Quantum State Discrimination and Quantum Cloning: Optimization and Implementation

by

Lazy bum

A dissertation submitted to the Graduate Faculty in Physics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2015

This manuscript has been read and accepted for the
Graduate Faculty in Physics in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy.

Date	Prof. Jï£¡nos A. Bergou Chair of Examining Committee
Date	Prof. Igor L. Kuskovsky Executive Officer

Supervisory Committee:

Prof. Mark Hillery

Prof. Steven Greenbaum

Prof. Ed Feldman

Prof. Neepa T. Maitra

THE CITY UNIVERSITY OF NEW YORK

Abstract

Quantum State Discrimination and Quantum Cloning: Optimization and Implementation

by

Lazy Bum

Advisor: Janos A. Bergou

This thesis reflects works previously published by the author and materials hitherto unpublished on the subject of quantum information theory. Particularly, results in optimal discrimination, cloning, and separation of quantum states, and their relationships, are discussed. Via Neumark's theorem [**?**], our description of these unitary processes can be implemented with linear optical devices.

Acknowledgements

THANKS BRUVS, MUCH LOVE!

Contents

List of Figures

CHAPTER 1

Introduction

1.1. On Why Quantum Information Theory Is Interesting

Classical bits versus quantum bits: instead of just a 0 or 1, quantum bits can maintain a superposition state

The probabalistic nature of detection: only orthogonal states can be discriminated perfectly

Quantum Computing

Quantum Communication factorization work of Peter Schor [?] and quantum key distribution protocols such as B92 [?].

1.2. State Representation and Measurements

We begin by describing pure quantum states ψ as vectors in a Hilbert space and an ensemble of states as $\rho = \sum_i \eta_i |\psi_i\rangle\langle\psi_i|$ where $\sum \eta_i = 1$. The evolution of this ensemble is the Schrodinger equation as

$$i\hbar \frac{\partial \rho}{\partial t} = [H, \rho].$$

Solving this for evolution of initial state $\rho(t = 0)$ we get

$$\rho(t) = U(t)\rho(0)U(t)^\dagger.$$

and where the unitary matrix U obeys $UU^\dagger = I$. There are several ways to view this formula. The first is by decomposing the unitary into a set of Kraus operators A_i such that $\sum A_i A_i^\dagger = I$. This allows us to write the evolution of pure states interacting through unitary evolution as

$$U(|\psi_A\rangle \otimes |\phi_B\rangle) = \sum_i A_i |\psi_A\rangle \otimes |i_B\rangle. \tag{1.2.1}$$

This equation describes the alternative effects of the evolution between states $|\psi_A\rangle$ and $|\phi_B\rangle$ in the space $H_1 \otimes H_2$ as a Kraus operator A_i for each outcome direction $|i_B\rangle$ of the ancilliary Hilbert space, such as a measurement device. For example a particle striking one detector as opposed to another can be described by the different operators A_i.

The other is due to Neumark [?]. A unitary acts on a pure state ψ to make state ϕ, as in $U|\psi\rangle = |\phi\rangle$. Before we describe such processes particularly let us describe the mathematics of these structures. The ensemble previously described can be describes as a density matrix. This is a generalized state that is a statistical collection of pure states defined by four properties:

(1) $\rho = \sum_i \eta_i |\psi_i\rangle\langle\psi_i|$ where $\sum \eta_i = 1$,

(2) It is Hermitian,

(3) $Tr\rho = 1$,

(4) $\langle\phi_i|\rho|\phi_i\rangle \geq 0$.

The Kaus operators can be associated with measurement operators Π_i such that $\Pi_i = A_i^\dagger A_i$. Hence the Π_i are a decomposition of the identity in terms of positive semi-definite matrices. A measurement can be either a projector onto an eigenvector of the Hilbert space or a generalized measurement. In the latter case it must fulfill only 2 properties:

(1) $\sum \Pi_i = 1$

(2) They are Hermitian. This corresponds to real eigenvalues (measurement outcomes).

Since non-orthogonal states cannot be discriminated perfectly, we can speak of the probability of a given outcome: $\langle\Pi_i\rangle = \sum_j \eta_j \langle\psi_j|\Pi_i|\psi_j\rangle = Tr(\Pi_i\rho)$ This further justifies the motivation for the measurement operators.

CHAPTER 2

Discrimination of Pure States

Consider the problem of discriminating between two states $|\psi\rangle_1$ and $|\psi\rangle_2$. If we could always perform this perfectly then we should be able to write a unitary U such that

$$U|\psi_1\rangle = |1\rangle, \tag{2.0.2}$$

$$U|\psi_2\rangle = |2\rangle, \tag{2.0.3}$$

where the states $|1\rangle$ and $|2\rangle$ are orthogonal in the basis of the input states, and each result is associated with the respective input. However since the unitary is inner-product preserving, taking the product of the first with the second's adjoint shows this unitary and hence such a measurement are impossible unless the input states are orthogonal:

$$\langle\psi_2|U^\dagger U|\psi_1\rangle = \langle\psi_2|\psi_1\rangle = \langle 1|2\rangle = 0.$$

We can make a similar demonstration using the operator method. If

$$\Pi_1|\psi_2\rangle = 0, \tag{2.0.4}$$

$$\Pi_2|\psi_1\rangle = 0, \tag{2.0.5}$$

then using $\Pi_1 + \Pi_2 = I$ and inner product of these equations, we get the same result:

$$0 = \langle\psi_2|\Pi_1 + \Pi_2|\psi_1\rangle = \langle\psi_1|\psi_2\rangle$$

Since the two constraints of measurement, orthogonality of the measurement vectors and their spanning the space, proved contradictory, we must give up one of these two functions in order to perform a physical measurement. Hence we must choose a unitary that performs this task optimally according to some figure of merit, typically a probability measure. For

all future discussion we assume that the input states ψ_1 and ψ_2 are provided one at a time, and with known probabilities η_1 and η_2 respectively, such that $\eta_1 + \eta_2 = 1$.

2.1. Minimum Error Discrimination

Historically, the first solution to this problem is due to Helstrom [?]. Now known as the Minimum Error (ME) strategy, the figure of merit is the average rate of mistakingly identifying one state for the other. Using the density matrix and trace notation, this average probability of error can be written as

$$P_e = \eta_1 tr[\rho_1 \Pi_2] + \eta_2 tr[\rho_2 \Pi_1],$$

where we associate the outcome Π_i with state ψ_i. Analagously, the average probability of success is simply

$$P_s = \eta_1 tr[\rho_1 \Pi_1] + \eta_2 tr[\rho_2 \Pi_2],$$

Because we want a result every time a state is sent to us the operators Π_i must span the space, so $\Pi_1 + \Pi_2 = I$ which implies $P_e + P_s = 1$. The minimum error is attained at the Helstrom bound:

$$P_E = \frac{1}{2}[1 - \sqrt{1 - 4\eta_1\eta_2|\langle\psi_1|\psi_2\rangle|^2}]. \tag{2.1.1}$$

The detectors and states can are graphically represented in Fig. 1 for the case when the given states are qubits and occur with equal likelihood.

