There are $\binom{N}{2}$ such measurements. If none of these are optimal, then we know $k > 2$, and so on. This requires constructing $\sum_{k=1}^{4} \binom{N}{k}$ (i.e. $O(N^4)$) candidate Γ operators, and for each one checking $O(N)$ conditions, thus we require $O(N^5)$ operations, in the worst case. Our detailed results for the case of three symmetric states with arbitrary priors, which we discuss in Chapter 4, indicate that for almost all prior probabilities the optimal measurement has only two outcomes. Thus we expect that in many cases an optimal measurement will be found faster than $O(N^5)$ operations.

For a given set of states, the method we present here allows us to characterise the entire parameter space of prior probabilities, beginning with the no-measurement solution, through those regions in which a two-outcome measurement is optimal, and constructing three- and then four-outcome solutions for the remaining regions, as shown in the example above. We note also that for specific cases numerical methods can also be used to determine which states are identified by an optimal measurement, and once this is known our method may be used to find an exact analytical solution for the optimal probability of correctly identifying the state and to find optimal measurements.

We have introduced a new analytical method, complementary to the geometric approach in the literature, for constructing optimal measurements for minimum error state discrimination problems [39]. Our method constructs linear constraints on the so-called Lagrange operator Γ, and its inverse Γ^{-1}, from which the optimal Γ may readily be found for any qubit state discrimination problem. Although the constraints we present appear elsewhere in the literature in a different context, it seems not to have been recognised that these together give enough information to construct optimal measurements. We have further shown that these are applicable to both pure and mixed states in the qubit case.

In this chapter we have discussed the qubit case in detail. We expect that the linear constraints given on Γ may also be applied in higher dimensions. The constraints on Γ^{-1} may be applied to pure states in higher dimensions, although the mixed state case appears less straight-forward, as it is no longer obvious how to decompose a general mixed state into a combination of the identity and rank-one components, as in equation (3.19).

Chapter 4

Optimal measurement strategies for the trine states with arbitrary prior probabilities

4.1 Introduction

In this chapter, we give a complete analysis of the problem of state discrimination for the trine states with arbitrary prior probabilities, for both the minimum error and maximum confidence figures of merit [73]. Each of these are amenable to analytic solutions; in the minimum error case, which we begin with, this is made possible by recent developments [11, 12, 39], including the work of the previous chapter. We continue by investigating the maximum confidence measurement [7] for the trine states with arbitrary prior probabilities and obtain an expression for the probability of correctly identifying each signal state using this measurement scheme.

In this chapter we are concerned with the trine states, qubit states associated with three equidistant points on any great circle of the Bloch sphere [17]. We will place the trine states on the equator of the Bloch sphere, so that:

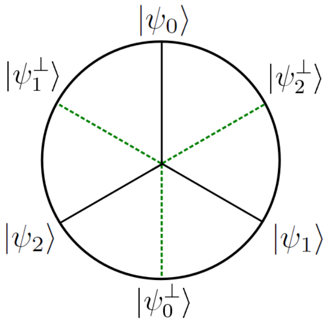

Figure 4.1: The trine states on the equator of the Bloch sphere. The dotted lines show the anti-trine measurement basis - each POVM element is aligned so that it is orthogonal to one of the potential states. For instance, if we get a "click" at the POVM element at $|\psi_0^\perp\rangle$, we know with certainty that state $|\psi_0\rangle$ was not prepared. This is the basis for a system of quantum cryptography described in [2]

$$\begin{aligned}|\psi_0\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle),\\ |\psi_1\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + e^{i\frac{2\pi}{3}}|1\rangle),\\ |\psi_2\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + e^{i\frac{4\pi}{3}}|1\rangle),\end{aligned}$$

where the states $|0\rangle$ and $|1\rangle$ correspond to the poles on the Bloch sphere. These trine states can be visualised on the Bloch sphere as shown in figure 4.1. For equal prior probabilities ($p_0 = p_1 = p_2 = \frac{1}{3}$), it is known that the optimal measurement of the trine states for minimising the probability of error is to measure along the states themselves [34], that is, making a measurement of the form $\pi_j = \frac{2}{3}|\psi_j\rangle\langle\psi_j|$. This is known as the trine measurement.

In contrast to a two-state system, intriguingly, if we wish to maximise the mutual information gained by our measurement, we must use a different POVM: in this case, we perform the so-called anti-trine measurement [36], as shown in Figure 4.1. This involves three measurement outcomes, each of which is perpendicular to one of the trine states; this is therefore an eliminatory measurement, as it tells us with certainty that the system was *not* prepared in a particular state, with the other two possible states equally likely to be the signal state.

Throughout this chapter, and without loss of generality, we assume $p_0 \geq p_1 \geq p_2$.

4.2 Minimum-error measurement

In the case of the trine states, the minimum-error measurement must have either two or three elements: a one-element measurement, that is $\pi_k = \mathbb{1}$ for some k, corresponding to the "no-measurement" strategy, can never be optimal for pure state ensembles [50], as condition (2.6) cannot be satisfied for $j \neq k$. Furthermore, as each measurement outcome corresponds to identifying one of the potential states, the number of outcomes cannot exceed the number of states: any extra elements will be redundant.

In light of this, we split the problem into two parts: we ask when a two-element POVM is optimal, as this is a relatively easy problem to solve, and then we consider the remaining parameter space. In the region where the two-outcome measurement does not give the minimum error, we know that a three-element POVM of some form will be optimal. In this region, we construct the optimal measurement by applying the strategy outlined in Chapter 3. Surprisingly, the two-element POVM is optimal for almost the whole parameter space. We show, explicitly, that all optimal measurements on the trine states are unique - that is, for any choice of initial probabilities $\{p_i\}$, there is one and only one measurement which is optimal.

4.2.1 Conditions for a two-element POVM to be optimal

We know that when $p_2 = 0$, a two-element POVM must be optimal. This problem has a well-known solution, with the optimal probability of correctness given by the Helstrom bound [4, 5]:

$$\mathrm{P}_{\text{2-el}} = \frac{1}{2}(1 + \sqrt{1 - 4p_0p_1|\langle\psi_0|\psi_1\rangle|^2}), \tag{4.1}$$

where "2-el" is short for two-element. It is readily shown that this is achieved by a measurement of the form $\pi_{0,1} = |\Theta_{0,1}\rangle\langle\Theta_{0,1}|$, $\pi_2 = 0$, where

$$|\Theta_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\theta}|1\rangle),$$
$$|\Theta_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - e^{i\theta}|1\rangle),$$

with

$$\tan\theta = \frac{-\sqrt{3}p_1}{2p_0 + p_1}. \tag{4.2}$$

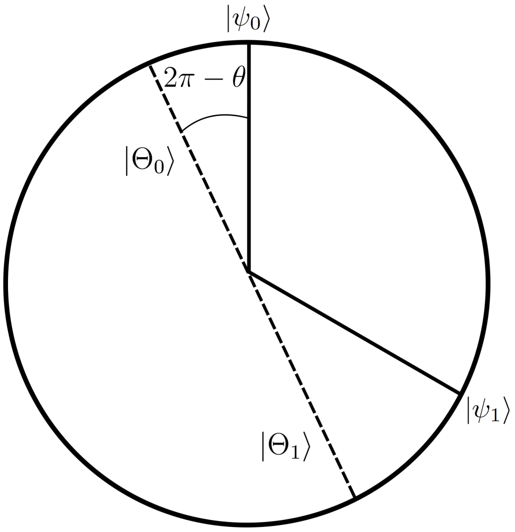

FIGURE 4.2: The two signal states we are trying to discriminate between ($|\psi_0\rangle$ and $|\psi_1\rangle$, solid lines) and the optimal measurement for doing so (dotted lines), where θ is as defined in equation (4.2). Note that the signal states need not be symmetrical with respect to the measurement states - if one state is *a priori* more likely, the optimal measurement will be biased towards it. In this example, $|\psi_0\rangle$, is more likely to be sent than $|\psi_1\rangle$.

Figure 4.2 shows the measurement states on the Bloch sphere. Writing the probability of correctly guessing the state in terms of only p_0 and p_1 gives:

$$\mathrm{P}_{\text{2-el}} = \frac{1}{2}(p_0 + p_1 + \sqrt{p_0^2 + p_0 p_1 + p_1^2}) \tag{4.3}$$

As described in [11, 12, 39], we know that if state ρ_2 is added to this ensemble with a small enough probability, the number of POVM elements necessary for minimum-error measurement remains unchanged. Intuitively, if p_2 is small enough, we do not gain anything by identifying ρ_2, and the minimum-error measurement favours the more likely states. We can use the Helstrom conditions to define precisely what "small enough" means in this context, and put conditions on p_0, p_1 and p_2 which state when a two-element POVM is sufficient and when a three-element POVM is required.

To find the values for p_0 and p_1 for which a two-element POVM is the optimal measurement, we investigate the other Helstrom condition, shown in equation (2.6). This is trivial for $j = 0, 1$, as we already know this must be the optimal measurement when these are the only signal states. Therefore, it suffices to check the positivity of the matrix

$$M = \sum_i p_i \rho_i \pi_i - p_2 \rho_2. \tag{4.4}$$

It follows from the conditions for $j = 0, 1$ that $\sum_i p_i \rho_i \pi_i$ is a positive operator. Further, as ρ_2 is a pure state, M has at most one negative eigenvalue, and to check positivity of M we can calculate the sign of the determinant: when $\det(M)$ is positive, the two-element POVM described above is optimal. This is straightforward, and the determinant of the matrix is found to be

$$\begin{aligned}\det(M) = &-3p_0^4 - 3p_1^4 - 10p_0^3p_1 - 10p_0p_1^3 + 6p_0^3 + 6p_1^3 \\ &- 13p_0^2p_1^2 + 12p_0^2p_1 + 12p_0p_1^2 - 3p_0^2 - 3p_1^2 - 2p_0p_1.\end{aligned} \tag{4.5}$$

To find the boundary of the region where the two-element measurement is optimal, it is useful to parameterise the probabilities as follows: $p_0 = p + \delta, p_1 = p - \delta, p_2 = 1 - 2p$, where the ordering $p_0 \geq p_1 \geq p_2$ implies $\delta \geq 0, \delta \leq 3p - 1, \delta \leq p$.

After a little algebra, we find that the determinant is simply a quadratic in δ^2, with roots $\pm\delta_{c\pm}$, where

$$\delta_{c\pm}^2 = 2 - 6p + 5p^2 \pm 2\sqrt{1 - 6p + 16p^2 - 24p^3 + 16p^4} \tag{4.6}$$

There are four roots for δ, only two of which give physically-realisable probability distributions; these two simply swap p_0 for p_1 and vice-versa (the other two roots correspond to unphysical distributions with, e.g., $p_0 > 1$). Imposing our condition that $p_0 \geq p_1 \geq p_2$, we find that $\det(M) \geq 0$ for $\delta \leq \delta_{c-}$. That is, a two-element POVM is optimal when $\delta < (2 - 6p + 5p^2 - 2\sqrt{1 - 6p + 16p^2 - 24p^3 + 16p^4})^{\frac{1}{2}}$. Otherwise some three-element POVM (discussed in the next section) is optimal. The parameter regions for which the optimal measurement has two or three outcomes are shown in Figure 4.3.

It is apparent that a three-element POVM is only optimal when close to a symmetric ensemble, i.e. p_1 very close to p_2. For $p_1 = p_0 \in [\frac{1}{3}, \frac{4}{9+\sqrt{3}})$ and for all $p_1 = p_2$, the symmetric three-element measurement outlined in [6] is optimal.

An interesting consequence of equation (4.2) is that there is not a one-to-one correspondence between ensembles and optimal measurements. As we can increase p_2 from zero without changing the optimal measurement, there are many different ensembles with the same optimal measurement strategy. In this region, where the two-element POVM is optimal, the optimal measurement depends only on the relative frequency of occurrence of p_0 and p_1 (i.e. the ratio between p_0 and p_1). For fixed measurement angle θ, the

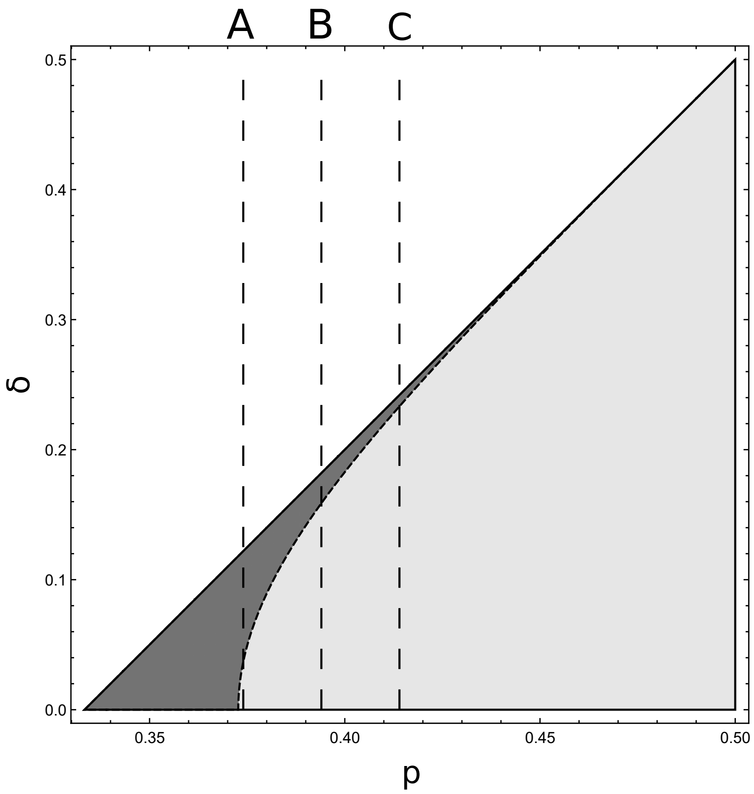

Figure 4.3: Graph showing the sign of the determinant of matrix M in equation (4.4) as a function of p and δ. The dark region corresponds to a negative determinant, and hence shows the region where a 3-element POVM is optimal. The light area displays the rest of the allowable parameter space, where the 2-element POVM we have discussed is optimal. The three dashed vertical lines A, B and C correspond to the three plots A, B and C shown in Figure 4.4. Note that the diagonal line $\delta = 3p - 1$ corresponds to $p_1 = p_2$. It is also important to note that as p increases, the threefold symmetry of the weightings of the states breaks down, resulting in the shift from a three-elements POVM being optimal to a two-element POVM being optimal.

probability of correctness increases linearly with $p_0 + p_1$. We also note that this effect does not happen in the two-state discrimination case, where, given two states and a measurement which is known to be optimal, there is only one p_0 – and hence only one complementary p_1 – which will satisfy the Helstrom conditions.

4.2.2 Optimal three-element POVM

We now turn our attention to the region in which we know a three-element POVM must be optimal. This region is hard to analyse due to its lack of symmetry, but the problem can be solved analytically by using the Helstrom conditions constructively, following the approach developed in Chapter 3. Recall that:

$$\langle\psi_j|\Gamma^{-1}|\psi_j\rangle = \frac{1}{p_j}. \tag{4.7}$$

By writing Γ^{-1} in the form $\frac{1}{2}(a\mathbb{1}+\vec{b}\cdot\hat{\sigma})$, we find three linear equations in three unknowns. As described in Chapter 3 and [71], we may assume from symmetry that the optimal POVM will be in the same plane as the states, so $b_z = 0$, and hence find a, b_x, b_y. Thus we can find Γ and hence $\mathrm{P}_{\mathrm{Corr}}$, the optimal probability of correctly identifying the state which was sent, as $\mathrm{P}_{\mathrm{Corr}} = \sum_k p_k \operatorname{Tr}(\rho_k \pi_k) = \operatorname{Tr}(\Gamma) = \frac{4a}{a^2-|b|^2}$. In fact, because we know that $\Gamma - p_j\rho_j = c_j|\phi_j^\perp\rangle\langle\phi_j^\perp|$, we can also explicitly find the POVM elements and hence extract the optimal measurement directly from the Helstrom conditions. Furthermore, as Γ is known to be unique for a given set of states [9], this POVM will be unique for this ensemble of $\{p_j\}$ and $\{\rho_j\}$, as the vector solution $|\phi_j^\perp\rangle$ is unique.

It is sufficient for our purposes to simply calculate $\mathrm{P}_{\text{3-el}}$. From the above, we obtain

$$\begin{aligned} a &= \frac{2}{3}\left(\frac{1}{p_0}+\frac{1}{p_1}+\frac{1}{p_2}\right) \\ b_x &= \frac{2}{3}\left(\frac{2}{p_0}-\frac{1}{p_1}-\frac{1}{p_2}\right) \\ b_y &= \frac{2}{\sqrt{3}}\left(\frac{1}{p_1}-\frac{1}{p_2}\right), \end{aligned}$$

which yields:

$$\mathrm{P}_{\text{3-el}} = \frac{2(p_0p_1+p_0p_2+p_1p_2)}{2-\left(\frac{p_0p_1}{p_2}+\frac{p_0p_2}{p_1}+\frac{p_1p_2}{p_0}\right)}. \tag{4.8}$$

If we compare this to the expression for $\mathrm{P}_{\text{2-el}}$ given by the optimal two-element POVM then we find that they meet at the boundary when the two-element POVM stops being optimal, as we would expect.

Our expression for $\mathrm{P}_{\text{3-el}}$ has the interesting property that, in parts of the region where we know the two-element POVM to be optimal, the expression for $\mathrm{P}_{\text{3-el}}$ yields a greater value than $\mathrm{P}_{\text{2-el}}$. We also obtain some values for $\mathrm{P}_{\text{3-el}}$ which are greater than 1, which is clearly incorrect. These anomalies are due to the fact that our method of obtaining Γ does not *strictly* impose the conditions for POVM elements; specifically, not every POVM element π_i is a positive semi-definite operator. This may be seen by comparing our measurement to the analogous measurement detailed in [6]. This is not a problem, of course, as these regions in the parameter space are readily determined. At $\delta = 0$, we have $p_0 = p_1$ and the optimal measurement includes a POVM element of the form $(1-a^2)|\psi_2\rangle\langle\psi_2|$, with $a = \frac{\sqrt{3}p}{4-9p}$. Clearly the factor of $(1-a^2)$ becomes negative for $p > \frac{4}{9+\sqrt{3}}$, and so our attempted measurement no longer fulfils the POVM criteria, giving spurious results. It is at this point that the two-element POVM becomes optimal. Thus we can conclude that our optimal three-element POVM does indeed become invalid in the region where we know a two-element POVM must be optimal.