We provide a derivation of this result to demonstrate the Neumark formalism. Since both ψ_1 and ψ_2 can now evolve to either 0 or 1, we must write the unitary equations as

$$U|\psi_1\rangle = \sqrt{p_1}|1\rangle + \sqrt{r_1}|2\rangle \tag{2.1.2}$$

$$U|\psi_2\rangle = \sqrt{p_2}|2\rangle + \sqrt{r_2}|1\rangle, \tag{2.1.3}$$

where p_i and r_i are the individual success and error probabilities of the measurement. Taking the inner product of these two equations with themselves we find $p_i + r_i = 1$, and taking the inner product with each other we get the overlap constraint

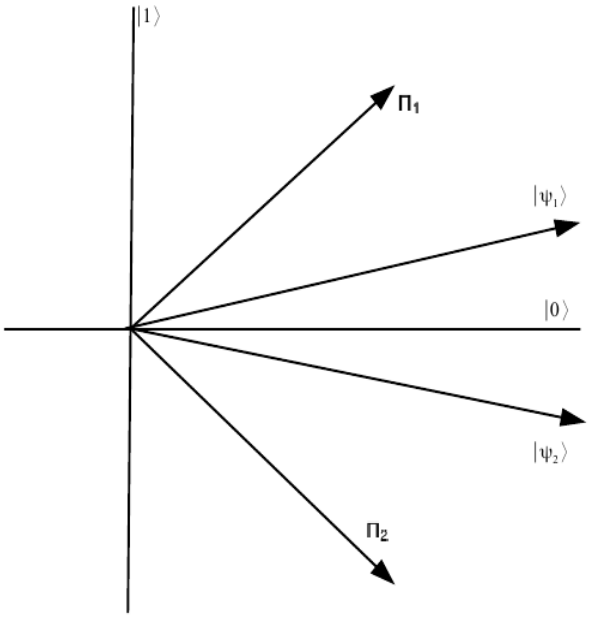

FIGURE 2.1.1. The states ψ_i and projectors $\Pi_i = |\phi_i\rangle\langle\phi_i|$ which minimizes the error rate for detection. along the states $\{|\psi_1\rangle, |\psi_2\rangle\}$ for $\eta_1 = \eta_2 = \frac{1}{2}$.

$$s = \langle\psi_1|\psi_2\rangle = \sqrt{p_1 r_2} + \sqrt{p_2 r_1}, \tag{2.1.4}$$

We wish to minimize the average error rate

$$P_E = \eta_1 r_1 + \eta_2 r_2, \tag{2.1.5}$$

subject to the constraint in (2.1.4). We solve this two variable problem using the method of Lagrange multipliers. For details on this method see Appendix 2. The constrained error equation can be written as

$$F(r_1, r_2, \lambda) = \eta_1 r_1 + \eta_2 r_2 + \lambda \left[s - \sqrt{(1-r_1) r_2} - \sqrt{(1-r_2) r_1} \right]. \tag{2.1.6}$$

Our extrema will be found when all partial derivatives of this equation are zero, therefore

$$\frac{\partial F_E}{\partial r_1} = \eta_1 + \frac{1}{2}\left[\sqrt{\frac{r_2}{1-r_1}} - \sqrt{\frac{1-r_2}{r_1}} \right] = 0,$$

and

$$\frac{\partial F_E}{\partial r_2} = \eta_2 + \frac{1}{2}\left[-\sqrt{\frac{r_1}{1-r_2}} + \sqrt{\frac{1-r_1}{r_2}}\right] = 0.$$

We notice that these two equations may be re-arranged such that the left hand side is dependent on a single variable:

$$\frac{2\eta_1}{\lambda}\sqrt{r_1(1-r_1)} = \sqrt{r_1 r_2} - \sqrt{(1-r_1)(1-r_2)}, \tag{2.1.7}$$

$$\frac{2\eta_2}{\lambda}\sqrt{r_2(1-r_2)} = \sqrt{r_1 r_2} - \sqrt{(1-r_1)(1-r_2)}. \tag{2.1.8}$$

The right hand sides of Eq.(2.1.7) and (2.1.8) can be set to a constant $\frac{2\eta_i}{\lambda}\sqrt{r_i(1-r_i)} \equiv C$, which can later be determined from the unitarity constraint 2.1.4,

$$r_i = \frac{1}{2}\left(1 \pm \sqrt{1 - \frac{\lambda^2 C^2}{\eta_i^2}}\right) = \frac{1}{2}\left(1 - \sqrt{1 - \frac{\delta^2}{\eta_i^2}}\right), \tag{2.1.9}$$

$$r_i = \frac{1}{2}\left[1 - A_i\right], \tag{2.1.10}$$

where $A_i \equiv \sqrt{1 - \frac{\delta^2}{\eta_i^2}}$ and $\delta^2 \equiv \lambda^2 C^2$. The smaller r_i is picked (lower sign in 2.1.9) as this represents error rate, which is to be minimized. Now replace r_i into the constraint (2.1.4) and solve for δ:

$$\begin{aligned}
s &= \sqrt{(1-r_1)r_2} + \sqrt{(1-r_2)r_1}, \\
2s &= \sqrt{(1+A_1)(1-A_2)} + \sqrt{(1-A_1)(1+A_2)}, \\
2s^2 &= 1 - A_1 A_2 + \sqrt{(1-A_1^2)(1-A_2^2)}, \\
2s^2 &= 1 - A_1 A_2 + \frac{\delta^2}{\eta_1 \eta_2}, \\
(2s^2 - 1 - \tfrac{\delta^2}{\eta_1\eta_2})^2 &= 1 - \frac{\delta^2}{\eta_1^2} - \frac{\delta^2}{\eta_2^2} + \frac{\delta^4}{\eta_1^2 \eta_2^2}.
\end{aligned}$$

After some tedious but trivial algebra:

$$\delta^2 = \frac{4s^2(1-s^2)\eta_1^2\eta_2^2}{1-4\eta_1\eta_2 s^2}. \tag{2.1.11}$$

Now substitute the value of δ from (2.1.11) into (2.1.9) to get the explicit form of the individual error rates,

$$r_i = \frac{1}{2}\left[1 - \frac{1-2\eta_i s^2}{\sqrt{1-4\eta_1\eta_2 s^2}}\right] \tag{2.1.12}$$

Inserting r_1 and r_2 into (**??**) Helstrom bound is retrieved [**?**]

$$\begin{aligned} P_E &= \frac{1}{2}\left[1 - \frac{\eta_1 - 2\eta_1\eta_2 s^2}{\sqrt{1-4\eta_1\eta_2 s^2}} - \frac{\eta_2 - 2\eta_1\eta_2 s^2}{\sqrt{1-4\eta_1\eta_2 s^2}}\right], \\ P_E &= \frac{1}{2}\left[1 - \sqrt{1-4\eta_1\eta_2 s^2}\right]. \end{aligned} \tag{2.1.13}$$

2.2. Unambiguous Discrimination

It was noticed by *UD CITations** that we may completely eliminate error from the measurement results by giving up the constraint that our two measurement operators span the whole space, $\Pi_1 + \Pi_2 = I$. This creates an additional result which is not associated with the state being either ψ_1 or ψ_2 . It is called the inconclusive or failure outcome Π_0, and the new decomposition of the identity reads $\Pi_0 + \Pi_1 + \Pi_2 = I$.

$$\begin{aligned} \Pi_1 &= \frac{p_1}{|\langle\psi_1|\psi_2^\perp\rangle|^2}|\psi_2^\perp\rangle\langle\psi_2^\perp|, \\ \Pi_2 &= \frac{p_2}{|\langle\psi_2|\psi_1^\perp\rangle|^2}|\psi_1^\perp\rangle\langle\psi_1^\perp|. \end{aligned} \tag{2.2.1}$$

To determine the failure operator, insert (2.2.1) into (**??**):

$$\Pi_0 = I - \Pi_1 - \Pi_2 = I - \frac{p_1}{|\langle\psi_1|\psi_2^\perp\rangle|^2}|\psi_2^\perp\rangle\langle\psi_2^\perp| - \frac{p_2}{|\langle\psi_2|\psi_1^\perp\rangle|^2}|\psi_1^\perp\rangle\langle\psi_1^\perp|. \tag{2.2.2}$$

The eigenvalues of Π_0 must be non-negative, giving us the inequality constraint between the individual failure rates as

$$q_1 q_2 \geq |\langle\psi_1|\psi_2\rangle|^2. \tag{2.2.3}$$

where we used $q_i = 1 - p_i$.

Now our value of merit will be the average failure rate $Q = \eta_1 q_1 + \eta_2 q_2$. Since there is no error the success and failure add to one: $P_s + Q = 1$. Hence we wish to minimize the failure rate by taking the equality in Eq. (2.2.3), giving us the minimum failure rate at

$$Q \equiv Q_0 = 2\sqrt{\eta_1 \eta_2}|\langle\psi_1|\psi_2\rangle|. \tag{2.2.4}$$

This solution is valid for $q_i \leq 1$. Outside of this bound we ignore the state with the high rate of failure by removing that detector and reducing our measurement strategy back to projective measurements. We project onto the more likely state. Therefore the total UD solution is

$$Q_{\rm c} = \begin{cases} \eta_1 + \eta_2 \cos^2\theta, \text{if } \eta_1 < \dfrac{\cos^2\theta}{1+\cos^2\theta} \equiv \eta_1^{(l)}, \\ \eta_2 + \eta_1 \cos^2\theta, \text{if } \eta_1 > \dfrac{1}{1+\cos^2\theta} \equiv \eta_1^{(r)}, \\ 2\sqrt{\eta_1\eta_2}\cos\theta \equiv Q_0, \text{if } \eta_1^{(l)} \leq \eta_1 \leq \eta_1^{(r)}\ , \end{cases} \tag{2.2.5}$$

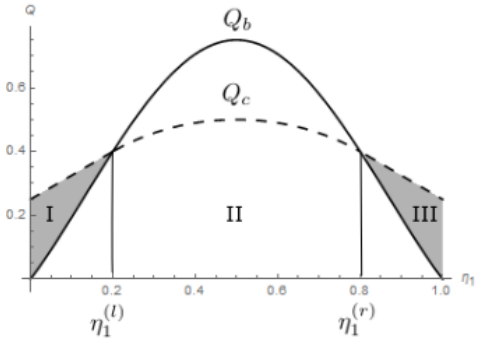

FIGURE 2.2.1. $Q_{\rm c}$ (dashed line, Eq. (2.2.5)) and $Q_{\rm b}$ (solid line, Eq. (**??**)) vs. η_1 for $\theta = \pi/3$. Measurements can be optimized in the area under the dashed line, Q_c. Measurements in the area above Q_c are suboptimal. In the shaded areas between Q_c and Q_b (regions I, left, and III, right) the optimal FRIO measurement is a projective measurement, in the unshaded area below Q_c (region II) the optimal measurement is a POVM.

2.3. Interpolative Discrimination

There are many reasons to consider the problem of allowing both failure and error results in your measurement. First, experimental implementations are never without error. Second, success rates may significantly increase for small changes in the error rate. This motivates an intermediate strategy

2.3.1. Operator Transformation. We review the original problem in a single two-dimensional Hilbert space, and the solution that involves a transformation that eliminates the failure operator from the discrimination problem. We are given two pure states $\rho_1 = |\psi\rangle_1\langle\psi_1|$ and $\rho_2 = |\psi\rangle_2\langle\psi_2|$ with a-priori probabilities η_1 and η_2 respectively. These two probabilities add to one: $\eta_1 + \eta_2 = 1$. We wish to optimize the success rate $P_s = \eta_1 tr[\Pi_1\rho_1] + \eta_2 tr[\Pi_2\rho_2]$ for a fixed failure rate $Q = tr[\Pi_0(\eta_1\rho_1 + \eta_2\rho_2)]$ where the measurement operators are the Π_i that span the Hilbert space: $\Pi_1 + \Pi_2 + \Pi_0 = I$. The transformation we implement is

$$\Omega^{-1/2}[\Pi_1 + \Pi_2]\Omega^{-1/2} = I, \tag{2.3.1}$$

where $\Omega = I - \Pi_0$.

Calling the transformed operators $\widetilde{\Pi}_i$ for $i = 1, 2$, we can find the corresponding transformed density matrices $\widetilde{\rho}_i$ and a-priori probabilities $\widetilde{\eta}_i$ to make this a new ME problem that can be readily solved. The minimized error probability as a function of the failure rate is

$$P_e = \frac{1}{2}(1 - Q - \sqrt{(1-Q)^2 - (Q - Q_0)^2}) \tag{2.3.2}$$

for $Q \leq Q_0 = 2\sqrt{\eta_1\eta_2}\cos\theta$ where Q_0 is the maximum failure rate allowed in the optimization scheme and it corresponds to the best measurement in the UD case. We will provide an in-depth description of this solution as it pertains to our problem of discriminating mixed states with a Jordan basis structure in the next chapter.

2.3.2. Neumark Solution of Interpolation. Since our measurement results must now include the three degrees of freedom from the successful measurements and failure result, in addition to the two degrees of freedom provided us by the pure states, we require

an extra ancillary degree of freedom. This we add by extending our Hilbert space with a direct sum extension.

$$U|\psi_1\rangle = \sqrt{p_1}|1\rangle + \sqrt{r_1}|2\rangle + \sqrt{q_1}|0\rangle \tag{2.3.3}$$

$$U|\psi_2\rangle = \sqrt{r_2}|1\rangle + \sqrt{p_2}|2\rangle + \sqrt{q_2}|0\rangle \tag{2.3.4}$$

Here p_i is the probability that the state i is correctly identified when it is sent into the measurment apparatus, r_i the error rate (mistaking one state for the other) and q_i the failure rate, or not getting a conclusive measurement result. By sandwiching the preceding equations with their adjoints we confirm that $q_i + r_i + p_i = 1$, the sum of various probabilities is one.