To summarise, this gives us the following functions for the probability of correctly guessing the signal state using the minimum-error measurement scheme. In the case where $\delta < (2 - 6p + 5p^2\ -2\sqrt{1-6p+16p^2-24p^3+16p^4})^{\frac{1}{2}}$, we have:

$$\begin{aligned} \mathrm{P}_{\text{2-el}} &= \frac{1}{2}(p_0 + p_1 + \sqrt{p_0^2 + p_0 p_1 + p_1^2}) \\ &= p + \frac{1}{2}\sqrt{3p^2 + \delta^2}. \end{aligned} \tag{4.9}$$

Otherwise:

$$\begin{aligned} \mathrm{P}_{\text{3-el}} &= \frac{2(p_0p_1 + p_0p_2 + p_1p_2)}{2 - (\frac{p_0p_1}{p_2} + \frac{p_0p_2}{p_1} + \frac{p_1p_2}{p_0})} \\ &= \frac{2(1-2p)(p^2-\delta^2)(3p^2+\delta^2-2p)}{9p^4 - 4p^3 + 6p^2\delta^2 - 12p\delta^2 + 4\delta^2 + \delta^4}. \end{aligned} \tag{4.10}$$

We can therefore plot the optimal probability of correctness for discriminating between the trine states for arbitrary prior probabilities. These results are shown in Figure 4.4 and Figure 4.5, for various values of p and δ.

This solves the problem of minimum-error state discrimination between the trine states for *all possible* probability distributions, and highlights some differences between two-state and three-state discrimination. Firstly, for the two-state case we always require two POVM elements and, indeed, these are both simple projectors. In this case each signal state has a measurement outcome associated with it. This is not the case for the three-state problem, for which it is sometimes beneficial to simply never measure one of the signal states. Indeed, for most of the parameter space, a two-outcome measurement is optimal. Our solution also shows that, for three states, there is not a one-to-one correspondence between ensembles and optimal measurements; a certain measurement strategy may be optimal for multiple probability distributions of the trine states, whereas in the two-state case each optimal measurement strategy is unique to its corresponding probability distribution.

For $\delta = 0$ our results agree with previous work [6, 39]. As we have produced an analytic solution, it is also possible to use this to solve problems where state discrimination arises as a smaller part of a problem, as occurs when multiple copies are available. We shall see this in Chapter 5

4.3 Maximum confidence measurement

Maximum confidence measurements, as described in §2.5, may be viewed as a generalisation of unambiguous discrimination [7, 32]: whereas the latter is only possible when the states to be measured are linearly independent [16], maximum confidence is a viable strategy for linearly dependent states. While the maximum confidence measurement does not have the advantage of giving an answer which is *guaranteed* to be correct (as unambiguous discrimination does), it offers a "middle ground" where, if a given state is identified, it is with the lowest possible probability of error for that state; otherwise the output is an inconclusive outcome, similarly to unambiguous discrimination. It has the advantage of an analytic solution for the elements of the optimal POVM in general, and is also related in certain cases to the minimum-error strategy. Understanding the maximum confidence measurement for the trines with arbitrary priors provides insight into the form of our minimum-error results.

The maximum confidence measurement scheme has already been described for three equiprobable symmetric states on the Bloch sphere [7, 65], and we extend this to the case with arbitrary prior probabilities.

In this measurement scheme, we have $\pi_i \propto \rho^{-1}\rho_i\rho^{-1}$, where $\rho = \sum_j p_j|\psi_j\rangle\langle\psi_j|$. Note that the figure of merit for this strategy is the probability of outcome π_i correctly identifying the state ρ_i, given by Bayes:

$$\mathrm{P}(\rho_i|\pi_i) = \frac{p_i\mathrm{P}(\pi_i|\rho_i)}{\mathrm{P}(\pi_i)} = \frac{\mathrm{P}(\pi_i,\rho_i)}{\mathrm{P}(\pi_i,\rho_i) + \sum_{j\neq i}\mathrm{P}(\pi_i,\rho_j)}. \tag{4.11}$$

This is independent of the constant of proportionality multiplying π_i, which may therefore be chosen arbitrarily. It is always possible to choose the constants of proportionality such that $\sum_j \pi_j \leq \mathbb{1}$. If necessary, a complete measurement may then be formed by adding an inconclusive outcome $\pi_? = \mathbb{1} - \sum_j \pi_j$. The probability that each measurement outcome accurately reflects the state of the system is, however, independent of how we

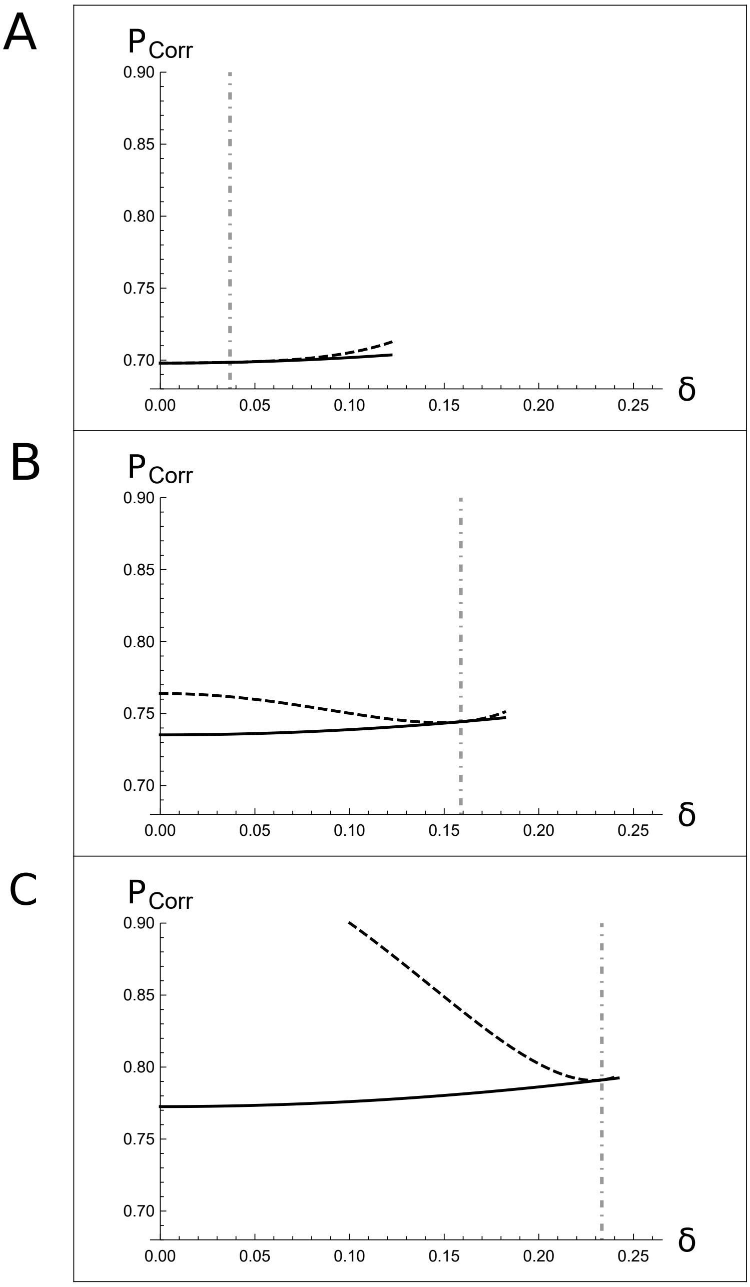

FIGURE 4.4: Comparisons of $\mathrm{P}_{\mathrm{Corr}}$ given by the optimal two-element POVM (bold line) and the results given by our method for finding the optimal three-element POVM (dotted line) for fixed values of p. In graphs A, B and C, respectively, p has values 0.374, 0.394, and 0.414, corresponding to the lines A, B and C shown in Figure 4.3. The dot-dashed grey vertical lines show when the determinant in equation (4.5) becomes negative and thus a three-element POVM becomes physically realisable. That is, the three-element POVM is only viable to the right of the dot-dashed grey line. Note that, when physically viable, the three-element POVM does not significantly outperform the two-element POVM.

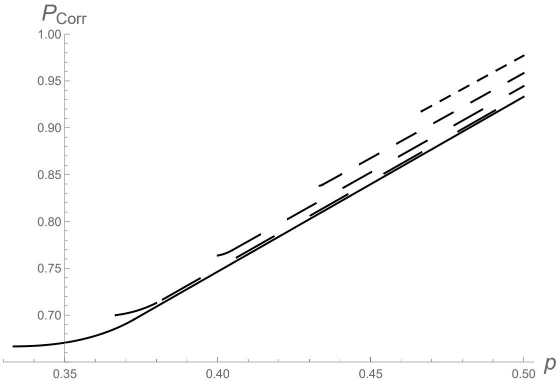

FIGURE 4.5: Graph showing the probability of correctly identifying the signal state using the optimal measurement strategy for $p \in [\frac{1}{3}, \frac{1}{2}]$. The lines, in increasing amounts of dashing - and lowest to highest - correspond to $\delta = 0, \delta = 0.1, \delta = 0.2, \delta = 0.3$ and $\delta = 0.4$.

choose to complete the measurement.

It is convenient to note that, for qubits, $\rho^{-1} \propto \sum_j p_j |\psi_j^{\perp}\rangle\langle\psi_j^{\perp}|$. This may be seen by considering ρ as a point within the Bloch sphere; ρ^{-1} must therefore correspond to the antipodal point in the Bloch sphere - this antipodal point is given by $\sum_j p_j |\psi_j^{\perp}\rangle\langle\psi_j^{\perp}|$. It is perhaps useful to think of the decompositions of $\rho = \frac{1}{2}(\mathbb{1} + \vec{b} \cdot \hat{\sigma})$ and $\rho^{-1} \propto (\mathbb{1} - \vec{b} \cdot \hat{\sigma})$. In fact, we can go further by noting that $\rho^{-1} = [1 - \text{Tr}(\rho^2)]^{-1}(\mathbb{1} - \vec{b} \cdot \hat{\sigma})$: this means ρ must be a mixed state, otherwise ρ^{-1} has no physical meaning. Using $\rho^{-1} \propto \sum_j p_j |\psi_j^{\perp}\rangle\langle\psi_j^{\perp}|$, therefore, we may write

$$\pi_i \propto \sum_{j,k} p_j p_k |\psi_j^{\perp}\rangle\langle\psi_j^{\perp}|\psi_i\rangle\langle\psi_i|\psi_k^{\perp}\rangle\langle\psi_k^{\perp}|. \tag{4.12}$$

The numerator of equation (4.11) in the general case is $p_i\langle\psi_i|\pi_i|\psi_i\rangle \propto p_i(\sum_m p_m |\langle\psi_m^{\perp}|\psi_i\rangle|^2)^2$. Due to the symmetry of the trine ensemble, it is readily verified that $|\langle\psi_j^{\perp}|\psi_i\rangle|^2 = \frac{3}{4}(1 - \delta_{ij})$. The numerator, in this instance, is therefore $\frac{9}{16}p_i(1 - p_i)^2$. The other piece of this expression takes the following form, where the last two lines are dependent on the number of states we are discriminating between and their overlaps:

$$\begin{aligned}\sum_{j\neq i} \mathrm{P}(\pi_i,\rho_j) &= \sum_{j\neq i} p_j \langle\psi_j|\pi_i|\psi_j\rangle \\ &\propto \sum_{j\neq i} p_j \left|\sum_m p_m \langle\psi_j|\psi_m^\perp\rangle\langle\psi_m^\perp|\psi_i\rangle\right|^2 \\ &= \frac{9}{16}\sum_{j\neq i} p_j \sum_{m\neq i,j} p_m^2 \\ &= \frac{9}{16}(1-p_i)\prod_{j\neq i} p_j.\end{aligned}$$

The final line may not be obvious at first, but can be verified by setting, e.g., $i=0$ and noting that m can only take one value - if $j=1$, $m=2$ and vice versa. We therefore obtain

$$\mathrm{P}(i)_{\mathrm{Corr}} = \left(1+\frac{\prod_{j\neq i} p_j}{p_i\sum_{j\neq i} p_j}\right)^{-1}, \tag{4.13}$$

which has some attributes we might expect: when any individual p_j is set equal to zero, the probability of correctly identifying the state ρ_i $(i\neq j)$ becomes unity, as the set of possible states is now linearly independent, allowing unambiguous discrimination to be performed. When p_i is zero, there is zero chance of that state being correctly identified, as one might anticipate.

We plot the confidence of correctly identifying each state using this measurement scheme, and compare this to the confidence using the minimum-error strategy. These can be seen in Figures 4.6 and 4.7 (note that Figure 4.7 only uses the two-outcome measurement, for simplicity). In both cases, the minimum-error measurement is close to optimal for ρ_0 and ρ_1. Also note how low the confidence for ρ_2 gets as p increases - this indicates why this state is not identified in the minimum-error measurement.

4.4 Conclusion

We have investigated the optimal measurement strategies for the minimum-error and maximum confidence figures of merit for three equidistant states on the equator of the Bloch sphere with arbitrary prior probabilities, providing values for the optimal probability of correctly identifying the state in each case. The most surprising result is that,

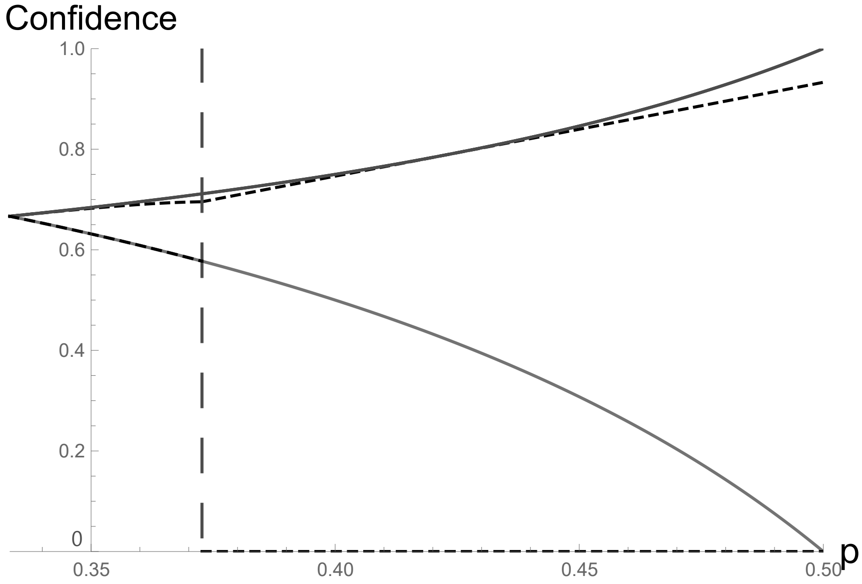

FIGURE 4.6: Graph showing the confidence in correctly identifying the signal state given the outcome of the minimum-error strategy (black dotted line) and the maximum confidence measurement for $p \in [\frac{1}{3}, \frac{1}{2}]$ and $\delta = 0$. The lighter lines (from darkest to lightest and top to bottom) represent the maximum confidence strategy on states ρ_0, ρ_1 and ρ_2 - note that as $\delta = 0$, the states ρ_0 and ρ_1 are equally likely, and so their values for confidence completely overlap. Also, the minimum-error measurement and maximum confidence measurement are identical for ρ_2, so give the same confidence value, resulting in only 3 lines being visible. The dotted vertical line corresponds to the crossover point at which the minimum-error measurement stops being a three-outcome measurement and starts being a two-outcome measurement.

for much of the parameter space of probabilities, the optimal minimum-error measurement is a simple von Neumann measurement, and this allows optimal discrimination between these states with a minimum of resources. However, this is in keeping with previous results: for a completely unknown qubit state, the best measurement to estimate the state is simply a von Neumann measurement in any basis [74]; furthermore, the optimal intercept-resend strategy for an eavesdropper in the BB84 quantum key distribution protocol - which has four signal states - is a von Neumann measurement in the so-called Breidbart basis [75]. This was also noted by Andersson *et. al.*, in a case with restricted symmetry [6]. We have shown that the region of parameter space for which a POVM measurement is needed is rather small. This indicates that cases requiring POVM measurements are perhaps rather special, which might have implications for quantum key distribution, scalability in quantum computing, and quantum sensing.