2.3.2.1. *Equal Weights Solution.* The solution when $\eta_1 = \eta_2 = \frac{1}{2}$ is very simple and neat. Here we have $p_1 = p_2$, $q_1 = q_2$ and $r_1 = r_2$. Hence $Q = \eta_1 q_1 + \eta_2 q_2 = q_1$

The overlap between the two states $\langle\psi_2|\psi_1\rangle = s$ can be also found by finding the adjoint of one of the Unitary expressions and sandwiching it with the other expression, and since $UU^\dagger = 1$ we find

$$s = 2\sqrt{pr} + Q = s\sqrt{p(1-p-q)} + Q$$

Solving for p by squaring we find

$$(s-Q)^2 = 4p(1-p-q)$$

so

$$p^2 - p(1-Q) + (s-Q)^2/4 = 0$$

and the positive root gives us the answer as

$$p = \frac{1}{2}[1 - Q - \sqrt{(1-Q)^2 - (s-Q)^2}]$$

2.3.3. Full Solution. The total probability of success, error and failure we denote as $P_s = \eta_1 p_1 + \eta_2 p_2$ and $P_r = \eta_1 r_1 + \eta_2 r_2$ and $Q = \eta_1 q_1 + \eta_2 q_2$ respectively.

Taking the inner product

$$s = \langle \psi_1 | \psi_2 \rangle \ = \sqrt{p_1 r_2} + \sqrt{p_2 r_1} + \sqrt{q_1 q_2}$$

$$= \sqrt{(1 - r_1 - q_1) r_2} + \sqrt{(1 - r_2 - q_2) r_1} + \sqrt{q_1 q_2}$$

This is our overlap constraint.

We now use the Lagrange multiplier method to minimize P_e subject to the overlap constraint.

$$F_e = \eta_1 r_1 + \eta_2 r_2 + \lambda(s - OverlapConstraint)$$

Step one:

$$dF/dr_i = 0$$

Solve for $r_i(\lambda)$

Step two: substitute $s = \sqrt{p_1 r_2} + \sqrt{p_2 r_1} + \sqrt{q_1 q_2}$ into $r_i(\lambda)$ and solve for λ, then plug λ into $r_i(\lambda)$. The details of this calculation are included in Appnedix A.

A lot of algebra gives us the individual error rates as:

$$r_1 = \frac{1}{2}[\alpha_1 - \frac{[2\eta_2\omega - \alpha_1(1-Q)]}{\sqrt{(1-Q)^2 - 4\omega\eta_1\eta_2}}]$$

and

$$r_2 = \frac{1}{2}[\alpha_2 - \frac{[2\eta_1\omega - \alpha_2(1-Q)]}{\sqrt{(1-Q)^2 - 4\omega\eta_1\eta_2}}]$$

where $\alpha_i \equiv 1 - q_i$ and $\omega \equiv (s - \sqrt{q_1 q_2})^2$

These are not the final result because they were optimized individually. Let us now put these into the expression of minimum error and do a final optimization. First

$$P_E = \eta_1 r_1 + \eta_2 r_2$$

$$= \frac{1}{2}[(1-Q) - \sqrt{(1-Q)^2 - 4\eta_1\eta_2(s - \sqrt{q_1 q_2})^2}]$$

We can further minimize this expression since total Q is fixed but we can choose how much of it to place on q_1 and q_2

To minimize the above expression we can maximize the term under the square root

$$(1-Q)^2 - 4\eta_1\eta_2(s - \sqrt{q_1 q_2})^2$$
$$= (1-Q)^2 - 4\eta_1\eta_2(s - \sqrt{\frac{q_1(Q - \eta_1 q_1)}{\eta_2}})^2$$

Since the only part of this expression that is dependent on the variables q_i is the inner square root, we can find an extremum of this part. A simple partial derivative with respect to q_1 tells us the optimality condition is $Q = 2\eta_1 q_1$ or equivalently $\eta_1 q_1 = \eta_2 q_2$

This simplifies our solution to be

$$P_E = \frac{1}{2}[(1-Q) - \sqrt{(1-Q)^2 - (Q - Q_0)^2}]$$

CHAPTER 3

Discrimination of Mixed States

Here we consider the problem of discriminating between two mixed states

$$\rho_1 = \sum_i r_i |r_i\rangle\langle r_i|, \tag{3.0.5}$$

$$\rho_2 = \sum_i s_i |s_i\rangle\langle s_i|, \tag{3.0.6}$$

3.1. Minimum Error Discrimination

We introduce the subject of discriminating mixed states

3.2. Mixed States with Jordan Structure

3.2.1. Jordan Basis Structure. In this section we discuss the interpolative discrimination of a class of mixed states with a Jordan structure. As before, they are to be discriminated with a-priori probabilities η_1, η_2 but now their density matricies lie pairwise such that there are only two vectors per subspace:

$$\langle r_i|s_j\rangle = \delta_{ij}\cos\theta_i \tag{3.2.1}$$

In each 2d subspace i there lie an $|r_i\rangle$ and $|s_i\rangle$ with a-priori probabilities now $\eta_{1,i} = \eta_1 r_i$ and $\eta_{2,i} = \eta_2 s_i$. This structure can be physically interpreted as the transmission of two input states over several fiber optic cables. Each cable contains two degrees of freedom, that could be horizontal and vertical polarization.

The generalization to multiple subspaces is straightforward at first so we begin with the two subspaces example.

Here our density matrices are in four dimensions that can be described as two tensor product spaces:

$$\begin{aligned} \rho_1 &= r_1|r_1\rangle\langle r_1| + r_2|r_2\rangle\langle r_2| \\ \rho_2 &= s_1|s_1\rangle\langle s_1| + s_2|s_2\rangle\langle s_2|. \end{aligned} \tag{3.2.2}$$

These mixed states are shown in Fig. 1.

Defining our measurement operators for the first subspace as

$$\widetilde{\Pi}_{1,1} + \widetilde{\Pi}_{2,1} = I_1, \tag{3.2.3}$$

where I_1 is the identity matrix of the first subspace. We define the failure rate for the first subspace as $Q_1 = \xi_1|0\rangle_{11}\langle 0|$,which in terms of a measurement probability is also

$$Q_1 = \xi_1[\eta_{1,1}\cos^2\phi_1 + \eta_{2,1}\cos^2(\theta_1 - \phi_1)], \tag{3.2.4}$$

where θ_1 is the overlap angle between the two states in subspace 1, and ϕ_1 is the angle $|r_1\rangle$ makes with respect to $|0\rangle_1$. The error rate in that subspace is

$$P_{e,1} = \eta_{1,1}\langle r_1|\Pi_2|r_1\rangle + \eta_{2,1}\langle s_1|\Pi_1|s_1\rangle. \tag{3.2.5}$$

For our simplification trick to work we introduce the normalized state vector

$$|\widetilde{r_1}\rangle = \frac{\Omega^{1/2}|r_1\rangle}{\sqrt{\langle r_1|\Omega|r_1\rangle}} \tag{3.2.6}$$

and normalized coefficients

$$\widetilde{\eta_{1,1}} = \frac{\eta_{1,1}\langle r_1|\Omega|r_1\rangle}{\eta_{1,1}\langle r_1|\Omega|r_1\rangle + \eta_{2,1}\langle s_1|\Omega|s_1\rangle} \tag{3.2.7}$$

to get

$$P_{e,1} =$$

$$[\eta_{1.1}\langle r_1|\Omega|r_1\rangle + \eta_{2,1}\langle s_1|\Omega|s_1\rangle](\widetilde{\eta_{1,1}}\langle\widetilde{r_1}|\widetilde{\Pi_2}|\widetilde{r_1}\rangle + \widetilde{\eta_{2,1}}\langle\widetilde{s_1}|\widetilde{\Pi_1}|\widetilde{s_1}\rangle).$$

We notice that the expression in the ( ) with all tildes contains a pure state minimum error problem, while with the notation $\eta_{1,1}+\eta_{2,1}=\omega_1$ the left hand set of [ ] can be reworked into $\omega_1 - Q_1$ to rewrite the error rate as

$$P_{e,1} = \frac{1}{2}[\omega_1 - Q_1](1 - \sqrt{1 - 4\widetilde{\eta_{1,1}}\widetilde{\eta_{2,1}}|\langle\widetilde{r_1}|\widetilde{s_1}\rangle|^2}).$$

If we substitute and simplify we find this equals to

$$= \frac{1}{2}(\omega_1 - Q_1 - \sqrt{(\omega_1 - Q_1)^2 - (Q_{0,1} - Q_1 \sin 2\phi)^2}),$$

where we used the notation $Q_{0,1} = 2\sqrt{\eta_{1,1}\eta_{2,1}}\cos\theta_1$ and $\sin\phi = \frac{\sqrt{\eta_{2,1}}\cos(\theta_1-\phi_1)}{\sqrt{\eta_{1,1}\cos^2(\phi_1)+\eta_{2,1}\cos^2(\theta_1-\phi_1)}}$. Minimization of the error rate as a function of ϕ_1 tells us to set $\phi_1 = \frac{\pi}{4}$ so finally

$$P_{e,1} = \frac{1}{2}(\omega_1 - Q_1 - \sqrt{(\omega_1 - Q_1)^2 - (Q_{0,1} - Q_1)^2}). \tag{3.2.8}$$

This result agrees with the single subspace limit and is simply the optimized solution for that subspace alone. We can derive a similar result for the other subspace, so we can consider an optimal distribution of failure among the two subspaces. However, we want to treat this distribution problem for n subspaces so we first generalize our preceding solution to 2n dimensions.

Recognizing that the likelihood of finding a particle in a subspace isn't 1, we want to normalize our problem so that we can solve it like the 2d case where we had $P_e + P_s + Q = 1$. Instead, in our problem we have $P_{e,i} + P_{s,i} + Q_i = \eta_{1,i} + \eta_{2,i} = \omega_i$ where $P_{s,i}$ and Q_i are the success and the failure probabilities in that subspace. Since our measurements span the Hilbert space of this subspace, the total probability of a particle being measured therein we call ω_i, or the weight of that subspace.

We can define weighted result probabilities

$$\bar{P_{e,i}} + \bar{P_{s,i}} + \bar{Q_i} = 1 \tag{3.2.9}$$

with $\bar{\bullet} = \frac{\bullet}{\omega_i}$. We can define new constants $\bar{\eta_{1,i}}$ and $\bar{\eta_{2,i}}$ that still sum 1, so that the states and measurements in (3.8) don't change. Now it is straightforward to apply the 2d solution

to each subspace, where our error rate becomes

$$\bar{P_{e,i}} = \frac{1}{2}(1 - \bar{Q}_i - \sqrt{(1 - \bar{Q}_i)^2 - (\bar{Q_{0,i}} - \bar{Q}_i)^2}),$$

where $\bar{Q_{0,i}} = 2\sqrt{\bar{\eta_{1,i}}\bar{\eta_{2,i}}}\cos\theta_i$. If we remove the bars, this becomes the generaized version of the solution we derived for one subspace in (3.7):

$$P_{e,i} = \frac{1}{2}(\omega_i - Q_i - \sqrt{(\omega_i - Q_i)^2 - (Q_{0,i} - Q_i)^2}). \tag{3.2.10}$$

3.2.2. Subspaces formalism. To discuss the more general case of higher-dimensional input states we consider a Jordan Basis structure. Two states are now to be discriminated with a-priori probabilities η_1, η_2 such that $\rho_1 = \sum r_i|r_i\rangle\langle r_i|$ and $\rho_2 = \sum s_i|s_i\rangle\langle s_i|$ with $\langle r_i|s_j\rangle = \delta_{ij}\cos\theta_i$. In each 2d subspace i there lie an $|r_i\rangle$ and $|s_i\rangle$ with a-priori probabilities now $\eta_{1,i} = \eta_1 r_i$ and $\eta_{2,i} = \eta_2 s_i$. This structure can be physically interpreted as the transmission of two input states over several fiber optic cables. Each cable contains two degrees of freedom, that could be horizontal and vertical polarization.

The generalization to multiple subspaces is straightforward at first so we begin with the two subspaces example.

3.2.2.1. *Two Subspaces.* Here our density matrices are in four dimensions that can be described as two tensor product spaces:

$$\rho_1 = r_1|r_1\rangle\langle r_1| + r_2|r_2\rangle\langle r_2| \tag{3.2.11}$$

$$\rho_2 = s_1|s_1\rangle\langle s_1| + s_2|s_2\rangle\langle s_2|. \tag{3.2.12}$$

Defining our measurement operators for the first subspace as

$$\widetilde{\Pi}_{1,1} + \widetilde{\Pi}_{2,1} = I_1, \tag{3.2.13}$$

where I_1 is the identity matrix of the first subspace. We define the failure rate for the first subspace as $Q_1 = \xi_1|0\rangle_{11}\langle 0|$,which in terms of a measurement probability is also

$$Q_1 = \xi_1[\eta_{1,1}\cos^2\phi_1 + \eta_{2,1}\cos^2(\theta_1 - \phi_1)], \tag{3.2.14}$$

where θ_1 is the overlap angle between the two states in subspace 1, and ϕ_1 is the angle $|r_1\rangle$ makes with respect to $|0\rangle_1$. The error rate in that subspace is

$$P_{e,1} = \eta_{1,1}\langle r_1|\Pi_2|r_1\rangle + \eta_{2,1}\langle s_1|\Pi_1|s_1\rangle. \tag{3.2.15}$$

For our simplification trick to work we introduce the normalized state vector

$$|\widetilde{r_1}\rangle = \frac{\Omega^{1/2}|r_1\rangle}{\sqrt{\langle r_1|\Omega|r_1\rangle}} \tag{3.2.16}$$

and normalized coefficients

$$\widetilde{\eta_{1,1}} = \frac{\eta_{1,1}\langle r_1|\Omega|r_1\rangle}{\eta_{1,1}\langle r_1|\Omega|r_1\rangle + \eta_{2,1}\langle s_1|\Omega|s_1\rangle} \tag{3.2.17}$$

to get

$$P_{e,1} =$$

$$[\eta_{1.1}\langle r_1|\Omega|r_1\rangle + \eta_{2,1}\langle s_1|\Omega|s_1\rangle](\widetilde{\eta_{1,1}}\langle\widetilde{r_1}|\widetilde{\Pi_2}|\widetilde{r_1}\rangle + \widetilde{\eta_{2,1}}\langle\widetilde{s_1}|\widetilde{\Pi_1}|\widetilde{s_1}\rangle).$$