This chapter solves the problem of optimal state discrimination between the trine states for arbitrary prior probabilities analytically; we have also shown that, for given probabilities p_0, p_1, p_2, there is one and only one optimal measurement - when a two-outcome

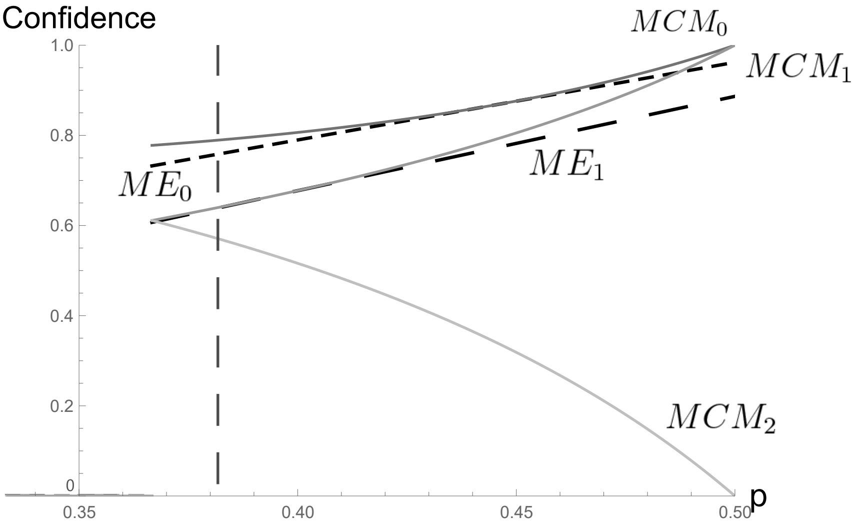

FIGURE 4.7: Graph showing the confidence in correctly identifying the signal state given the outcome of the two-element minimum-error strategy (black dotted lines, ME_0 and ME_1) and the maximum confidence measurement for $p \in [\frac{1}{3}, \frac{1}{2}]$ and $\delta = 0.1$. The lighter lines (from darkest to lightest and top to bottom) represent the maximum confidence measurement (MCM) strategy on states ρ_0, ρ_1 and ρ_2 (labelled MCM_0, MCM_1, and MCM_2). The higher dotted line corresponds to the minimum-error strategy on ρ_0 (i.e., ME_0), while the lower one corresponds to the same minimum-error measurement on ρ_1 (ME_1). The dotted vertical line corresponds to the crossover point at which the minimum-error measurement stops being a three-outcome measurement and starts being a two-outcome measurement. Note that, as predicted, the most likely states are the easiest to detect in this measurement scheme. We ignore the three-outcome minimum-error measurement, as it is only optimal for a small region of the space (c.f. Figure 4.3.)

measurement is optimal we know it is unique, as the measurement angle is fixed by equation (4.2), and, as already discussed, equation (3.14) shows that the three-element POVM must also be a unique solution. This also shows that there is no region where two- and three-outcome measurements are simultaneously optimal. This work provides a complement to that of Hunter [58, 76], which found the minimum-error strategy for arbitrary equiprobable signal states. Subsequent work presented analytical and geometric methods for arbitrary priors [10–12, 39]; what is surprising about the results presented here is the simplicity of the expressions for the optimal probability of success given in equations (4.9) and (4.10).

This chapter also gives the maximum confidence that it is possible for a measurement to achieve on each of the trine states with arbitrary prior probabilities. This helps to identify situations in which it is sub-optimal for the minimum-error strategy to identify every signal state, as the maximum confidence possible for the least likely state tends to zero.

We hope that this work leads to new and interesting results, and we look forward to seeing other ways in which our method for tackling minimum-error discrimination problems is used.

Chapter 5

Optimal measurement strategies for multiple copies of the trine states

5.1 Introduction

In this chapter, we investigate the difference between local and global measurement schemes when more than one copy of the state is provided [77]. Understanding when joint control is really necessary for optimal – or close-to-optimal – performance is of considerable practical interest. We investigate the dependence of the optimal minimum-error strategy for two copies of the trine states on the measurement efficiency for non-ideal measurements. In doing so, we find that the gap in efficacy between the optimal global and local measurement schemes is small – given that global measurements are hard to implement, this suggests that it is hard to motivate an experiment using such a measurement. For particularly low efficiencies, the optimal local measurement in fact outperforms the optimal joint measurement.

We conclude by examining local and global measurements in the context of the maximum confidence measurement strategy (which in this case is also unambiguous discrimination) of multiple copies of the trine states, in a natural extension of the work of Chefles [30]. Here we also account for measurements with non-unit efficiency.

5.2 Minimum-error discrimination of the double-trine ensemble

Until this point, all of our work has assumed that we have an ideal photon detector (or other measurement device). In reality, detectors and other measurements are not 100% efficient. This particularly affects the problem of multi-partite state discrimination, where the optimal measurement may depend on the outcomes of previous rounds of measurement. An analytic solution for minimum-error state discrimination, such as that described in Chapter 3, allows us to investigate how the optimal measurement changes based on which outcome occurs. Note that we are discussing *measurement* efficiency and not *detector* efficiency.

For the sake of simplicity, we consider the case where Alice sends Bob two copies of the signal state ρ_j (one of the trine states) with *a priori* probability $\frac{1}{3}$. Bob then makes two rounds of measurements to determine which state was sent, with the second measurement allowed to depend on the result of the first. If both measurements fail, Bob has to guess with accuracy $\frac{1}{3}$. We show that the optimal measurement strategy varies depending on the measurement efficiency η. This process is shown schematically in Figure 5.1. The probability of Bob correctly guessing the state based on this measurement scheme is therefore

$$\mathrm{P_{Corr}} = \eta^2 \mathrm{P}^B_{\mathrm{Corr}} + \eta(1-\eta)(\mathrm{P}^A_{\mathrm{Corr}} + \mathrm{P}^C_{\mathrm{Corr}}) + \frac{1}{3}(1-\eta)^2. \tag{5.1}$$

It follows that there is a tradeoff between the η^2 and $\eta(1-\eta)$ terms - for high η, we wish to maximise $\mathrm{P}^B_{\mathrm{Corr}}$, whereas for low η the $(\mathrm{P}^C_{\mathrm{Corr}} + \mathrm{P}^A_{\mathrm{Corr}})$ term dominates.

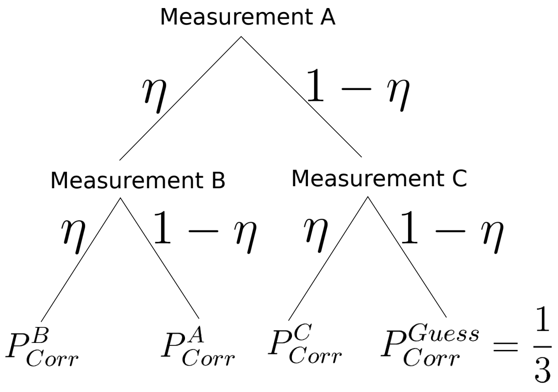

FIGURE 5.1: Decision tree showing the general form of Bob's measurements. Moving to the left indicates that the measurement was successful, while moving to the right indicates failure. These events occur with probability η and $1-\eta$, respectively. Note that measurements B and C are, in general, different.

5.2.1 Efficient measurements, i.e., $\eta \to 1$

We begin by reviewing known cases, and then discuss the general case. For perfect efficiency, the optimal sequential measurement is known [33, 39]. On the first copy of the state, the best choice of initial measurement is the antitrine measurement, shown in Figure 4.1. Measurement operators for the antitrine measurement A have form $\pi_j^{(A)} = \frac{2}{3}|\phi_j\rangle\langle\phi_j|$, where

$$|\phi_j\rangle = \frac{1}{\sqrt{2}}(|0\rangle - e^{i\frac{2\pi}{3}j}|1\rangle) \tag{5.2}$$

and $j = 0, 1, 2$. The antitrine measurement, if successful, rules out one possible state, leaving the other two states equally probable. We will denote this with $\bar{j}$. For example, if outcome $\bar{0}$ is given, our updated probabilities become $p_0 = 0, p_1 = \frac{1}{2}, p_2 = \frac{1}{2}$.

On the second copy, as the two remaining states are equally probable, the probability of the Helstrom measurement correctly identifying the state is

$$\begin{aligned} \mathrm{P}^{\mathrm{H}}_{\mathrm{Corr}} &= \frac{1}{2}(1 + \sqrt{1 - |\langle\psi_0|\psi_1\rangle|^2}) \\ &= \frac{1}{2}\left(1 + \frac{\sqrt{3}}{2}\right). \end{aligned} \tag{5.3}$$

This is therefore the overall probability of success for measurements of perfect efficiency.

For η close to 1, therefore, it is a reasonable strategy to attempt this optimal measurement. In the case in which the first measurement fails, with probability $1-\eta$, we simply perform the trine measurement on the second copy, which identifies the state correctly with probability $\frac{2}{3}$, i.e.

$$\mathrm{P}^{\mathrm{T}}_{\mathrm{Corr}} = \frac{2}{3}. \tag{5.4}$$

This also succeeds with probability η (close to 1). If, on the other hand, the first measurement succeeds but the Helstrom measurement fails, then we are simply guessing between two equally probable states, i.e.

$$\mathrm{P}^{\mathrm{A}}_{\mathrm{Corr}} = \frac{1}{2}. \tag{5.5}$$

This measurement scheme is portrayed in the decision tree depicted in Figure 5.2. Combining this information with equation (5.1) gives us a probability of correctly guessing the state of

$$\mathrm{P}_{\mathrm{Corr}}^{\eta\to 1} = \frac{1}{12}\left\{4 + \eta[6 + \eta(-4 + 3\sqrt{3})]\right\}. \tag{5.6}$$

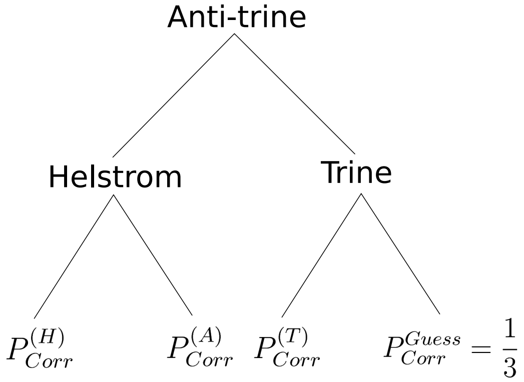

FIGURE 5.2: Decision tree showing the optimal measurement strategy when $\eta \to 1$. Moving to the left indicates that the measurement was successful, while moving to the right indicates failure.

5.2.2 Inefficient measurements, i.e., $\eta \to 0$

In the case in which the measurement efficiency is low, measurements fail most of the time; it is therefore clearly a better strategy to optimise the single-copy measurement. Thus, we perform the trine measurement, which is correct with probability $\frac{2}{3}$, on the first copy. That is,

$$\mathrm{P}_{\mathrm{Corr}}^{\mathrm{T}} = \frac{2}{3}. \tag{5.7}$$

In the (rare) case where the initial measurement is a success, we again perform the minimum-error measurement, but with our priors updated based on the outcome of the previous measurement. The priors are updated according to the Bayesian update rule

$$p_{i|j} = \frac{\frac{2}{3}|\langle\psi_i|\psi_j\rangle|^2}{\frac{2}{3}\sum_k |\langle\psi_j|\psi_k\rangle|^2}, \tag{5.8}$$

which yields an updated prior of $\frac{2}{3}$ if $i = j$, and $\frac{1}{6}$ otherwise. This new weighting describes a mirror symmetric set of states [6], discussed earlier. Using [6], or our results from Chapter 4, we find:

$$\mathrm{P}_{\mathrm{Corr}}^{\mathrm{Symm}} = \sum_k p_k \langle\psi_k|\pi_k^{(\mathrm{Symm})}|\psi_k\rangle \tag{5.9}$$

$$= \frac{4}{5}. \tag{5.10}$$

In the (much more likely) event of the first measurement failing, the optimal second measurement is simply to repeat the trine measurement (i.e. the minimum error measurement for the equally weighted trine states). As before, this gives

$$\mathrm{P}^{\mathrm{T}}_{\mathrm{Corr}} = \frac{2}{3}. \tag{5.11}$$

Combining these probabilities in the same manner as in the high efficiency case, the total success probability is

$$\begin{aligned}\mathrm{P}^{\eta\to 0}_{\mathrm{Corr}} &= \eta^2 \mathrm{P}^{(\mathrm{Symm})}_{\mathrm{Corr}} + 2\eta(1-\eta)\mathrm{P}^{(\mathrm{T})}_{\mathrm{Corr}} + \frac{1}{3}(1-\eta)^2 \\ &= \frac{4}{5}\eta^2 + \frac{4}{3}(1-\eta)\eta + \frac{1}{3}(1-\eta)^2. \end{aligned} \tag{5.12}$$

5.2.3 General case

It is clear from the above discussion that the best measurement strategy to employ depends on the efficiency η. We now consider the general case, and show how our results from the previous sections may be used to understand how the optimal measurement changes with η. To this end, we define a general POVM as

$$\pi_k(\phi) = \frac{2}{3}|\Phi_k\rangle\langle\Phi_k|, \tag{5.13}$$

where

$$|\Phi_k\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + e^{i(\frac{2\pi}{3}k+\phi)}|1\rangle\right). \tag{5.14}$$

This is a general form of the trine measurement: for $\phi = 0$ we have the trine measurement, and for $\phi = \frac{\pi}{3}$ we have the antitrine measurement. To avoid degeneracies, we limit $\phi \in [0, \frac{2\pi}{3}]$. We know that such a POVM must be optimal for the measurement on the first copy, as it shares the same symmetry as the states, as discussed in §2.6.

If the first measurement using this general POVM is successful, which occurs with probability η, our updated priors are

$$\begin{aligned} p_0 &= \frac{1}{3}\big(1+\cos\phi\big) \\ p_1 &= \frac{1}{3}\big(1+\cos(\phi-\frac{2\pi}{3})\big) \\ p_2 &= \frac{1}{3}\big(1+\cos(\phi+\frac{2\pi}{3})\big). \end{aligned} \tag{5.15}$$

Assuming, without loss of generality, $p_0 \geq p_1 \geq p_2$ means we only need consider $\phi \in [0, \frac{\pi}{3}]$. An example of the states with updated priors is shown in Figure 5.3. If the first measurement is unsuccessful, we continue with the trine measurement in the same way we did with for the low efficiency case – in both cases, our prior probabilities are $p = \frac{1}{3}$, and this is the minimum-error measurement. When the first measurement is successful, however, we must determine the optimal new measurement given the dependence of the priors on arbitrary angle ϕ.

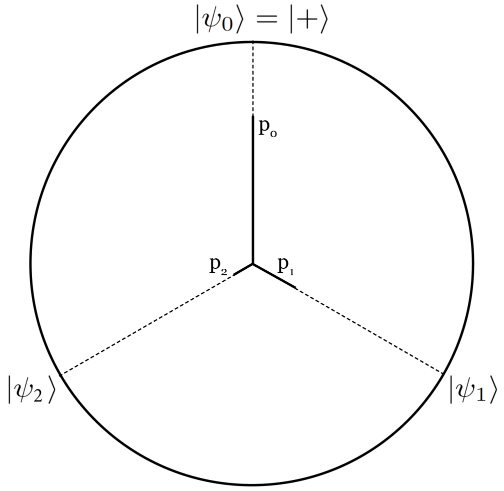

FIGURE 5.3: Depiction of the trine states with updated priors, as described in equation (5.15), where the length of the line is indicative of the probability of that state being sent. In this example, $\phi = 0.2$, giving $p_0 = 0.66$, $p_1 = 0.23$, and $p_2 = 0.11$.

Using the results from Chapter 4, we can determine the values of ϕ for which a 2-element POVM is optimal. We ultimately find that the conditions for choosing the optimal POVM are:

$$\phi < 0.0121351 \implies 3 - \text{element POVM optimal}$$

$$\phi > 0.0121351 \implies 2 - \text{element POVM optimal.}$$

The probability of correctly identifying the signal state using the 2- and 3-element POVMs is given by equations (4.9) and (4.10), given here in terms of ϕ:

$$\begin{aligned}
\mathrm{P}_{\text{2-el}}(\phi) &= \frac{1}{12}\left(4+\cos\phi+\sqrt{3}\sin\phi+\sqrt{15+6\cos\phi+6\sqrt{3}\sin\phi}\right) \\
&= \frac{1}{12}\left[4+2\sin(\phi+\frac{\pi}{6})+\sqrt{15+12\sin(\phi+\frac{\pi}{6})}\right] \\
\mathrm{P}_{\text{3-el}}(\phi) &= \frac{1}{8-\frac{27}{4}\sec^2\frac{3\phi}{2}}.
\end{aligned}$$

These probabilities correspond to the success probability for two sequential measurements on the trine states, where the first measurement (i.e., Measurement A in Figure 5.1) is a rotated trine measurement given by π_k^ϕ and the second measurement (Measurement B in Figure 5.1) is the minimum-error measurement given the subsequent updated priors. These expressions are plotted in Figure 5.4. All that remains is simply to combine these probabilities with the probability that the second ϕ-dependent measurement fails, in which case we guess the most likely state based on the updated priors from the first round of measurement.

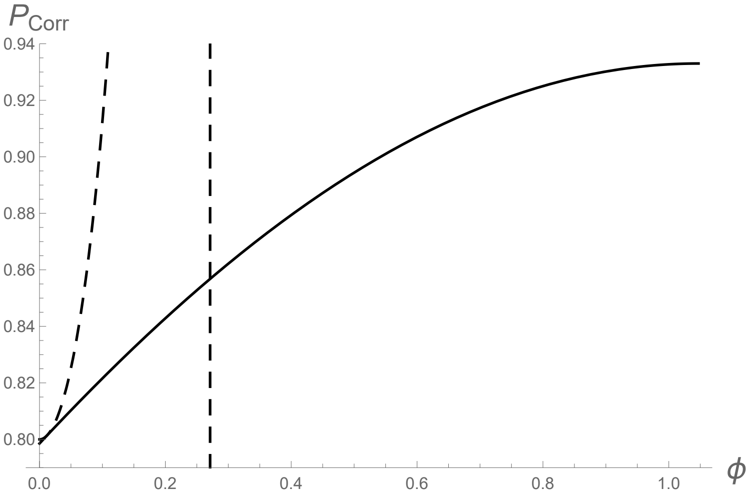

FIGURE 5.4: The expressions $\mathrm{P}_{\text{2-el}}(\phi)$ (solid line) and $\mathrm{P}_{\text{3-el}}(\phi)$ (dashed line) plotted against the angle ϕ from equation (5.14). $\mathrm{P}_{\text{3-el}}(\phi)$ is only viable, i.e. corresponds to a physically-realisable measurement scheme, for $\phi < 0.0121351$. See the discussion in §4.2.2 for more details. It is important to note that the dotted line on the right is a continuation of the dotted line on the right: the unphysical nature of this curve – tending to infinity – is a result of "forcing" a three-element solution where a two-element measurement is optimal.

The full expression for the probability of correctness of this strategy is therefore:

$$\mathrm{P_{Corr}} = \eta^2 \mathrm{P_{n\text{-}el}}(\phi) + \frac{1}{3}\eta(1-\eta)(3+\cos\phi) + \frac{1}{3}(1-\eta)^2 \tag{5.16}$$

where n may be 2 or 3, depending on the value of ϕ.

From Figure 5.4, we note that even in the very small region in which the three-outcome measurement is optimal, the performance of the best two-outcome measurement is almost indistinguishable from the optimal strategy. For this reason we are justified in the remaining in only considering the two-outcome measurement on the second copy. Therefore we arrive at:

$$\begin{aligned}\mathrm{P_{Corr}} =& \frac{\eta^2}{12}\left[4 + 2\sin(\phi+\frac{\pi}{6}) + \sqrt{15 + 12\sin(\phi+\frac{\pi}{6})}\right] \\ &+ \frac{\eta}{3}(1-\eta)(3+\cos\phi) + \frac{1}{3}(1-\eta)^2.\end{aligned} \tag{5.17}$$

Using this expression, we can differentiate to find the optimal angle ϕ to use for any given efficiency η. This is not analytically solvable for ϕ as a function of η. However, it is readily solved for η as a function of the corresponding optimal ϕ, and this turns out to be sufficient for our needs. We find:

$$\frac{\eta}{2(1-\eta)} = \frac{\sin\phi}{\cos(\phi+\frac{\pi}{6})}\left[1 + \frac{3}{\sqrt{15+12\sin(\phi+\frac{\pi}{6})}}\right]^{-1}. \tag{5.18}$$

This gives us a plot of optimal angle ϕ for a given efficiency η, shown in Figure 5.5.

For any given η we can thus find the corresponding ϕ which defines the optimal strategy. Using this relationship, we can further plot (as a parametric plot) the overall optimal probability of success as a function of η, shown in Figure 5.6. Also shown in Figure 5.6 is the performance of various strategies for fixed ϕ as a function of η. We see that the performance is not very sensitive to ϕ, which is encouraging for experimental applications, and that the optimal measurement is well-approximated by choosing ϕ to be 0, $\frac{\pi}{6}$ and $\frac{\pi}{3}$ at low, medium and high measurement efficiencies, respectively.

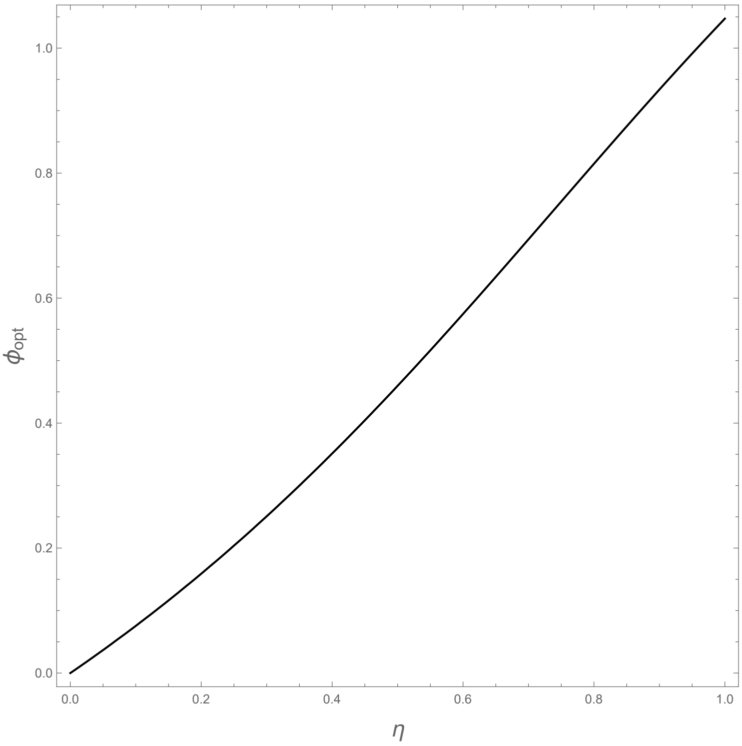

FIGURE 5.5: Optimal angle ϕ for a given efficiency η. Note that there is a one-to-one correspondence between the two; every value for efficiency η has a corresponding optimal ϕ, which is unique.

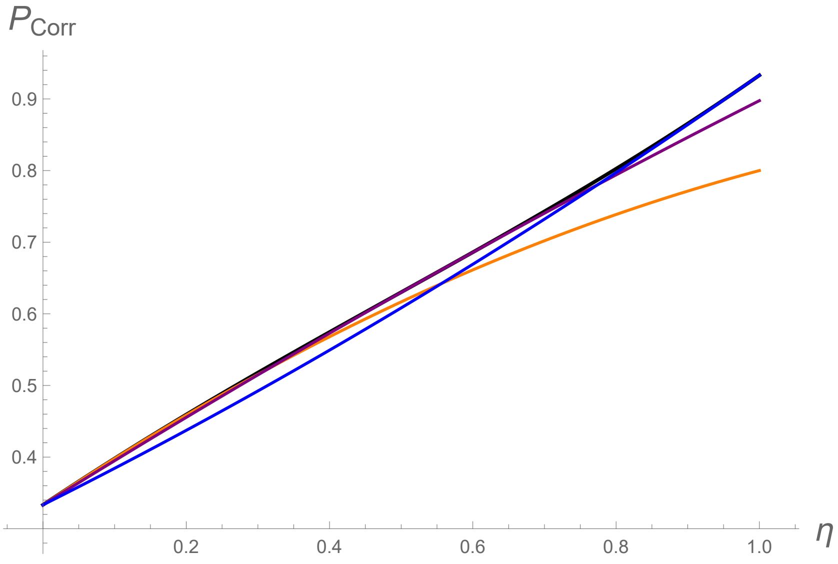

FIGURE 5.6: The probability of correctly guessing the signal state using the optimal sequential measurement (solid black line) plotted against that for various fixed-ϕ measurement schemes. These measurements have had ϕ fixed at 0 (orange line), $\frac{\pi}{6}$ (purple line), and $\frac{\pi}{3}$ (blue line). These are the optimal measurements for low, medium, and high efficiencies, respectively. Note that by using only these three measurement schemes, one can effectively coarse-grain the optimal measurement for any efficiency η. In fact, fixing $\phi = \frac{\pi}{6}$ is a very good approximation over the whole parameter space. Figure 5.7 below shows the same information in more detail.

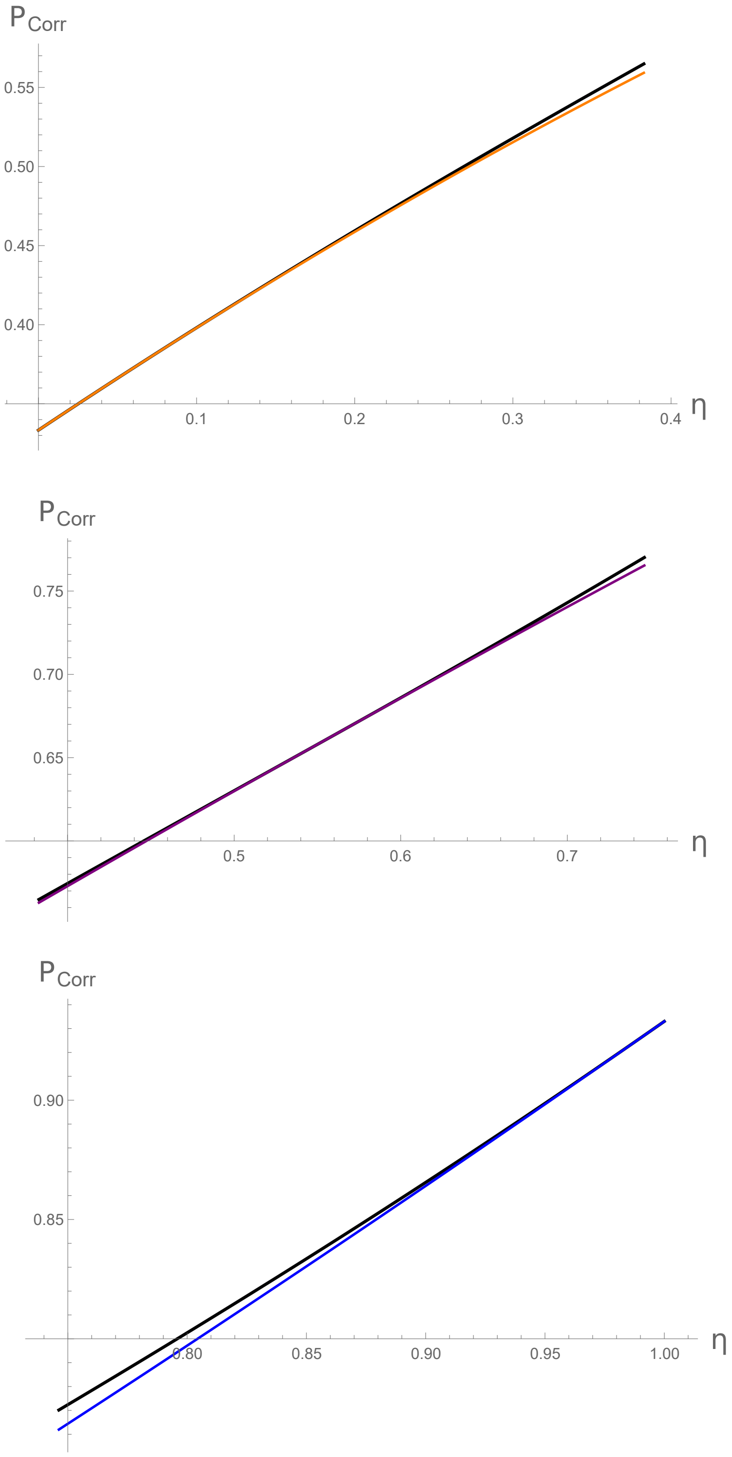

FIGURE 5.7: Detailed close-up of Figure 5.6. The probability of correctly guessing the signal state using the optimal sequential measurement (solid black line) has been plotted against that for various fixed-ϕ measurement schemes. These measurements have had ϕ fixed at 0 (orange line), $\frac{\pi}{6}$ (purple line), and $\frac{\pi}{3}$ (blue line). These are the optimal measurements for low, medium, and high efficiencies, respectively – this can be seen by how close in efficacy they are to the optimal sequential measurement.

5.2.4 Joint Measurement

With this information, it is also possible to compare the optimal local measurement strategy to the optimal global strategy. We consider the simplest possible way of introducing inefficiencies to the joint measurement: the measurement succeeds with probability η and fails with probability $1-\eta$. Note that the optimal global strategy will not change with η, as there is only one round of measurement and therefore we cannot update our prior probabilities. As was shown by Chitambar and Hsieh [33], the optimal joint measurement on the double trine ensemble is $\pi_k^{\text{Joint}} = |\Omega_k\rangle\langle\Omega_k|$, where

$$|\Omega_k\rangle = \frac{1}{\sqrt{3}}[|00\rangle + \frac{1}{\sqrt{2}}e^{i\frac{2\pi}{3}k}(|01\rangle + |10\rangle) + e^{i\frac{4\pi}{3}k}|11\rangle]. \tag{5.19}$$

Similar to the single-state measurements, we then consider the cases where this measurement succeeds and fails in order to determine the total success probability. In the case where the joint measurement is successful, the probability of correctly guessing the state is

$$\begin{aligned} \mathrm{P}_{\text{Corr}}^{\pi^{\text{Joint}}} &= \frac{1}{3}\sum_k \langle\psi_k|\pi_k^{\text{Joint}}|\psi_k\rangle \\ &= \frac{1}{3}\left(1+\frac{1}{\sqrt{2}}\right)^2. \end{aligned} \tag{5.20}$$

If this joint measurement fails, we then guess the outcome of the measurement, which will be correct with probability

$$\mathrm{P}_{\text{Corr}}^{\text{Guess}} = \frac{1}{3}. \tag{5.21}$$

Thus the total probability of correctly identifying the state using the joint measurement is

$$\begin{aligned} \mathrm{P}_{\text{Corr}}^{\text{Joint}} &= \eta \mathrm{P}_{\text{Corr}}^{\pi^{\text{Joint}}} + (1-\eta)\mathrm{P}_{\text{Corr}}^{\text{Guess}} \\ &= \frac{1}{6}[2 + (1+2\sqrt{2})\eta)]. \end{aligned} \tag{5.22}$$

The optimal sequential measurement is compared to the joint measurement over all efficiencies η in Figure 5.8 and Figure 5.9. Note that this is not strictly a like-for-like comparison, as joint measurements are more technically challenging and thus generally have a low efficiency and lower fidelities. However, given the technical challenges associated with joint measurements and the relatively small gap between the efficacies of global and local measurements, this suggests that a local measurement is, in practice, sufficient for optimal measurement in this case. For very low efficiencies we even find that the local scheme outperforms the global scheme. This of course may be expected: it is well known that quantum information protocols requiring entanglement are sensitive to loss [78].

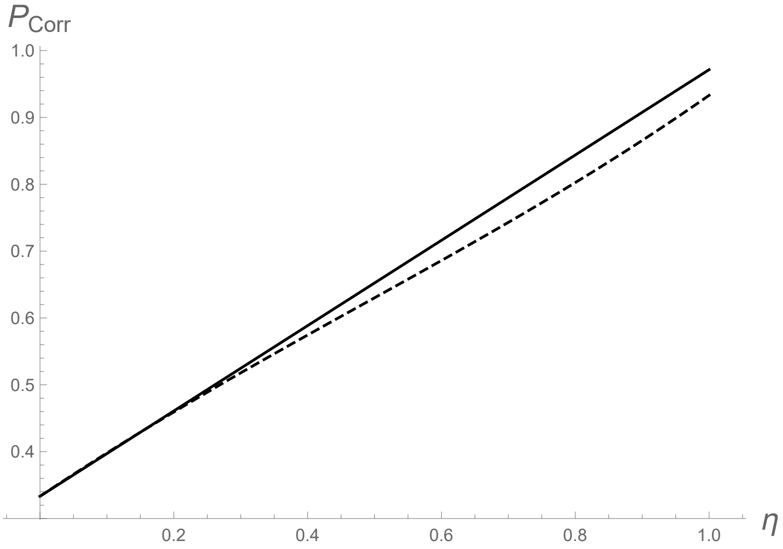

FIGURE 5.8: The probability of correctly guessing the signal state using the optimal joint measurement (bold) plotted against that for the optimal sequential measurement (dotted) for all values of η. Note that the sequential measurement is, for low efficiencies, better than the joint measurement. This may be more clearly seen in Figure 5.9.

5.3 Maximum confidence measurement of multiple trine copies

It is also possible to compare the effectiveness of local and global measurements for Maximum Confidence Measurement of the double trine ensemble – and indeed extend this to general n copies. As the states are linearly independent, a global measurement will simply take the form of unambiguous state discrimination [16, 30] and therefore will give the correct answer with 100% confidence - although, for the double-trine ensemble, such an outcome will only be given with probability 0.75 [30]. Taking into account the efficiency of the detection apparatus, η, we therefore have an expression for the

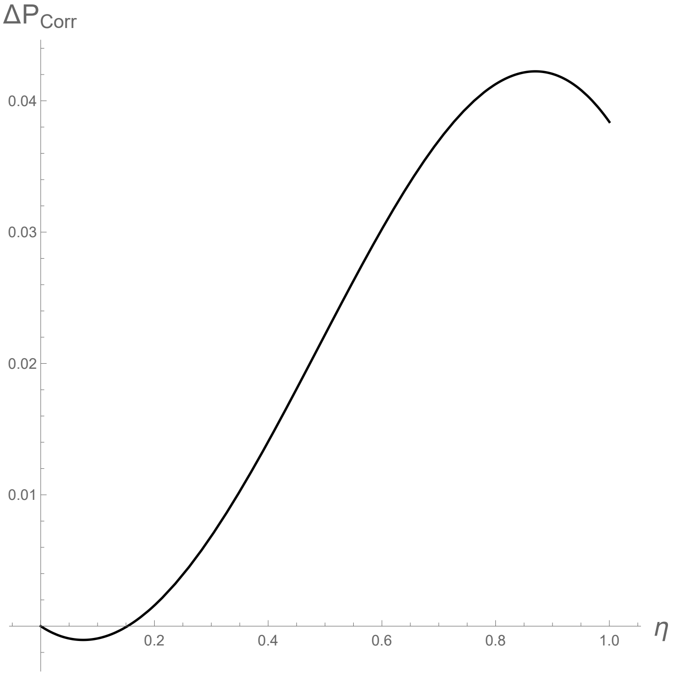

FIGURE 5.9: The difference between the probability of correctly guessing the signal state using the optimal joint measurement and the optimal sequential measurement. The small negative region corresponds to the values of η for which it is better to make the sequential measurement.

maximum probability of unambiguous discrimination globally, $\mathrm{P}_{\mathrm{Max}}^{\mathrm{Global}}$:

$$\mathrm{P}_{\mathrm{Max}}^{\mathrm{Global}} = 0.75\eta. \tag{5.23}$$

The local measurement scheme is also conceptually simple - if both sets of measurements are in the anti-trine basis, we simply need to calculate the probability that we obtain two different measurement results. For instance, if the first measurement tells us that the state ρ_0 was not sent (which we denote with $\bar{0}$), and the second tells us that ρ_1 was not sent (i.e., outcome $\bar{1}$), we know with certainty that ρ_2 was the signal state. Thus the local measurement scheme may also reach 100% confidence.

If both rounds of measurement are successful – which occurs with probability η^2 – there is a probability of $\frac{1}{2}$ that they will give different measurement outcomes. Any other set of outcomes will give an inconclusive result. We therefore have:

$$\mathrm{P}_{\mathrm{Max}}^{\mathrm{Local}} = \frac{1}{2}\eta^2. \tag{5.24}$$

We may extend this further to n copies of the trine states. As shown by Chefles [30], the global probability of unambiguous discrimination with measurement efficiency η is given by

$$\mathrm{P}_{\mathrm{Max}}^{\mathrm{Global}}(n) = \begin{cases} \eta(1-2^{-n}) & \text{if } n \text{ is even} \\ \eta(1-2^{-(n-1)}) & \text{if } n \text{ is odd,} \end{cases} \tag{5.25}$$

which has the curious property that, e.g., 3 copies are no more likely to be discriminated than 2 copies.