We notice that the expression in the ( ) with all tildes contains a pure state minimum error problem, while with the notation $\eta_{1,1}+\eta_{2,1} = \omega_1$ the left hand set of [ ] can be reworked into $\omega_1 - Q_1$ to rewrite the error rate as

$$P_{e,1} = \frac{1}{2}[\omega_1 - Q_1](1 - \sqrt{1 - 4\widetilde{\eta_{1,1}}\widetilde{\eta_{2,1}}|\langle\widetilde{r_1}|\widetilde{s_1}\rangle|^2}).$$

If we substitute and simplify we find this equals to

$$= \frac{1}{2}(\omega_1 - Q_1 - \sqrt{(\omega_1 - Q_1)^2 - (Q_{0,1} - Q_1\sin 2\phi)^2}),$$

where we used the notation $Q_{0,1} = 2\sqrt{\eta_{1,1}\eta_{2,1}}\cos\theta_1$ and $\sin\phi = \frac{\sqrt{\eta_{2,1}}\cos(\theta_1-\phi_1)}{\sqrt{\eta_{1,1}\cos^2(\phi_1)+\eta_{2,1}\cos^2(\theta_1-\phi_1)}}$. Minimization of the error rate as a function of ϕ_1 tells us to set $\phi_1 = \frac{\pi}{4}$ so finally

$$P_{e,1} = \frac{1}{2}(\omega_1 - Q_1 - \sqrt{(\omega_1 - Q_1)^2 - (Q_{0,1} - Q_1)^2}). \tag{3.2.18}$$

This result agrees with the single subspace limit and is simply the optimized solution for that subspace alone. We can derive a similar result for the other subspace, so we can consider an optimal distribution of failure among the two subspaces. However, we want to treat this distribution problem for n subspaces so we first generalize our preceding solution to 2n dimensions.

3.2.2.2. *Generalization to n subspaces.* Recognizing that the likelihood of finding a particle in a subspace isn't 1, we want to normalize our problem so that we can solve it like the 2d case where we had $P_e + P_s + Q = 1$. Instead, in our problem we have $P_{e,i} + P_{s,i} + Q_i = \eta_{1,i}+\eta_{2,i} = \omega_i$ where $P_{s,i}$ and Q_i are the success and the failure probabilities in that subspace. Since our measurements span the Hilbert space of this subspace, the total probability of a particle being measured therein we call ω_i, or the weight of that subspace.

We can define weighted result probabilities

$$\bar{P_{e,i}} + \bar{P_{s,i}} + \bar{Q_i} = 1 \tag{3.2.19}$$

with $\bar{\bullet} = \frac{\bullet}{\omega_i}$. We can define new constants $\bar{\eta_{1,i}}$ and $\bar{\eta_{2,i}}$ that still sum 1, so that the states and measurements in (3.8) don't change. Now it is straightforward to apply the 2d solution to each subspace, where our error rate becomes

$$\bar{P_{e,i}} = \frac{1}{2}(1 - \bar{Q_i} - \sqrt{(1 - \bar{Q_i})^2 - (\bar{Q_{0,i}} - \bar{Q_i})^2}),$$

where $\bar{Q_{0,i}} = 2\sqrt{\bar{\eta_{1,i}}\bar{\eta_{2,i}}}\cos\theta_i$. If we remove the bars, this becomes the generaized version of the solution we derived for one subspace in (3.7):

$$P_{e,i} = \frac{1}{2}(\omega_i - Q_i - \sqrt{(\omega_i - Q_i)^2 - (Q_{0,i} - Q_i)^2}). \tag{3.2.20}$$

3.2.3. Lagrangian Optimization. Since each subspace failure rate can vary independently we are interested in the optimal values for Q_i as a function of fixed Q. If we consider this a Lagrange Multiplier problem of $P_{e,i}$ and constraint $\sum Q_i = Q$ then we get the constrained function

$$F = P_{e,i} - \lambda(\sum Q_i - Q). \tag{3.2.21}$$

We find the minimum of this equation as a function of Q_i, substitute into the constraint equation and solve for λ to find the optimized value of the individual failure rate as

$$Q_i = \frac{Q_{0,i} - \omega_i Q_0 + Q(\omega_i - Q_{0,i})}{1 - Q_0} \tag{3.2.22}$$

Now the optimized subspace error rate is

$$P_{e,i} = \frac{1}{2}(\omega_i - Q_i - (\omega_i - Q_{0,i})\sqrt{\frac{1 + Q_0 - 2Q}{1 - Q_0}}), \tag{3.2.23}$$

with the total optimal error rate $P_e = \sum P_{e,i}$ is

$$P_e = \frac{1}{2}(1 - Q - \sqrt{(1-Q)^2 - (Q - Q_0)^2}). \tag{3.2.24}$$

We remind ourselves that while this appears identical to the 2d solution (2.2), it in fact contains parameters that are summed over all subspaces. This means that there is an onto relationship between N dimensional and 2d solutions that allows us to construct a variety of subspace strategies that replicate any 2d solution.

3.2.4. Threshold Structure. The range of the failure rate solution for subspaces previously derived is valid strictly for more than one subspace and while the upper bound at the UD limit ($Q = Q_0$) is always valid for these equations, the lower bound at the ME solution (Q=0) is not. This limit is restricted by the positivity of Q_i: as we decrease the overall failure rate Q in equation (4.2) we notice that negative solutions are attainable. Since these are not physical we must prevent Q_i from dropping below 0. To find the total failure rate at which a subspace's failure rate vanishes we set $Q_i = 0$ in (4.2) to find the critical value of

Q for that subspace to be

$$Q = Q_c^i = \frac{\omega_i Q_0 - Q_{0,i}}{\omega_i - Q_{0,i}}. \tag{3.2.25}$$

When Q falls below Q_c^j we fix $Q_j = 0$ and discard that subspace from our optimization. We realize that after this first threshold we must re-do the optimization with the remaining subspaces.

It is worthwile to consider also the positivity of the Q_c^i, which would make it a real candidate for elimination. Since $\omega_i - Q_{0,i} \geq 0$ we analyze the positivity of $\omega_i Q_0 - Q_{0,i}$. For this to be positive we need $Q_0 \geq \bar{Q_{0,i}}$ which means that the UD failure rate of that normalized subspace should be smaller than the total UD failure rate of the system of subspaces.

3.2.4.1. *First iteration.* After one subspace failure rate is set to zero, the set of subspaces contributing to the optimization decreases causing changes in the formulas. To elucidate suppose we order the subspaces such that the highest has the largest Q_c^i, and have discarded the Nth subspace associated with Q_N and ω_N . This ordering is immutable as will be proven in the subsequent subsection. An analogous optimization over remaining subspaces gives us the failure rates as

$$Q_i^{(1)} = \frac{Q_{0,i}\Lambda_{N-1} - \omega_i F_{N-1} + Q(\omega_i - Q_{0,i})}{\Lambda_{N-1} - F_{N-1}} \tag{3.2.26}$$

between $Q_c^N \geq Q \geq Q_c^{(1)N-1}$ where we've introduced the notation $\Lambda_k = \sum_1^k \omega_i$ and $F_k = \sum_1^k Q_{0,i}$, and the '1' in parenthesis in $Q_i^{(1)}$ indicates the number of subspaces removed from the Lagrangian optimization.