Extending the expression for local measurements to n copies requires a little more work. Suppose our first measurement is unsuccessful, with probability $1-\eta$. Our discrimination problem from this point on is now identical to the discrimination of $n-1$ copies of the signal state. We therefore expect the expression for $\mathrm{P}_{\mathrm{Max}}^{\mathrm{Local}}(n)$ to contain a term of the form $(1-\eta)\mathrm{P}_{\mathrm{Max}}^{\mathrm{Local}}(n-1)$.

Now suppose that state ρ_2 was sent, and that the first measurement is successful - i.e., we obtained outcome $\bar{0}$ or $\bar{1}$. This will occur with probability η. We now wish to find the probability that, after $n-1$ subsequent measurement rounds, we will obtain $\bar{1}$ or $\bar{0}$, respectively. In the second measurement round, only a successful measurement with a different outcome will suffice - this will occur with probability $\frac{\eta}{2}$. We may also obtain the same outcome, with probability $\frac{\eta}{2}$, or the measurement may fail with probability $1-\eta$. From here the problem takes on a recursive quality, with each subsequent measurement round having an $\frac{\eta}{2}$ chance of unambiguously identifying the signal state, and a $1-\frac{\eta}{2}$ chance of providing no new information (either by giving us the same outcome as before or by failing entirely). This may be seen in Figure 5.10.

Suppose our first measurement is successful, and we obtain our desired outcome after $m \leq n$ rounds of measurement. This means that there were exactly $m-2$ instances of the measurement giving no new information (with probability $1-\frac{\eta}{2}$). The probability of this happening is $\frac{\eta^2}{2}(1-\frac{\eta}{2})^{m-2}$. Bringing this together with the possibility of the first measurement failing, we obtain the following expression:

$$\mathrm{P}_{\mathrm{Max}}^{\mathrm{Local}}(n) = \sum_{m=0}^{n-2} \frac{\eta^2}{2}\left(1-\frac{\eta}{2}\right)^m + (1-\eta)\mathrm{P}_{\mathrm{Max}}^{\mathrm{Local}}(n-1) \tag{5.26}$$

with $\mathrm{P}^{\mathrm{Local}}_{\mathrm{Max}}(2) = \frac{1}{2}\eta^2$. We can simplify this further:

$$\mathrm{P}^{\mathrm{Local}}_{\mathrm{Max}}(n) = \sum_{j=1}^{n} \eta(1-\eta)^{j-1}\big[1-(1-\frac{\eta}{2})^{n-j}\big]. \tag{5.27}$$

This may be seen by realising that the term $\eta(1-\eta)^{j-1}$ gives the probability of receiving *any* answer (i.e. $\bar{i}$) after j rounds of measurement, while the term $1-(1-\frac{\eta}{2})^{n-j}$ gives the probability of receiving the other answer necessary for a definitive outcome (i.e. $\overline{i+1}$) after $n-j$ rounds of measurement.

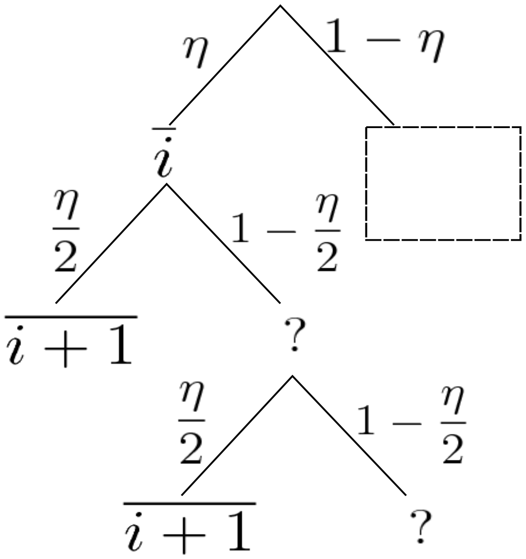

FIGURE 5.10: The probability tree for three copies of signal state ρ_2. Addition involving the index i is modulo 2, and the question mark denotes an inconclusive outcome (i.e. receiving outcome $\bar{i}$ again or the measurement failing). The dashed box represents the probability tree for two copies of the signal state. Note that the "fork" below $\bar{i}$ is repeated beneath the question mark - the tree has started recursing. This tree could therefore be extended to n copies of the signal state, with the dashed box representing the tree for $n-1$ copies.

It is worth noting that, in the case of the first $n-1$ measurement attempts failing, the maximum confidence measurement reverts to being the trine measurement, as it is simply the case of a single copy of one of the trine states [7]. However, we have neglected this from our analysis as we are only concerned with measurement outcomes which yield 100% confidence.

5.4 Conclusion

We have investigated bi-partite minimum-error discrimination and multi-partite maximum confidence measurement, and the effect of measurement efficiencies on these; this was made possible by the results of the previous chapter, further demonstrating the utility of an analytical solution to quantum state discrimination problems. We found that, for minimum-error measurement, the optimal strategy does not vary considerably with measurement efficiency η (in fact, one measurement is very close to optimal for all values of η), and that the joint measurement, while almost always better than the local measurement, is not sufficiently effective to justify the extra difficulty in performing experimentally. It is also interesting to note that, for low efficiencies, the sequential measurement actually outperforms the joint measurement. In the case of maximum confidence measurement, the probability of obtaining an unambiguous answer with local measurements increases each time we increase the number of copies of the states we possess, in contrast to the global measurement where it increases with every even number of copies possessed.

Chapter 6

Optimal sequential measurements for bi-partite state discrimination

6.1 Introduction

The work in this chapter is a useful illustration of the necessity and power of an analytical solution to the problem of single-qubit state discrimination, and how this utility can extend into the realm of bi-partite state discrimination [42]. As we will see, solving a bi-partite state discrimination problem will sometimes necessitate the solution of a single-qubit discrimination problem. In such instances, a simple analytical solution is very useful. The work in this chapter will also use the measurement trees first used in the previous chapter. In fact, this chapter may be seen as an extension of the work in the previous chapter: instead of looking at arbitrary detector efficiencies, we now investigate how to characterise the allowed measurements in the completely general case with arbitrary signal states.

State discrimination is a useful test problem with which to clarify the power and limitations of different classes of measurement. For information encoded across multiple quantum systems, the ability to measure jointly is strictly more powerful (but in general technologically more challenging) than the ability to measure each subsystem independently, even if many rounds of classical communication between systems are allowed. Intuitively, one might expect the difference in performance to be more pronounced when information is encoded in entangled states. That this is not necessarily the case was first revealed through two state discrimination problems. The first, so-called "nonlocality without entanglement," gave a set of multipartite orthogonal *product* states between which perfect discrimination is not possible using only local measurements and classical communication [41]. The second, complementary and no less surprising, showed that any two orthogonal pure states, regardless of entanglement or multipartite structure, may be perfectly discriminated using only sequential measurement, i.e., local measurement on each system, with classical feed-forward [40]. This was later extended to show that any two generally non-orthogonal pure states may be discriminated optimally by sequential measurement of the subsystems, according to the commonly used minimum error [79] and unambiguous discrimination strategies [80–82].

Beyond the two-state examples, the situation becomes much less clear: for the next simplest example of discriminating three possible qubit states given two copies, it was postulated by Peres and Wootters in 1991 that local measurement was strictly weaker than joint measurement on both copies [83], and only 20 years later was it finally proved that such a gap exists for this problem, for the minimum error strategy [33].

In this chapter we consider sequential measurements on a bipartite system; i.e. subsystem A and B are measured in turn, and the choice of measurement performed on subsystem B is allowed to depend in general on the result of measurement of A. This is often a physically relevant class of measurement; for example, if A and B are in different laboratories it is easy to imagine that feed-forward of measurement results from laboratory A to laboratory B would be practical, but many rounds of classical communication could become unfeasible. Alternatively, if A and B interact only weakly or not at all (e.g., photons), joint measurements are difficult to perform, while classical feed-forward from one detector to another apparatus is relatively easily achieved with current technology (see, e.g., Ref. [84] for such an experiment in the state discrimination context). It is natural then to ask how well information can be retrieved with this restriction on the measurement strategy that may be employed. Furthermore, implementations of joint measurement strategies for extracting information may provide applications for small quantum processors [85], and it is useful to understand when the additional experimental challenge of joint measurement may provide a significant advantage over local measurement strategies. For simplicity, we restrict to bipartite instead of the more general multipartite state discrimination.

We begin with the case where the bipartite state is simply a two-copy state. We construct necessary conditions that a given sequential measurement must satisfy to be optimal in the sense of minimising the error in determining the state, analogous to the well-known Helstrom conditions, equations (2.5) and (2.6) [29, 48]. We further find a condition which is both necessary and sufficient, but which requires optimisation over an arbitrary measurement on one subsystem. We illustrate the two-copy case through the example of the trine states considered in [33, 83], and give the probabilities of correctly identifying the state for sequential and global strategies, as well as discussing features of the optimal measurements in each case.

We extend the discussion to arbitrary bipartite states, and as an example give the optimal sequential strategies for discriminating three Bell states.

6.2 Two-copy state discrimination with sequential measurement

6.2.1 Necessary conditions

Let us consider the two-copy case, with sequential measurement. Suppose therefore we are provided with two copies of a state drawn from a known set $\{\rho_i\}$ with associated probabilities $\{p_i\}$. As we saw in the previous chapter, the allowed measurement procedures are as follows: make a measurement described by some POVM $\{M_j^A\}$ on system A; given outcome j, make a measurement on system B. This is shown in the tree in Figure 6.1. As the choice of measurement on system B can in general depend on the outcome of measurement on A, we denote the associated POVM $\{N_{i|j}^B\}$, where for all i and j, $N_{i|j}^B \geq 0$, and for each j,

$$\sum_i N_{i|j}^B = \mathbb{1}^B.$$

The measurement on the joint AB system is thus of the form $\{\pi_i = \sum_j M_j^A \otimes N_{i|j}^B\}$, with the probability of correctly identifying the state given by

$$\begin{aligned} \mathrm{P}_{\mathrm{Corr}} &= \sum_{ij} p_i \,\mathrm{Tr}_{AB}\left(\rho_i^A \otimes \rho_i^B M_j^A \otimes N_{i|j}^B\right) \\ &= \sum_{ij} p_i \,\mathrm{Tr}_A\left(\rho_i^A M_j^A\right) \mathrm{Tr}_B\left(\rho_i^B N_{i|j}^B\right). \end{aligned} \tag{6.1}$$

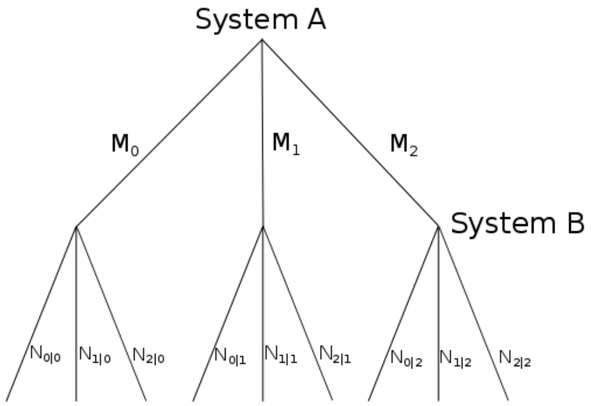

FIGURE 6.1: Probability tree showing sequential measurement notation. The measurement described by POVM $\{M_j\}$ is performed on system A. Given outcome j, the measurement described by POVM $\{N_{i|j}\}$ is performed on system B.

In the following we drop the superscripts A, B, whenever it is not confusing to do so. We begin by pointing out that each of $\{M_j\}$, $\{N_{i|j}\}$ may be interpreted as an optimal measurement for an appropriately defined discrimination problem, as follows. We first note that, given measurement result j on system A, we can update the probabilities as follows, using Bayes' rule:

$$\mathrm{P}(i|M_j) = \frac{\mathrm{P}(i, M_j)}{\mathrm{P}(M_j)} = \frac{p_i \operatorname{Tr}_A(\rho_i M_j)}{\sum_k p_k \operatorname{Tr}_A(\rho_k M_j)} = p_{i|j}. \tag{6.2}$$

Thus given result j on system A, the possible states $\{\rho_i\}$ of system B occur with probabilities $p_{i|j}$. Clearly $\{N_{i|j}\}$ should thus be optimal for discriminating the states ρ_i with the updated priors $p_{i|j}$, and thus a necessary condition is

$$\sum_i p_{i|j}\rho_i N_{i|j} - p_{k|j}\rho_k \geq 0, \quad \forall k,$$

or equivalently, using equation (6.2),

$$\sum_i p_i \operatorname{Tr}_A(\rho_i M_j)\rho_i N_{i|j} - p_k \operatorname{Tr}_A(\rho_k M_j)\rho_k \geq 0, \quad \forall k, \tag{6.3}$$

which must hold for each j. This set of conditions is necessary, but not sufficient (we have not done any optimisation over M_j). Finally, summing over j gives

$$\operatorname{Tr}_A\left(\sum_{i,j} p_i(\rho_i \otimes \rho_i)(M_j \otimes N_{i|j}) - p_k\rho_k \otimes \rho_k\right) \geq 0, \quad \forall k, \tag{6.4}$$

which is rather similar to the Helstrom condition (2.6), but with a partial trace over system A.

Conversely, we can rewrite equation (6.1) as follows:

$$\begin{aligned}
\mathrm{P}_{\mathrm{Corr}} &= \sum_j \mathrm{Tr}_A \left[\sum_i p_i \, \mathrm{Tr}_B \left(\rho_i^B N_{i|j}^B \right) \rho_i^A M_j^A \right] \\
&= \sum_j c_j \, \mathrm{Tr}_A \left(\sigma_j^A M_j^A \right),
\end{aligned}$$

where we have defined

$$\sigma_j^A = \frac{\sum_i p_i \, \mathrm{Tr}_B \left(\rho_i^B N_{i|j}^B \right) \rho_i^A}{\sum_k p_k \, \mathrm{Tr}_B \left(\rho_k^B N_{k|j}^B \right)}, \tag{6.5}$$

$$c_j = \sum_i p_i \, \mathrm{Tr}_B \left(\rho_i^B N_{i|j}^B \right). \tag{6.6}$$

We can interpret the trace one operators $\{\sigma_j\}$ as density operators, and if we further define probabilities $q_j = c_j / \left(\sum_i c_i \right)$, it follows that $\{M_j^A\}$ must be optimal for discriminating the states $\{\sigma_j^A\}$ with probabilities $\{q_j\}$. The Helstrom condition (2.6) then gives

$$\sum_j q_j \sigma_j^A M_j^A - q_k \sigma_k^A \geq 0,$$

which may be rewritten as

$$\sum_j \left[\sum_i p_i \, \mathrm{Tr}_B \left(\rho_i^B N_{i|j}^B \right) \rho_i^A \right] M_j^A - \sum_i p_i \, \mathrm{Tr}_B \left(\rho_i^B N_{i|k}^B \right) \rho_i^A \geq 0. \tag{6.7}$$

Finally, we obtain

$$\mathrm{Tr}_B \left(\sum_{ij} p_i (\rho_i^A \otimes \rho_i^B)(M_j^A \otimes N_{i|j}^B) - \sum_i p_i (\rho_i^A \otimes \rho_i^B)(\mathbb{1}^A \otimes N_{i|k}^B) \right) \geq 0. \tag{6.8}$$

Again, this is necessary, but not sufficient (this time we have not done any optimisation over $N_{i|j}$). One might hope that the conditions (6.3) and (6.8) when taken together are also sufficient, and could then imagine that it may be possible to construct an iterative procedure for numerical solution of the optimization problem. However, this turns out not to be the case; we will return to this point later. Each of conditions (6.3) and (6.8), however, have a clear interpretation. Note that it might have been expected that $\{N_{i|j}^B\}$

should be optimal for the updated priors given measurement of A; that M_j^A plays a complementary role for a different discrimination problem is less obvious *a priori*.

6.2.2 A necessary and sufficient condition

We now turn to the problem of simultaneously optimising both the measurement on A and that on system B. We find that the condition

$$\sum_{i,j} p_i \operatorname{Tr}_B(\rho_i N_{i|j})\rho_i M_j - \sum_k p_k \operatorname{Tr}_B(\rho_k \widetilde{N}_k)\rho_k \geq 0, \tag{6.9}$$

where $\{\widetilde{N}_k\}$ is any physically allowed measurement on system B, is both necessary and sufficient for optimality of $\{\pi_i = \sum_j M_j \otimes N_{i|j}\}$. Unfortunately, this still contains an arbitrary measurement on system B, and thus is not as readily applicable as the original Helstrom conditions to verify optimality of a candidate measurement. Nevertheless we will give examples in which it can be used to prove optimality analytically. We also note that the inclusion of an arbitrary measurement on one subsystem means that analysis beyond the bipartite case becomes complicated and our method is not readily extended to multipartite discrimination.

We will prove the necessity and sufficiency of this condition in a way which closely follows the proof of its single-qubit analogue, shown in §2.3.

We begin by proving the sufficiency of condition (6.9). If $\{\pi_i = \sum_j M_j \otimes N_{i|j}\}$ is optimal among sequential measurements, we require

$$\operatorname{Tr}_{AB}\left(\sum_{i,j} p_i(\rho_i \otimes \rho_i)(M_j \otimes N_{i|j})\right) \geq \operatorname{Tr}_{AB}\left(\sum_{k,l} p_k(\rho_k \otimes \rho_k)(M'_l \otimes N'_{k|l})\right),$$

for all $\{\pi'_k = \sum_l M'_l \otimes N'_{k|l}\}$. Inserting the identity $\sum_l M'_l \otimes \mathbb{1}$ and rearranging gives

$$\begin{aligned}\sum_l \operatorname{Tr}_{AB}\left[\left(\sum_{i,j} p_i(\rho_i \otimes \rho_i)(M_j \otimes N_{i|j}) - \sum_k p_k(\rho_k \otimes \rho_k)(\mathbb{1} \otimes N'_{k|l})\right) M'_l\right] &\geq 0,\\ \sum_l \operatorname{Tr}_A\left[\left(\sum_{i,j} p_i \operatorname{Tr}_B(\rho_i N_{i|j})\rho_i M_j - \sum_k p_k \operatorname{Tr}_B(\rho_k N'_{k|l})\rho_k\right) M'_l\right] &\geq 0.\end{aligned}$$

Condition (6.9) is therefore sufficient, if $\{\widetilde{N}_k\}$ is any allowed measurement on B.

That condition (6.9) is also necessary may be seen as follows: as in the unrestricted case, we introduce the manifestly Hermitian operator

$$\Gamma^A_{\text{sym}} = \sum_{i,j} p_i \operatorname{Tr}_B \left(\rho_i N_{i|j}\right) \frac{1}{2}\{\rho_i, M_j\}.$$

Suppose now that there exists some $|\lambda\rangle$ and some $\{\widetilde{N}_k\}$ such that

$$\langle\lambda|\Gamma^A_{\text{sym}} - \sum_k p_k \operatorname{Tr}_B(\rho_k \widetilde{N}_k)\rho_k|\lambda\rangle < 0.$$

We can construct a variation of $\{\pi_i = \sum_j M_j \otimes N_{i|j}\}$ as follows:

$$\begin{aligned}
M'_j &= (\mathbb{1} - \epsilon|\lambda\rangle\langle\lambda|)M_j(\mathbb{1} - \epsilon|\lambda\rangle\langle\lambda|), \quad 0 \le j < n \\
N'_{i|j} &= N_{i|j}, \quad 0 \le j < n \\
M_n &= \epsilon(2+\epsilon)|\lambda\rangle\langle\lambda|, \\
N_{i|n} &= \widetilde{N}_i,
\end{aligned}$$

where $0 < \epsilon \ll 1$. Note that if $\{M^A_j\}$ has n outcomes, the primed measurement on system A has $n+1$ outcomes. Now note that

$$\begin{aligned}
\mathrm{P}_{\text{Corr}}\left(\{M'_j \otimes N'_{i|j}\}\right) &= \mathrm{P}_{\text{Corr}}\left(\{M_j \otimes N_{i|j}\}\right) \\
&-\epsilon \operatorname{Tr}_{AB}\left(\sum_{i,j} p_i\rho_i \otimes \rho_i \left(|\lambda\rangle\langle\lambda|M_j + M_j|\lambda\rangle\langle\lambda|\right) \otimes N_{i|j}\right) \\
&+2\epsilon \operatorname{Tr}_{AB}\left(\sum_i p_i\rho_i \otimes \rho_i(|\lambda\rangle\langle\lambda| \otimes \widetilde{N}_i)\right) + O(\epsilon^2) \\
&= \mathrm{P}_{\text{Corr}}\left(\{M_j \otimes N_{i|j}\}\right) - 2\epsilon\langle\lambda|\Gamma^A_{\text{sym}} - \sum_i p_i \operatorname{Tr}_B(\rho_i\widetilde{N}_i)\rho_i|\lambda\rangle + O(\epsilon^2) \\
&> \mathrm{P}_{\text{Corr}}\left(\{M_j \otimes N_{i|j}\}\right).
\end{aligned}$$

Finally, we note that, by virtue of the fact that $\{M_j\}$ is an optimal measurement for discriminating the states σ_j, it follows that $\Gamma^A_{\text{sym}} = \Gamma^A$, where Γ^A is defined as

$$\Gamma^A = \sum_{i,j} p_i \operatorname{Tr}_B \left(\rho_i N_{i|j}\right) \rho_i M_j.$$

Thus we require

$$\Gamma^A - \sum_k p_k \operatorname{Tr}_B \left(\rho_k \widetilde{N}_k\right) \rho_k \geq 0,$$

which completes our proof.

6.3 Example: The double trine ensemble

As an example we consider the so-called double trine ensemble discussed in Chapter 5: two copies of the trine states, for which $\rho_j = |\psi_j\rangle\langle\psi_j|$, and

$$|\psi_j\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + e^{2\pi j i/3}|1\rangle\right).$$

These each occur with prior probabilities $p_j = \frac{1}{3}$ and have the symmetry property

$$|\psi_j\rangle = U^j|\psi_0\rangle$$

where U is a rotation of $\frac{2\pi}{3}$ around the z axis in the Bloch sphere.

6.3.1 Optimal sequential measurement

For the two-copy case, Chitambar and Hsieh [33] showed that – as we saw in §5.2.1 – the optimal sequential measurement rules out one state of the three in the first step, and corresponds to the Helstrom measurement to distinguish between the remaining two states in the second step. We first briefly present this optimal measurement and then use it to demonstrate our conditions.

The optimal sequential measurement thus makes the antitrine measurement, described previously in §4.1 and Figure 4.1. Following this measurement, the updated priors become $p_{i|j} = \frac{1}{2}(1 - \delta_{ij})$, and $\{N_{i|j}\}$ is then the optimal measurement to distinguish the two remaining equiprobable pure states $\{|\psi_i\rangle, |\psi_k\rangle, i \neq j \neq k\}$. This is a case of the well-known Helstrom measurement and is a projective measurement in a basis located symmetrically around the signal states (see, e.g., §4.2.1, §5.2.1, and [4]). Thus for $i = j$,

$N_{i|j} = 0$, and for $i \neq j$ we denote $N_{i|j} = |\phi_{i|j}\rangle\langle\phi_{i|j}|$, where

$$\begin{aligned}
|\phi_{1|0}\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle + i|1\rangle\right), \\
|\phi_{2|0}\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle - i|1\rangle\right), \\
|\phi_{0|1}\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle + e^{i\pi/6}|1\rangle\right) = U|\phi_{2|0}\rangle, \\
|\phi_{2|1}\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle - e^{i\pi/6}|1\rangle\right) = U|\phi_{1|0}\rangle, \\
|\phi_{0|2}\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle + e^{-i\pi/6}|1\rangle\right) = U^2|\phi_{1|0}\rangle, \\
|\phi_{1|2}\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle - e^{-i\pi/6}|1\rangle\right) = U^2|\phi_{2|0}\rangle.
\end{aligned}$$

These states, along with the trine and antitrine states, are shown in the Bloch sphere picture in Fig. 6.2.

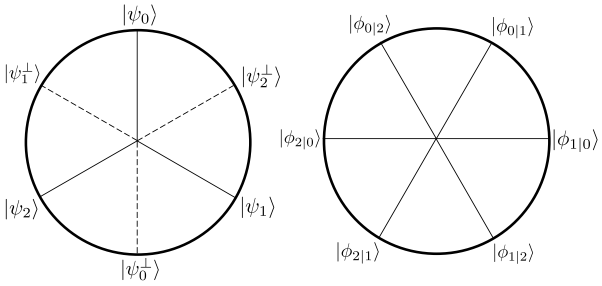

FIGURE 6.2: Trine and antitrine states shown on the equator of the Bloch sphere (left). Bases defined by the optimal Helstrom measurements in step two of the optimal sequential measurement procedure for discriminating the two-copy trine ensemble (right).

6.3.2 Necessary and sufficient conditions

We now use this strategy to illustrate the conditions presented in the previous section. From the symmetry we find that $\mathrm{Tr}(\rho_i N_{i|j}) = p_H(1 - \delta_{ij})$, for all i, j, where p_H is the probability of success of the Helstrom measurement distinguishing between two equiprobable states with overlap $|\langle\psi_i|\psi_k\rangle| = |\langle\psi_0|\psi_1\rangle| = 1/2$, i.e., from [5]:

$$p_H = \frac{1}{2}\left(1 + \sqrt{1 - |\langle\psi_0|\psi_1\rangle|^2}\right) = \frac{1}{2}\left(1 + \frac{\sqrt{3}}{2}\right).$$

By construction, this measurement strategy satisfies condition (6.3). To evaluate (6.8) and the necessary and sufficient condition (6.9), we first calculate Γ^A:

$$\begin{aligned}
\Gamma^A &= \sum_{i,j} p_i \operatorname{Tr}\left(\rho_i N_{i|j}\right) \rho_i M_j \\
&= \sum_{i,j} \frac{1}{3} p_H (1-\delta_{ij}) \left(|\psi_i\rangle\langle\psi_i|\right) \left(\frac{2}{3}|\psi_j^{\perp}\rangle\langle\psi_j^{\perp}|\right) \\
&= \frac{1}{3} p_H \left(\sum_i |\psi_i\rangle\langle\psi_i|\right) \left(\sum_j \frac{2}{3}|\psi_j^{\perp}\rangle\langle\psi_j^{\perp}|\right) \\
&= \frac{1}{2} p_H \mathbb{1} = \frac{1}{4}\left(1+\frac{\sqrt{3}}{2}\right)\mathbb{1},
\end{aligned}$$

where in the last line we have used $\sum_j \frac{2}{3}|\psi_j^{\perp}\rangle\langle\psi_j^{\perp}| = \sum_j \frac{2}{3}|\psi_j\rangle\langle\psi_j| = \mathbb{1}$. We first show that the strategy satisfies condition (6.8). We obtain

$$\begin{aligned}
\sum_i p_i \operatorname{Tr}\left(\rho_i N_{i|j}\right) \rho_i^A &= \frac{1}{3}\sum_i p_H (1-\delta_{ij}) \rho_i \\
&= \frac{1}{2} p_H \left(\mathbb{1} - \frac{2}{3}\rho_j\right),
\end{aligned}$$

from which it is clear that condition (6.8) is satisfied for each j. Finally, to prove that this is indeed the optimal strategy, we must show that it satisfies the necessary and sufficient condition (6.9). As we have shown that Γ^A is proportional to the identity, this amounts to showing that for any allowed measurement $\{\widetilde{N}_k\}$ on system B, the largest eigenvalue of the operator

$$\sum_k p_k \operatorname{Tr}\left(\rho_k \widetilde{N}_k\right) \rho_k$$

is bounded by $\frac{1}{2} p_H = \frac{1}{4}\left(1+\frac{\sqrt{3}}{2}\right)$. The proof that this holds is straight-forward using results from previous chapters, but needs a few steps; the details are given in Appendix 6.6.1.

The probability of correctly identifying the state using the optimal sequential measurement is given by

$$\mathrm{P}_{\text{Corr}}^{seq} = \operatorname{Tr}(\Gamma^A) = p_H = \frac{1}{2}\left(1+\frac{\sqrt{3}}{2}\right) \simeq 0.933.$$

6.3.3 Comparison of global and sequential schemes

For comparison we recall the globally optimal measurement strategy, also discussed in [33]. Recall that the double-trine ensemble satisfies $|\psi_i\rangle|\psi_i\rangle = (U\otimes U)^i|\psi_0\rangle|\psi_0\rangle$, where U is a rotation of $\frac{2\pi}{3}$ around the z axis in the Bloch sphere. For sets with such symmetry the optimal measurement was shown by Ban *et al* [34] to be given by the so-called square-root measurement (also known as the "pretty-good measurement" [52]), seen in §2.3.2. In this case, the optimal measurement corresponds to a projective measurement, with operators $\{\Pi_j = |\Phi_j\rangle\langle\Phi_j|\}$, where

$$|\Phi_j\rangle = \frac{1}{\sqrt{3}}\left(|0\rangle|0\rangle + e^{2\pi ji/3}\frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) + e^{4\pi ji/3}|1\rangle|1\rangle\right). \tag{6.10}$$

The probability of correctly identifying the state is

$$\mathrm{P}^{glob}_{\mathrm{Corr}} = \frac{1}{2} + \frac{\sqrt{2}}{3} \simeq 0.971.$$

Note that the probability of identifying the state correctly achieved by the optimal sequential measurement is greater than 96% of that achieved by the optimal global measurement. In systems where joint measurement is technologically challenging it is thus perhaps difficult to argue that the additional experimental effort is merited by the improvement in performance in this case.

We comment finally on the optimal sequential measurement as an approximation to the optimal global measurement. For the optimal sequential measurement, given above, we obtain

$$\begin{aligned}\pi_0 &= \frac{2}{3}\left(|\psi_1^\perp\rangle\langle\psi_1^\perp| \otimes |\phi_{0|1}\rangle\langle\phi_{0|1}| + |\psi_2^\perp\rangle\langle\psi_2^\perp| \otimes |\phi_{0|2}\rangle\langle\phi_{0|2}|\right)\\ \pi_1 &= (U\otimes U)\pi_0(U\otimes U)^\dagger\\ \pi_2 &= (U\otimes U)^2\pi_0((U\otimes U)^\dagger)^2.\end{aligned} \tag{6.11}$$

Considering π_0, after a little algebra we find

$$\begin{aligned}|\psi_1^\perp\rangle\otimes|\phi_{0|1}\rangle &= \frac{1}{2}e^{-\pi i/12}\left[\sqrt{1+2\cos^2\frac{\pi}{12}}|\alpha_0\rangle + i\sqrt{1+2\sin^2\frac{\pi}{12}}|\beta_0\rangle\right],\\ |\psi_2^\perp\rangle\otimes|\phi_{0|2}\rangle &= \frac{1}{2}e^{\pi i/12}\left[\sqrt{1+2\cos^2\frac{\pi}{12}}|\alpha_0\rangle - i\sqrt{1+2\sin^2\frac{\pi}{12}}|\beta_0\rangle\right],\end{aligned}$$

where

$$\begin{aligned}|\alpha_0\rangle &= \left(1+2\cos^2\frac{\pi}{12}\right)^{-1/2}\left(\cos\frac{\pi}{12}|00\rangle+\frac{1}{\sqrt{2}}\left(|01\rangle+|10\rangle\right)+\cos\frac{\pi}{12}|11\rangle\right),\\ |\beta_0\rangle &= \left(1+2\sin^2\frac{\pi}{12}\right)^{-1/2}\left(\sin\frac{\pi}{12}|00\rangle+\frac{1}{\sqrt{2}}\left(|01\rangle-|10\rangle\right)-\sin\frac{\pi}{12}|11\rangle\right).\end{aligned}$$

Thus we can write

$$\begin{aligned}\pi_0 &= \frac{1}{3}\left(1+2\cos^2\frac{\pi}{12}\right)|\alpha_0\rangle\langle\alpha_0|+\frac{1}{3}\left(1+2\sin^2\frac{\pi}{12}\right)|\beta_0\rangle\langle\beta_0|\\ &= \frac{1}{3}\left(2+\frac{\sqrt{3}}{2}\right)|\alpha_0\rangle\langle\alpha_0|+\frac{1}{3}\left(2-\frac{\sqrt{3}}{2}\right)|\beta_0\rangle\langle\beta_0|.\end{aligned}$$

Note that $\langle\alpha_0|\beta_0\rangle = 0$, and hence this is the eigendecomposition of the operator. We further note that $|\beta_0\rangle$ is orthogonal to the signal state $|\psi_0\rangle|\psi_0\rangle$ and thus does not contribute to the probability of identifying the state. The remaining eigenvector $|\alpha_0\rangle$ is an approximation to $|\Phi_0\rangle$, the state onto which the optimal global measurement projects, which is, in fact, an amazingly good one: it turns out $|\langle\alpha_0|\Phi_0\rangle|^2 = 0.9997$. Due to the weighting factor, the overlap between $|\Phi_0\rangle$ and π_0 is given by $\langle\Phi_0|\pi_0|\Phi_0\rangle = \frac{1}{3}\left(2+\frac{\sqrt{3}}{2}\right)|\langle\alpha_0|\Phi_0\rangle|^2 = 0.9551$.

The state $|\Phi_0\rangle$ is thus very close to a superposition of $|\psi_1^\perp\rangle\otimes|\phi_{0|1}\rangle$ and $|\psi_2^\perp\rangle\otimes|\phi_{0|2}\rangle$, with appropriate normalisation:

$$|\Phi_0\rangle \simeq |\alpha_0\rangle = \left(1+2\cos^2\frac{\pi}{12}\right)^{-1/2}\left(e^{\pi i/12}|\psi_1^\perp\rangle\otimes|\phi_{0|1}\rangle+e^{-\pi i/12}|\psi_2^\perp\rangle\otimes|\phi_{0|2}\rangle\right).$$

The optimal sequential measurement, on the other hand, is formed from a *mixture* of projectors onto these same states. It gives additional information — one state is ruled out with certainty — at the expense of a slightly lower probability of success.

6.3.4 A non-optimal sequential measurement

The example of the trine states is further illuminating, as there exists another measurement strategy which satisfies both necessary conditions (6.3) and (6.8) but which is not an optimal strategy, thus demonstrating that these two conditions, when taken together, are not sufficient to define the optimal measurement. This strategy is to perform the optimal minimum error measurement at each step, with Bayesian update of the probabilities in between measurements. Note that such a strategy is known to be optimal (and in fact performs as well as the best joint measurement) for a different set of states — the case of just two pure states [86, 87]. For the trine states, the measurement is as follows: $\{M_j\}$ is the optimal one-copy minimum error measurement, which consists of weighted projectors onto the trine states themselves [5, 34], $M_j = \frac{2}{3}|\psi_j\rangle\langle\psi_j|$. Note that for the trine states $|\langle\psi_i|\psi_j\rangle|^2 = \frac{1}{4}(1+3\delta_{ij})$, and thus the updated priors upon obtaining outcome j are, using equation (6.2),

$$p_{i|j} = \frac{\frac{2}{3}|\langle\psi_i|\psi_j\rangle|^2}{\frac{2}{3}\sum_k |\langle\psi_k|\psi_j\rangle|^2} = \frac{1}{6} + \frac{1}{2}\delta_{ij}.$$

For each j, the states with these probabilities have so-called mirror symmetry — the set is invariant under reflection about $|\psi_j\rangle$. For such a set, the minimum error problem was considered by Andersson *et al.* [6]. Using their results we find for $j = 0$ the optimal measurement is of the form:

$$\begin{aligned} N_{0|0} &= (1-a^2)|\psi_0\rangle\langle\psi_0|, \\ N_{1|0} &= \frac{1}{2}\left(a|\psi_0\rangle - i|\psi_0^\perp\rangle\right)\left(a\langle\psi_0| + i\langle\psi_0^\perp|\right), \\ N_{2|0} &= \frac{1}{2}\left(a|\psi_0\rangle + i|\psi_0^\perp\rangle\right)\left(a\langle\psi_0| - i\langle\psi_0^\perp|\right), \end{aligned}$$

where a depends on the geometry of the set and the prior probabilities [6].This is also discussed in §4.2.2. For our case we find $a = \frac{1}{5\sqrt{3}}$. The optimal measurements for $j = 1, 2$ are obtained by symmetry $\{N_{i|j} = U^j N_{i|0} (U^j)^\dagger\}$. Note that condition (6.3) is satisfied by construction. Turning to condition (6.8), we find

$$\begin{aligned} \mathrm{Tr}\left(\rho_0 N_{0|0}\right) &= \frac{74}{75}, \\ \mathrm{Tr}\left(\rho_1 N_{1|0}\right) &= \mathrm{Tr}\left(\rho_2 N_{2|0}\right) = \frac{32}{75}, \end{aligned}$$

with analogous results for $j = 1, 2$. Concisely, $\mathrm{Tr}\left(\rho_i N_{i|j}\right) = \frac{32}{75} + \frac{42}{75}\delta_{ij}$. Finally, we can calculate Γ^A:

$$\begin{aligned}\Gamma^A &= \sum_{i,j} p_i \,\mathrm{Tr}\left(\rho_i N_{i|j}\right)\rho_i M_j \\ &= \sum_{i,j} \frac{1}{3}\left(\frac{32}{75} + \frac{42}{75}\delta_{ij}\right)\left(|\psi_i\rangle\langle\psi_i|\right)\left(\frac{2}{3}|\psi_j\rangle\langle\psi_j|\right) \\ &= \frac{30}{75}\mathbb{1} = \frac{2}{5}\mathbb{1}.\end{aligned}$$

For $c_j\sigma_j$ we obtain

$$\begin{aligned}\sum_i p_i \,\mathrm{Tr}\left(\rho_i N_{i|j}\right)\rho_i^A &= \frac{1}{3}\sum_i \left(\frac{32}{75} + \frac{42}{75}\delta_{ij}\right)\rho_i \\ &= \frac{16}{75}\mathbb{1} + \frac{14}{75}\rho_j \\ &= \frac{2}{5}|\psi_j\rangle\langle\psi_j| + \frac{16}{75}|\psi_j^\perp\rangle\langle\psi_j^\perp|\end{aligned}$$

from which it is clear that condition (6.8) is satisfied for each j.