3.2.4.2. *General iteration.* We can iterate this process to find the nth order failure rates as

$$Q_i^{(n)} = \frac{Q_{0,i}\Lambda_{N-n} - \omega_i F_{N-n} + Q(\omega_i - Q_{0,i})}{\Lambda_{N-n} - F_{N-n}}. \tag{3.2.27}$$

For every iteration we can also find the nodes of the failure equations, which appear as

$$Q_c^{(n)i} = \frac{\omega_i F_{N-n} - Q_{0,i}\Lambda_{N-n}}{\omega_i - Q_{0,i}}. \tag{3.2.28}$$

This is similar to the first set of critical points found in (5.1). For this to be positive (and to be a candidate for elimination) we need $\frac{F_{N-n}}{\Lambda_{N-n}} \geq \frac{Q_{0,i}}{\omega_i}$ which states that the relative UD failure rate for that subspace be smaller than average to be considered for elimination.

We can derive the ordering for subspaces mentioned earlier from comparing the critical values of two subspaces for a general iteration, and simplify the condition $Q_c^{(n)i} > Q_c^{(n)j}$ to just $\bar{Q_{0,i}} < \bar{Q_{0,j}}$. Since the second inequality is iteration-independent we can conclude that the subspace with the lowest value of the normalized UD failure rate $\bar{Q_{0,i}}$ will be eliminated first, etc.

3.2.4.3. *Continuity and intersection.* It is worthwile to demonstrate the continuity of our solutions for the Q_i's. To do this we need to show that the optimal solutions match at the boundaries where a subspace is discarded, or

$$Q_i^{(n)}(Q = Q_c^{(n)N-n}) = Q_i^{(n+1)}(Q = Q_c^{(n)N-n}), \tag{3.2.29}$$

where we have chosen to consider the nth iteration of the solution and now have decided to discard the $N-n$th subspace. After we substitute for the expressions for critical points and failure rates, we multiply through by the denominators and group and eliminate like terms we get our desired result. Continuity allows a physical implementation with variable parameters to smoothly transition from one discrimination regime to the next.

Also interesting is the question of whether the Q_i ever intersect. We consider this problem in the scope of two subspaces. If $Q_{0,1} > Q_{0,2}$ and $\frac{dQ_1}{dQ} < \frac{dQ_2}{dQ}$ then the two lines will not cross. The second condition can be restated in terms of the weights of the subspaces as $\omega_1 < \frac{1+Q_{0,1}-Q_{0,2}}{2}$ or $\omega_2 > \frac{1+Q_{0,2}-Q_{0,1}}{2}$. We notice that by our first assumption, the right hand side of the first equation is greater than a half, and smaller than a half in the second equation. These are sufficient but not necessary conditions. We can also derive the condition for crossing by noting that if $Q_{0,1} > Q_{0,2}$ and $Q_c^1 > Q_c^2$ the lines will intersect. The second condition can be rewritten as $\bar{Q_{0,1}} < \bar{Q_{0,2}}$, or in terms of the weights as $\omega_1 > \frac{Q_{0,1}}{Q_0}$.

3.2.5. Single-State Domain. Each subspace failure rate also has a ceiling. For the majority of initial conditions the UD failure rate $Q_{0,i}$ sets this upper bound. For the other

cases, we find it from the constraint that $\Pi_{0,i} \leq |0\rangle_{ii}\langle 0|$. The equality limit is a full projector which eliminates another measurement and moves us from the POVM to the single-state domain (SSD).

For the single subspace case the equation for the critical ceiling is

$$Q = Q_c = \frac{2\eta_1\eta_2 sin^2\theta}{1 - Q_0}. \tag{3.2.30}$$

This result is derived from the constraint that $\xi \leq 1$ where $\Pi_0 = \xi|0\rangle\langle 0|$. Evaluating ξ for the optimal solution gives us $\xi \leq \frac{1-Q_0}{sin^2\theta}\frac{Q_0}{2\eta_1\eta_2}$ where we take the equality limit and set $\xi = 1$ to find the region in which the POVM strategy outperforms the projector measurement.

There are two regions that this occurs. Assuming $\eta_1 \geq \eta_2$, the SSD overlaps with the interpolation measurement in the region $\frac{1}{1+\cos^2\theta} \leq \eta_1$ and when $Q \geq Q_c$. For $\eta_2 \geq \eta_1$ this happens when $\frac{\cos^2\theta}{1+\cos^2\theta} \geq \eta_1$ and $Q \geq Q_c$. Because the failure operator points directly onto the less likely state in either of these cases, we find the failure rates to be simply $Q^< = \eta_2 + \eta_1\cos^2\theta$ and $Q^> = \eta_1 + \eta_2\cos^2\theta$ respectively.

To generalize to subspaces we return to the bar normalization that returned the subspace probabilities to 1. Remembering that $Q_i = \xi_i\langle 0_i|D_i|0_i\rangle$ where D_i is the full density matrix of the states in the ith subspace, $D_i = \eta_{1,i}\rho_{1,i} + \eta_{2,i}\rho_{2,i}$ we can conclude that $\bar{Q}_i = \xi_i\langle 0_i|\bar{D}_i|0_i\rangle$ where $\bar{D}_i = \bar{\eta_{1,i}}\rho_{1,i} + \bar{\eta_{2,i}}\rho_{2,i}$ Now we have restored the summation of the a-priori probabilities for each subspace to 1 while leaving ξ_i unchanged, so the preceding arguments for the single subspace can be implemented to rewrite the inequality for ξ_i as

$$\xi_i \leq \frac{\omega_i - 2\sqrt{\eta_{1,i}\eta_{2,i}}\cos\theta_i}{1 - \cos^2\theta_i}\frac{\cos\theta_i}{\sqrt{\eta_{1,i}\eta_{2,i}}}.$$

We get the natural generalization of the critical ceiling to subspaces to be

$$Q_i = Q_i^{cc} = \frac{2\eta_{1,i}\eta_{2,i}sin^2\theta_i}{\omega_i - Q_{0,i}} \tag{3.2.31}$$

As Q_i is increased past this point we have $\Pi_{1,i} = |1\rangle_{ii}\langle 1|$ and $\Pi_{0,i} = |0\rangle_{ii}\langle 0|$. Now the condition for the overlap of the SSD onto the POVM region, assuming $\eta_{1,i} \geq \eta_{2,i}$ is

$$\frac{\omega_i}{1 + \cos^2\theta_i} \leq \eta_{1,i}, \tag{3.2.32}$$

with the maximum failure rate that can be generalized as: $Q_i^{<} = \eta_{2,i} + \eta_{1,i}\cos^2\theta_i$. Similarly for $\eta_{2,i} \geq \eta_{1,i}$ we get the condition

$$\frac{\omega_i \cos^2\theta_i}{1 + \cos^2\theta_i} \geq \eta_{1,i} \tag{3.2.33}$$

and the maximum failure rate as $Q_i^{>} = \eta_{2,i} + \eta_{1,i}\cos^2\theta_i$.