An analogous situation arises in state discrimination maximising the mutual information between sender and receiver. A necessary but not sufficient condition is known, and for the example of the trine states, is satisfied by both the trine measurement, which is not optimal [34], and the antitrine measurement, which is optimal [28]. We finally note that the probability of correctly identifying the state using this scheme, $\mathrm{Tr}(\Gamma^A) = \frac{4}{5}$, is considerably worse than that given by the optimal sequential measurement shown above.

6.4 General bi-partite case

6.4.1 Necessary and sufficient conditions

Above, for simplicity, we confined our discussion of optimal sequential measurement strategies to the case of two-copy state discrimination. The conditions obtained, however, are easily extended to the general bipartite case. Suppose, therefore, we are provided with a bi-partite state drawn from a known set $\{\rho_i^{AB}\}$, with known *a priori* probabilities $\{p_i\}$. If our measurement strategy is restricted to sequential measurements on each subsystem, with feed-forward, what is the best measurement to make? The allowed measurements on the joint AB system are again described by POVMs of the form $\{\pi_i = \sum_j M_j^A \otimes N_{i|j}^B\}$, and the probability of correctly identifying the state is expressed:

$$\mathrm{P_{Corr}} = \sum_{ij} p_i \, \mathrm{Tr}_{AB} \left(\rho_i^{AB} M_j^A \otimes N_{i|j}^B \right).$$

Following the same reasoning as in Sec. III, the necessary conditions, given in equations (6.3) and (6.8), become

$$\begin{aligned} \sum_i p_i \, \mathrm{Tr}_A \left(\rho_i^{AB} M_j \right) N_{i|j} - p_k \, \mathrm{Tr}_A \left(\rho_k^{AB} M_j \right) &\geq 0, \\ \sum_{i,j} p_i \, \mathrm{Tr}_B \left(\rho_i^{AB} N_{i|j} \right) M_j - \sum_i p_i \, \mathrm{Tr}_B \left(\rho_i^{AB} N_{i|k} \right) &\geq 0, \end{aligned}$$

with the following interpretation. Given a measurement M_j^A on system A, $\{N_{i|j}^B\}$ must be optimal for discriminating the updated states $\sigma_{i|j}^B$, occurring with probabilities $p_{i|j}$:

$$\begin{aligned} \sigma_{i|j}^B &= \frac{\mathrm{Tr}_A \left(\rho_i^{AB} M_j^A \right)}{\mathrm{Tr}_{AB} \left(\rho_i^{AB} M_j^A \right)}, \\ p_{i|j} &= \frac{p_i \, \mathrm{Tr}_{AB} \left(\rho_i^{AB} M_j^A \right)}{\sum_k p_k \, \mathrm{Tr}_{AB} \left(\rho_k^{AB} M_j^A \right)}. \end{aligned}$$

Similarly, given measurements $\{\{N_{i|j}^B\}\}$ on system B, $\{M_j^A\}$ must be optimal for discriminating the states σ_j^A, occurring with probabilities q_j:

$$\begin{aligned} \sigma_j^B &= \frac{\sum_i p_i \, \mathrm{Tr}_B \left(\rho_i^{AB} N_{i|j}^B \right)}{\sum_k p_k \, \mathrm{Tr}_{AB} \left(\rho_k^{AB} N_{k|j}^B \right)}, \\ q_j &= \frac{\sum_i p_i \, \mathrm{Tr}_{AB} \left(\rho_i^{AB} N_{i|j}^B \right)}{\sum_l \sum_k p_k \, \mathrm{Tr}_{AB} \left(\rho_k^{AB} N_{k|l}^B \right)}. \end{aligned}$$

Finally, following the same argument as in Sec. III, the necessary and sufficient condition for optimality of $\{\pi_j = \sum_j M_j \otimes N_{i|j}\}$ for discriminating the general bipartite states $\{\rho_i^{AB}\}$ becomes

$$\sum_{i,j} p_i \, \mathrm{Tr}_B \left(\rho_i^{AB} N_{i|j} \right) M_j - \sum_k p_k \, \mathrm{Tr}_B \left(\rho_k^{AB} \widetilde{N}_k \right) \geq 0, \tag{6.12}$$

where $\{\widetilde{N}_k\}$ is any physically allowed measurement on system B.

6.4.2 Example: Three Bell states

As an example of the general case, we consider the simple case of discriminating between three Bell states $\rho_i^{AB} = |\Psi_i\rangle\langle\Psi_i|$:

$$\begin{aligned} |\Psi_0\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle|0\rangle + |1\rangle|1\rangle\right), \\ |\Psi_1\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle|1\rangle + |1\rangle|0\rangle\right), \\ |\Psi_2\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle|0\rangle - |1\rangle|1\rangle\right), \end{aligned}$$

occurring with equal probabilities $p_i = \frac{1}{3}$. Although perfect discrimination between any two Bell states is possible by only local measurements and feed-forward (for example, to distinguish between $|\Psi_0\rangle$ and $|\Psi_1\rangle$ one need only measure both systems in the $\{|0\rangle, |1\rangle\}$ basis and look at the correlations between outcomes) it is known that for more than two states this is no longer possible [88, 89]. To distinguish between all three states, one strategy is to simply perform the measurement that perfectly distinguishes any two states and never identify the third. We show that this strategy is optimal in terms of minimising the probability of error.

Consider, therefore, the measurement

$$\begin{aligned} M_0 &= |0\rangle\langle 0|, & M_1 &= |1\rangle\langle 1|, \\ N_{0|0} &= |0\rangle\langle 0|, & N_{1|0} &= |1\rangle\langle 1|,\ N_{2|0} = 0, \\ N_{0|1} &= |1\rangle\langle 1|, & N_{1|1} &= |0\rangle\langle 0|,\ N_{2|1} = 0, \end{aligned}$$

that is, both Alice and Bob measure in the $\{|0\rangle, |1\rangle\}$ basis. Bob takes outcome 0 to indicate state $|\Psi_0\rangle$, and outcome 1 to indicate state $|\Psi_1\rangle$. State $|\Psi_2\rangle$ is never identified.

It is useful to rewrite equation (6.12) as follows:

$$\Gamma^A - \widetilde{c}\,\widetilde{\sigma} \geq 0,$$

where

$$\begin{aligned}
\Gamma^A &= \sum_{i,j} p_i \operatorname{Tr}_B \left(\rho_i^{AB} N_{i|j}\right) M_j, \\
\widetilde{c} &= \sum_k p_k \operatorname{Tr}_{AB} \left(\rho_k^{AB} \widetilde{N_k}\right), \\
\widetilde{\sigma} &= \frac{1}{\widetilde{c}} \sum_k p_k \operatorname{Tr}_B \left(\rho_k^{AB} \widetilde{N}_k\right).
\end{aligned}$$

Note that $\widetilde{\sigma}$ is a density operator. Furthermore, it is straight-forward to show that

$$\Gamma^A = \sum_{i,j} p_i \operatorname{Tr}_B \left(\rho_i^{AB} N_{i|j}\right) M_j = \frac{1}{3}\mathbb{1}^A$$

while

$$\begin{aligned}
\widetilde{c} &= \frac{1}{3} \sum_k \operatorname{Tr}_B \left(\operatorname{Tr}_A(\rho_k^{AB}) \widetilde{N}_k\right) \\
&= \frac{1}{3} \sum_k \operatorname{Tr}_B \left[\left(\frac{1}{2}\mathbb{1}^B\right) \widetilde{N}_k\right] \\
&= \frac{1}{3},
\end{aligned}$$

where the second line follows as the reduced density operator for system B in all cases is proportional to the identity $\mathbb{1}^B$, and the last line follows from the POVM condition $\sum \widetilde{N}_k = \mathbb{1}^B$. Thus the condition (6.12) becomes

$$\mathbb{1} - \widetilde{\sigma} \geq 0,$$

which is true for any arbitrary density operator $\widetilde{\sigma}$. Thus $\{M_j \otimes N_{i|j}\}$ is an optimal measurement among sequential strategies for discrimination of the three Bell states. A similar approach could be taken to the problem of discriminating all four Bell states, but this complicates the analysis without changing the basic conclusions.

6.5 Discussion

We have discussed the problem of extracting classical information from a set of bipartite states, when the measurement strategy is restricted to sequential measurements of each subsystem, with feed-forward of classical information in between measurements. As this is a physically well-motivated class, it is useful to understand how well it performs compared to the ability to perform arbitrary joint measurements, which in many physical systems is still technologically challenging. We have constructed an analogue of the Helstrom conditions for sequential measurement strategies. It is not obvious how to use this condition to construct an optimal measurement, but we show how for certain examples it is possible to use the condition to prove optimality of a candidate measurement procedure.

Our necessary and sufficient condition for optimality of a given sequential measurement still contains an arbitrary measurement on one subsystem. We have been unable to find a condition which is both necessary and sufficient and requires only the set of states and a candidate measurement. It would certainly be useful to find one, but in the absence of such, given a candidate optimal measurement our condition reduces the complexity of checking optimality from optimising over both systems to optimising over just one. It would also be interesting in the future to extend this analysis to other figures of merit, such as those which interpolate between minimum-error and unambiguous discrimination [90], or which maximise the success rate of discrimination while allowing for inconclusive results [71, 91].

For the two-copy trine case, the probability of success of the optimal sequential measurement is 96 % of the value achieved by the optimal global measurement [33]. The optimal sequential measurement sometimes rules out one of the states with certainty, thus providing information not given by the optimal global strategy, at the expense of a slightly higher probability of failure. Nonetheless, the difference in performance is arguably too small to motivate experimental implementation of the joint measurement.

6.6 Appendix

6.6.1 Proof of optimality for the double trine ensemble

To prove optimality of the sequential measurement scheme given in the text, we wish to show that the largest eigenvalue of

$$\widetilde{\sigma} = \frac{1}{3}\sum_k \mathrm{Tr}(\rho_k \widetilde{N}_k)\rho_k$$

is less than or equal to $\frac{1}{2}p_H = \frac{1}{4}\left(1 + \frac{\sqrt{3}}{2}\right)$ for any physically allowed measurement $\{\widetilde{N}_k\}$. We begin by writing the trine states ρ_j in the Bloch sphere representation:

$$\rho_j = \frac{1}{2}\left[I + \cos\left(\frac{2\pi j}{3}\right)\sigma_x + \sin\left(\frac{2\pi j}{3}\right)\sigma_y\right].$$

Writing $s_k = \frac{1}{3}\,\mathrm{Tr}\left(\rho_k \widetilde{N}_k\right)$ we thus obtain

$$\widetilde{\sigma} \;=\; \frac{1}{2}\left\{(s_0 + s_1 + s_2)I + \left[s_0 - \frac{1}{2}(s_1 + s_2)\right]\sigma_x + \frac{\sqrt{3}}{2}(s_1 - s_2)\sigma_y\right\}$$

with eigenvalues

$$\begin{aligned}
\lambda_\pm &= \frac{1}{2}(s_0 + s_1 + s_2) \pm \frac{1}{2}\left(\sqrt{\left[s_0 - \frac{1}{2}(s_1 + s_2)\right]^2 + \left[\frac{\sqrt{3}}{2}(s_1 - s_2)\right]^2}\right) \\
&= \frac{1}{2}(s_0 + s_1 + s_2) \pm \frac{1}{2}|s_0 + e^{\frac{2\pi i}{3}}s_1 + e^{-\frac{2\pi i}{3}}s_2|.
\end{aligned}$$

Thus it follows that there exists some θ such that the largest eigenvalue λ_+ may be written

$$\begin{aligned}
\lambda_+ &= \frac{1}{2}(s_0 + s_1 + s_2) \\
&\quad + \frac{1}{2}e^{i\theta}\left(s_0 + e^{\frac{2\pi i}{3}}s_1 + e^{-\frac{2\pi i}{3}}s_2\right), \\
&= \frac{1}{2}(1 + \cos\theta)s_0 + \frac{1}{2}\left[1 + \cos\left(\theta + \frac{2\pi}{3}\right)\right]s_1 + \frac{1}{2}\left[1 + \cos\left(\theta - \frac{2\pi}{3}\right)\right]s_2 \\
&= \frac{1}{2}\left[\sum_k q_k\,\mathrm{Tr}\left(\rho_k \widetilde{N}_k\right)\right],
\end{aligned}$$

where in the second equality we use the fact that λ_+ is real, and in the last line we have substituted for s_k, and defined $q_k = \frac{1}{3}\left[1+\cos\left(\theta+\frac{2\pi k}{3}\right)\right]$. Each strategy $\{\widetilde{N}_k\}$ thus defines a θ such that the above equalities hold. For each such θ, we can find an upper bound for λ_+ by considering the optimisation problem of discriminating the states $\{\rho_k\}$ occurring with priors q_k:

$$\lambda_+ \leq \frac{1}{2}\mathrm{P}_{\mathrm{Corr}}\left(\{q_k\rho_k\}\right).$$

We thus wish to find the optimal strategy $\{\pi_k\}$ for discriminating the trine states with *a priori* probabilities $\frac{1}{3}\left[1+\cos\left(\theta+\frac{2\pi k}{3}\right)\right]$, ultimately maximising the probability of correctness also over θ. Finally, if this maximum is achievable then we have succeeded in finding the optimal λ_+.

Thankfully we have already solved this problem, as these probabilities are exactly those given in equation (5.15), up to reordering. We find, as shown in Figure 5.4, that the maximum value of this function corresponds to (in this example) $\theta = -\frac{\pi}{3}$, corresponding to

$$\mathrm{P}_{\mathrm{Corr}}(\{q_k\rho_k\}) \quad \leq \quad \frac{1}{2}\left(1+\frac{\sqrt{3}}{2}\right).$$

Thus we obtain $\lambda_+ \leq \frac{1}{4}\left(1+\frac{\sqrt{3}}{2}\right)$, as desired.

Chapter 7

Conclusions

In this thesis we have investigated quantum state discrimination in a number of forms. We have primarily focussed on the problem of minimum-error state discrimination, finding a simple, closed-form general analytic solution for arbitrary single-qubit states [39]. We then used this solution to aid our understanding of related problems; for instance, in the problem of bi-partite state discrimination, where the prior probabilities of each signal state get updated depending on the measurement outcomes that have previously been received [42, 77] We also used our work to give the first full analytic solution to the problem of discriminating between the trine states with minimum error [73]. This is the first ensemble for which such a solution has been given since the two-state solution given by the Helstrom bound. We shall now briefly review the findings of each chapter in turn. In Chapter 3, we found a simple solution to the problem of minimum-error single-qubit state discrimination for arbitrary states (including mixed states) and arbitrary prior probabilities. This solution takes the form of a simple one-line equation – equation (3.19) – in contrast to the geometric approaches in the literature. This method produces a series of linear equations from which one can construct the Lagrange operator Γ and therefore the optimal POVM to use for minimum-error discrimination. This also has uses beyond the single-qubit case as, as we demonstrated in Chapter 5, we may use this to help solve multi-qubit state discrimination problems. This method has some element of trial-and-error attached to it, as it does not tell us which subset of states it is optimal to measure. However, if the wrong subset of states is inserted into equation (3.19), we expect that the candidate Γ may be seen to be erroneous by, e.g., giving a value for $\mathrm{P}_{\mathrm{Corr}}$ which is greater than 1, or yielding a negative operator as a POVM element.

In Chapter 4, we used the method detailed in Chapter 3 to perform an in-depth study of the minimum-error measurement strategy for the trine states with arbitrary prior probabilities, yielding the first such complete analysis for a set of states since the Helstrom bound was introduced. We found, somewhat surprisingly, that for most of the parameter space, the optimal POVM need not identify every signal state. We then investigated the maximum confidence measurement for the trine states with arbitrary prior probabilities, and compared this to the minimum-error strategy. This showed that the minimum-error measurement is very close in confidence to the maximum confidence measurement.

In Chapter 5, we extended the work from Chapter 4 to the double-trine ensemble, finding the optimal bipartite measurement for arbitrary detector efficiency, then comparing this to the optimal global measurement. In this, we found that one bipartite measurement strategy is close to optimal for all values of efficiency η. In this scheme, we first perform the measurement described in equation (5.14) with $\phi = \frac{\pi}{6}$, then follow this with the optimal two-element POVM, described in equation (4.2) with p_0 and p_1 given by equation (5.15). This measurement scheme is close in efficacy to the optimal global measurement described in equation (5.19) for all values of η. Given the additional difficulty in performing joint measurements, it is therefore questionable as to whether the additional effort exerted to perform a joint measurement is justifiable in this case. In fact, in a surprising result, we found that for certain small values of η, the optimal measurement is the sequential measurement. Also in Chapter 5, we investigated the maximum confidence measurement of multiple trine copies and again compared the local and global strategies for arbitrary detector efficiency η. In the local measurement, the probability of obtaining an unambiguous outcome increases with each copy of the states sent; however, in the global case, this probability only increases with each even number of copies possessed. In Chapter 6, we introduced a generalised version of the Helstrom conditions which can be used to test for optimality of bipartite sequential measurements. We then tested these conditions on the double-trine ensemble, for which we had already found the optimal measurement, and an ensemble of three of the four Bell states. However, we also showed that a non-optimal measurement can satisfy these conditions, showing that, when taken together, these conditions are not sufficient to identify the optimal measurement. Nevertheless, we did find a necessary and sufficient condition. We hope that a solution to this condition may further pave the way for future problems in multipartite state discrimination.

In the future, it would be interesting to see if our simple closed-form solution has any qudit analogue. Furthermore, the implications of our work beyond the bi-partite regime (i.e. into tri-partite state discrimination and beyond) could further augment our understanding of the process of quantum state discrimination. We hope such results will be found, and prove useful in the world of quantum computing and quantum information.

Bibliography

[1] Roger BM Clarke, Anthony Chefles, Stephen M Barnett, and Erling Riis. Experimental demonstration of optimal unambiguous state discrimination. *Physical Review A*, 63(4):040305, 2001.

[2] Simon JD Phoenix, Stephen M Barnett, and Anthony Chefles. Three-state quantum cryptography. *Journal of modern optics*, 47(2-3):507–516, 2000.

[3] Charles H Bennett. Quantum cryptography : Public key distribution and coin tossing. *International Conference on Computer System and Signal Processing, IEEE, 1984*, pages 175–179, 1984. URL http://ci.nii.ac.jp/naid/20001457561/en/.

[4] Stephen M Barnett and Sarah Croke. Quantum state discrimination. *Advances in Optics and Photonics*, 1(2):238–278, 2009.

[5] Carl W Helstrom. *Quantum detection and estimation theory.* Academic press, 1976.

[6] Erika Andersson, Stephen M Barnett, Claire R Gilson, and Kieran Hunter. Minimum-error discrimination between three mirror-symmetric states. *Physical Review A*, 65(5):052308, 2002.

[7] Sarah Croke, Erika Andersson, Stephen M Barnett, Claire R Gilson, and John Jeffers. Maximum confidence quantum measurements. *Physical review letters*, 96 (7):070401, 2006.

[8] Stephen M Barnett. Minimum-error discrimination between multiply symmetric states. *Physical Review A*, 64(3):030303, 2001.

[9] Joonwoo Bae and Won-Young Hwang. Minimum-error discrimination of qubit states: Methods, solutions, and properties. *Physical Review A*, 87(1):012334, 2013.

[10] Joonwoo Bae. Structure of minimum-error quantum state discrimination. *New Journal of Physics*, 15(7):073037, 2013.

[11] Donghoon Ha and Younghun Kwon. Complete analysis for three-qubit mixed-state discrimination. *Physical Review A*, 87(6):062302, 2013.

[12] Matthieu E Deconinck and Barbara M Terhal. Qubit state discrimination. *Physical Review A*, 81(6):062304, 2010.

[13] Igor D Ivanovic. How to differentiate between non-orthogonal states. *Physics Letters A*, 123(6):257–259, 1987.

[14] Dennis Dieks. Overlap and distinguishability of quantum states. *Physics Letters A*, 126(5):303–306, 1988.

[15] Asher Peres. How to differentiate between non-orthogonal states. *Physics Letters A*, 128(1):19, 1988.

[16] Anthony Chefles. Unambiguous discrimination between linearly independent quantum states. *Physics Letters A*, 6(239):339–347, 1998.

[17] Stephen Barnett. *Quantum information*, volume 16. Oxford University Press, 2009.

[18] William K Wootters and Wojciech H Zurek. A single quantum cannot be cloned. *Nature*, 299(5886):802–803, 1982.

[19] Dennis Dieks. Communication by epr devices. *Physics Letters A*, 92(6):271–272, 1982.

[20] Claude E Shannon. Communication theory of secrecy systems. *Bell Labs Technical Journal*, 28(4):656–715, 1949.

[21] Nicolas Gisin, Grégoire Ribordy, Wolfgang Tittel, and Hugo Zbinden. Quantum cryptography. *Reviews of modern physics*, 74(1):145, 2002.

[22] Charles H Bennett, Gilles Brassard, and Jean-Marc Robert. Privacy amplification by public discussion. *SIAM journal on Computing*, 17(2):210–229, 1988.

[23] Russell Impagliazzo, Leonid A Levin, and Michael Luby. Pseudo-random generation from one-way functions. In *Proceedings of the twenty-first annual ACM symposium on Theory of computing*, pages 12–24. ACM, 1989.

[24] Charles H Bennett, Gilles Brassard, Claude Crépeau, and Ueli M Maurer. Generalized privacy amplification. *IEEE Transactions on Information Theory*, 41(6): 1915–1923, 1995.

[25] Johan Håstad, Russell Impagliazzo, Leonid A Levin, and Michael Luby. A pseudorandom generator from any one-way function. *SIAM Journal on Computing*, 28 (4):1364–1396, 1999.

[26] Joonwoo Bae and Leong-Chuan Kwek. Quantum state discrimination and its applications. *Journal of Physics A: Mathematical and Theoretical*, 48(8):083001, 2015.

[27] Edward Davies. Information and quantum measurement. *IEEE Transactions on Information Theory*, 24(5):596–599, 1978.

[28] Masahide Sasaki, Stephen M Barnett, Richard Jozsa, Masao Osaki, and Osamu Hirota. Accessible information and optimal strategies for real symmetrical quantum sources. *Physical Review A*, 59(5):3325, 1999.

[29] H Yuen, R Kennedy, and Melvin Lax. Optimum testing of multiple hypotheses in quantum detection theory. *IEEE Transactions on Information Theory*, 21(2): 125–134, 1975.

[30] Anthony Chefles. Unambiguous discrimination between linearly dependent states with multiple copies. *Physical Review A*, 64(6):062305, 2001.

[31] Gregg Jaeger and Abner Shimony. Optimal distinction between two non-orthogonal quantum states. *Physics Letters A*, 197(2):83–87, 1995.

[32] Peter J Mosley, Sarah Croke, Ian A Walmsley, and Stephen M Barnett. Experimental realization of maximum confidence quantum state discrimination for the extraction of quantum information. *Physical review letters*, 97(19):193601, 2006.

[33] Eric Chitambar and Min-Hsiu Hsieh. Revisiting the optimal detection of quantum information. *Physical Review A*, 88(2):020302, 2013.

[34] Masashi Ban, Keiko Kurokawa, Rei Momose, and Osamu Hirota. Optimum measurements for discrimination among symmetric quantum states and parameter estimation. *International Journal of Theoretical Physics*, 36(6):1269–1288, 1997.

[35] Chih-Lung Chou. Minimum-error discrimination among mirror-symmetric mixed quantum states. *Physical Review A*, 70(6):062316, 2004.

[36] Roger BM Clarke, Vivien M Kendon, Anthony Chefles, Stephen M Barnett, Erling Riis, and Masahide Sasaki. Experimental realization of optimal detection strategies for overcomplete states. *Physical Review A*, 64(1):012303, 2001.

[37] Masoud Mohseni, Aephraim M Steinberg, and János A Bergou. Optical realization of optimal unambiguous discrimination for pure and mixed quantum states. *Physical review letters*, 93(20):200403, 2004.

[38] Bruno Huttner, Antoine Muller, Jean-Daniel Gautier, Hugo Zbinden, and Nicolas Gisin. Unambiguous quantum measurement of nonorthogonal states. *Physical Review A*, 54(5):3783, 1996.

[39] Graeme Weir, Stephen M. Barnett, and Sarah Croke. Optimal discrimination of single-qubit mixed states. *Phys. Rev. A*, 96:022312, Aug 2017. doi: 10.1103/PhysRevA.96.022312. URL https://link.aps.org/doi/10.1103/PhysRevA.96.022312.

[40] Jonathan Walgate, Anthony J Short, Lucien Hardy, and Vlatko Vedral. Local distinguishability of multipartite orthogonal quantum states. *Physical Review Letters*, 85(23):4972, 2000.

[41] Charles H Bennett, David P DiVincenzo, Christopher A Fuchs, Tal Mor, Eric Rains, Peter W Shor, John A Smolin, and William K Wootters. Quantum nonlocality without entanglement. *Physical Review A*, 59(2):1070, 1999.

[42] Sarah Croke, Stephen M. Barnett, and Graeme Weir. Optimal sequential measurements for bipartite state discrimination. *Phys. Rev. A*, 95:052308, May 2017. doi: 10.1103/PhysRevA.95.052308. URL https://link.aps.org/doi/10.1103/PhysRevA.95.052308.

[43] Eric Chitambar, Debbie Leung, Laura Mančinska, Maris Ozols, and Andreas Winter. Everything you always wanted to know about locc (but were afraid to ask). *Communications in Mathematical Physics*, 328(1):303–326, 2014.

[44] Nicolas Brunner, Miguel Navascués, and Tamás Vértesi. Dimension witnesses and quantum state discrimination. *Physical Review Letters*, 110(15):150501, 2013.

[45] Vladimir Bužek and Mark Hillery. Quantum copying: Beyond the no-cloning theorem. *Physical Review A*, 54(3):1844, 1996.

[46] Dagmar Bruss, Artur Ekert, and Chiara Macchiavello. Optimal universal quantum cloning and state estimation. *Physical Review Letters*, 81(12):2598, 1998.

[47] Dagmar Bruß, Mirko Cinchetti, G Mauro DAriano, and Chiara Macchiavello. Phase-covariant quantum cloning. *Physical Review A*, 62(1):012302, 2000.

[48] Alexander S Holevo. Statistical decision theory for quantum systems. *Journal of Multivariate Analysis*, 3(4):337–394, 1973.

[49] Stephen M Barnett and Sarah Croke. On the conditions for discrimination between quantum states with minimum error. *Journal of Physics A: Mathematical and Theoretical*, 42(6):062001, 2009.

[50] Kieran Hunter. Measurement does not always aid state discrimination. *Physical Review A*, 68(1):012306, 2003.

[51] Alexander Semenovich Holevo. On asymptotically optimal hypotheses testing in quantum statistics. *Teoriya Veroyatnostei i ee Primeneniya*, 23(2):429–432, 1978.

[52] Paul Hausladen and William K Wootters. A pretty good measurement for distinguishing quantum states. *Journal of Modern Optics*, 41(12):2385–2390, 1994.

[53] Paul Hausladen, Richard Jozsa, Benjamin Schumacher, Michael Westmoreland, and William K Wootters. Classical information capacity of a quantum channel. *Physical Review A*, 54(3):1869, 1996.

[54] Lane P Hughston, Richard Jozsa, and William K Wootters. A complete classification of quantum ensembles having a given density matrix. *Physics Letters A*, 183 (1):14–18, 1993.

[55] Carlos Mochon. Family of generalized pretty good measurements and the minimal-error pure-state discrimination problems for which they are optimal. *Physical Review A*, 73(3):032328, 2006.

[56] Yonina C Eldar, Alexandre Megretski, and George C Verghese. Optimal detection of symmetric mixed quantum states. *IEEE Transactions on Information Theory*, 50(6):1198–1207, 2004.

[57] Masahide Sasaki, Kentaro Kato, Masayuki Izutsu, and Osamu Hirota. Quantum channels showing superadditivity in classical capacity. *Physical Review A*, 58(1): 146, 1998.

[58] Kieran Hunter, Stephen M Barnett, Osamu Hirota, Patrik Öhberg, John Jeffers, and Erika Andersson. Results in optimal discrimination. In *AIP Conference Proceedings*, volume 734, pages 83–86. AIP, 2004.

[59] M Ježek, J Řeháček, and J Fiurášek. Finding optimal strategies for minimum-error quantum-state discrimination. *Physical Review A*, 65(6):060301, 2002.

[60] Jon Tyson. A remark on Deconinck and Terhal's "qubit state discrimination" and Kadison's "Order properties of bounded self-adjoint operators". *Private Communication*, 2013.

[61] William Karush. Minima of functions of several variables with inequalities as side constraints. *M. Sc. Dissertation. Dept. of Mathematics, Univ. of Chicago*, 1939.

[62] Harold W Kuhn and Albert W Tucker. Nonlinear programming. In *Traces and emergence of nonlinear programming*, pages 247–258. Springer, 2014.

[63] Donghoon Ha and Younghun Kwon. Discriminating n-qudit states using geometric structure. *Physical Review A*, 90(2):022330, 2014.

[64] Terry Rudolph, Robert W Spekkens, and Peter S Turner. Unambiguous discrimination of mixed states. *Physical Review A*, 68(1):010301, 2003.

[65] Sarah Croke, Peter J Mosley, Stephen M Barnett, and Ian A Walmsley. Maximum confidence measurements and their optical implementation. *The European Physical Journal D*, 41(3):589–598, 2007.

[66] Yonina C Eldar and G David Forney. On quantum detection and the square-root measurement. *IEEE Transactions on Information Theory*, 47(3):858–872, 2001.

[67] Stephen M Barnett and Erling Riis. Experimental demonstration of polarization discrimination at the helstrom bound. *Journal of Modern Optics*, 44(6):1061–1064, 1997.

[68] Yuqing Sun, János A Bergou, and Mark Hillery. Optimum unambiguous discrimination between subsets of nonorthogonal quantum states. *Physical Review A*, 66 (3):032315, 2002.

[69] Alexander S Holevo. *Probabilistic and statistical aspects of quantum theory.* North Holland Publishing Company, Amsterdam, 1982.

[70] Michael A Nielsen and Isaac L Chuang. *Quantum computation and quantum information.* Cambridge university press, 2010.

[71] Ulrike Herzog. Optimal measurements for the discrimination of quantum states with a fixed rate of inconclusive results. *Physical Review A*, 91(4):042338, 2015.

[72] Peter W Shor. On the number of elements needed in a povm attaining the accessible information. In *Quantum Communication, Computing, and Measurement 3*, pages 107–114. Springer, 2002.

[73] Graeme Weir, Catherine Hughes, Stephen M Barnett, and Sarah Croke. Optimal measurement strategies for the trine states with arbitrary prior probabilities. *Quantum Science and Technology*, 3(3):035003, 2018.

[74] Serge Massar and Sandu Popescu. Optimal extraction of information from finite quantum ensembles. *Physical review letters*, 74(8):1259, 1995.

[75] Charles H Bennett, Gilles Brassard, Seth Breidbart, and Stephen Wiesner. Quantum cryptography, or unforgeable subway tokens. In *Advances in Cryptology*, pages 267–275. Springer, 1983.

[76] Kieran Hunter. *Optimal generalised measurement strategies,* University of Strathclyde *Ph.D. Thesis.* 2004.

[77] Graeme Weir, Catherine Hughes, Stephen Barnett, and Sarah Croke. Optimal measurement strategies for multiple copies of the trine states. *In preparation.*

[78] Wojciech Hubert Zurek. Decoherence and the transition from quantum to classicalrevisited. In *Quantum Decoherence*, pages 1–31. Springer, 2006.

[79] Shashank Virmani, Massimiliano F Sacchi, Martin B Plenio, and Damian Markham. Optimal local discrimination of two multipartite pure states. *Physics Letters A*, 288 (2):62–68, 2001.

[80] Yi-Xin Chen and Dong Yang. Optimal conclusive discrimination of two nonorthogonal pure product multipartite states through local operations. *Physical Review A*, 64(6):064303, 2001.

[81] Yi-Xin Chen and Dong Yang. Optimally conclusive discrimination of nonorthogonal entangled states by local operations and classical communications. *Physical Review A*, 65(2):022320, 2002.

[82] Zhengfeng Ji, Hongen Cao, and Mingsheng Ying. Optimal conclusive discrimination of two states can be achieved locally. *Physical Review A*, 71(3):032323, 2005.

[83] Asher Peres and William K Wootters. Optimal detection of quantum information. *Physical Review Letters*, 66(9):1119, 1991.

[84] Yang Lu, Nick Coish, Rainer Kaltenbaek, Deny R Hamel, Sarah Croke, and Kevin J Resch. Minimum-error discrimination of entangled quantum states. *Physical Review A*, 82(4):042340, 2010.

[85] Robin Blume-Kohout, Sarah Croke, and Michael Zwolak. Quantum data gathering. *Scientific reports*, 3, 2013.

[86] Dorje Brody and Bernhard Meister. Minimum decision cost for quantum ensembles. *Physical Review Letters*, 76(1):1, 1996.

[87] Antonio Acín, Emili Bagan, Marià Baig, Ll Masanes, and Ramon Munoz-Tapia. Multiple-copy two-state discrimination with individual measurements. *Physical Review A*, 71(3):032338, 2005.

[88] Sibasish Ghosh, Guruprasad Kar, Anirban Roy, Aditi Sen, Ujjwal Sen, et al. Distinguishability of bell states. *Physical Review Letters*, 87(27):277902, 2001.

[89] Jonathan Walgate and Lucien Hardy. Nonlocality, asymmetry, and distinguishing bipartite states. *Physical Review Letters*, 89(14):147901, 2002.

[90] Anthony Chefles and Stephen M Barnett. Strategies for discriminating between non-orthogonal quantum states. *Journal of Modern Optics*, 45(6):1295–1302, 1998.

[91] Jaromír Fiurášek and Miroslav Ježek. Optimal discrimination of mixed quantum states involving inconclusive results. *Physical Review A*, 67(1):012321, 2003.