We notice that with more subspaces, the condition for the overlap region of SSD over the POVM does not change for individual subspaces as the bar transformation would show us. We show this structure Fig. 2, where the shaded regions represents the SSD domains.

3.2.6. Example. It is worthwhile to show a numerical example of this method in detail. We consider three subspaces with $\eta_1 = \eta_2$ and these parameters:

Subspace (i)	r_i	s_i	θ_i	ω_i	$Q_{0,i}$	$Q_{c,i}$
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{\pi}{4}$	$\frac{1}{4}$	$\frac{\sqrt{2}}{8}$	-.39
2	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{\pi}{6}$	$\frac{1}{4}$	$\frac{\sqrt{2}}{8}$	.48
3	$\frac{5}{8}$	$\frac{3}{8}$	$\frac{\pi}{6}$	$\frac{1}{2}$	$\frac{\sqrt{2}}{8}$	-.48

Subspace 1 has its maximum failure rate as $Q_{0,1} = \sqrt{2}/8 \approx .17$. For subspace 2, the maximum failure rate isn't $Q_{0,2} = \sqrt{6}/16 \approx .15$ because it fails one of the SSD conditions and instead $Q_2^{<} = 7/32 \approx .21$, and for subspace 3 the maximum failure rate isn't $Q_{0,3} = 3\sqrt{5}/16 \approx .42$, because it fails the other SSD condition and instead $Q_3^{>} = 29/64 \approx .45$. The failure rate maximum $Q^{MAX} = Q_{0,1} + Q_2^{<} + Q_3^{>} \approx .87$ while $Q_0 \approx .75$ and $\sum \omega_i = 1$ as it should.

3.2.6.1. *First elimination.* To find which subspace to discard first we find the critical Q's: $Q_c^1 \approx .14$, $Q_c^2 \approx .35$, and $Q_c^3 < 0$, so subspace 2 is discarded first when $Q \approx .35$. This means

that $Q_2 = 0$ when $Q = Q_c^2$ and we do not allow the value of Q_2 to vary afterward. At Q_c^2 we find the values of the other failure rates to be $Q_1 \approx .06$ and $Q_3 \approx .29$

3.2.6.2. *Second elimination.* It may be clear that Q_1 will reach 0 first and indeed this is so. Before we find the second set of critical values we find our new constants as: $\Lambda_2 = \sum^{1,3} \omega_i = 3/4$; $F_2 = \sum^{1,3} Q_{0,i} \approx .6$. Now the critical values read $Q_c^{(1)1} \approx .22$ and $Q_c^{(1)3} < 0$ so when $Q = Q_c^{(1)1}$ we discard subspace 1 and reduce the optimization problem to the single subspace case, where $Q^{(2)3} = Q$. This process is depicted in the graph below.

3.3. Interpolative Mixed Qubit Discrimination

Here we discriminate between

$$\rho_1 = p|\psi_1\rangle\langle\psi_1| + \frac{(1-p)}{2}I$$

and

$$\rho_1 = d|\psi_2\rangle\langle\psi_2| + \frac{(1-d)}{2}I$$

Where the pure states $|\psi_1\rangle = c_1|0\rangle + s_1|1\rangle$ and $|\psi_2\rangle = c_2|0\rangle + s_2|1\rangle$ form a 2 dimensional space and p,d are the purities of the two mixed states.

The a-priori probabilities are as usual η_1 and η_2

We want to implement the transformation that gives us

$$\widetilde{\rho_1} = \frac{\Omega^{1/2}\rho_1\Omega^{1/2}}{Tr(\Omega\rho_1)}$$
$$\widetilde{\eta_1} = \frac{\eta_1 Tr(\Omega\rho_1)}{1-Q}$$

So that we still have $Tr\widetilde{\rho_1} = 1$ and $\widetilde{\eta_1} + \widetilde{\eta_2} = 1$ and the error rate is

$$\widetilde{P_e} = \frac{1}{2}(1 - Tr\left|\widetilde{\Lambda}\right|)$$

where $\widetilde{\Lambda} = \widetilde{\eta_1}\widetilde{\rho_1} - \widetilde{\eta_2}\widetilde{\rho_2}$

Hence

$$Tr\left|\widetilde{\Lambda}\right| = \frac{1}{1-Q} Tr\left|\Omega^{1/2}\Lambda\Omega^{1/2}\right|$$

where $\Omega^{1/2} = \begin{pmatrix} \sqrt{1-\xi} & 0 \\ 0 & 1 \end{pmatrix}$ and

$$\Lambda = \begin{pmatrix} p\eta_1 c_1^2 - d\eta_2 c_2^2 + \frac{\eta_1(1-p)-\eta_2(1-d)}{2} & p\eta_1 c_1 s_1 - d\eta_2 c_2 s_2 \\ p\eta_1 c_1 s_1 - d\eta_2 c_2 s_2 & p\eta_1 s_1^2 - d\eta_2 s_2^2 + \frac{\eta_1(1-p)-\eta_2(1-d)}{2} \end{pmatrix}$$

so

$$\Omega^{1/2}\Lambda\Omega^{1/2} =$$

$$\begin{pmatrix} (1-\xi)[p\eta_1 c_1^2 - d\eta_2 c_2^2 + \frac{\eta_1(1-p)-\eta_2(1-d)}{2}] & \sqrt{1-\xi}(p\eta_1 c_1 s_1 - d\eta_2 c_2 s_2)) \\ \sqrt{1-\xi}(p\eta_1 c_1 s_1 - d\eta_2 c_2 s_2)) & p\eta_1 s_1^2 - d\eta_2 s_2^2 + \frac{\eta_1(1-p)-\eta_2(1-d)}{2} \end{pmatrix}$$

To find the sum of the absolute values of its eigenvalues we first write the characteristic equation as

$$\lambda^2 - b\lambda + c = 0$$

where

$$b = \eta_1 - \eta_2 - \xi[p\eta_1 c_1^2 - d\eta_2 c_2^2 + \frac{\eta_1(1-p) - \eta_2(1-d)}{2}]$$

and

$$c = (1-\xi)[\frac{1 - 4\eta_1\eta_2 - (p\eta_1 - d\eta_2)^2}{4} - pd\eta_1\eta_2(c_1 s_2 - c_2 s_1)^2]$$

Where we can rewrite the term $(c_1 s_2 - c_2 s_1)^2 = sin^2\theta$

And we can find ξ by solving

$$1 - Q = Tr\Omega\rho = 1 - \xi[p\eta_1 c_1^2 + d\eta_2 c_2^2 + \frac{1 - p\eta_1 - d\eta_2}{2}]$$

We can solve the quadratic equation $\lambda^2 - b\lambda + c$ for the two eigenvalues, but we notice that for $p = d = 1$ the value of c is $c = (1-\xi)[-\eta_1\eta_2 sin^2\theta]$ is negative and for $p = d = 0$ the value of c is $c = (1-\xi)[\frac{(\eta_1-\eta_2)^2}{4}]$ is positive.

Hence the sum of the eigenvalues for $c < 0$ is

$$\sum |\lambda_i| = \sqrt{b^2 - 4c}$$

And the sum of the eigenvalues for $c \geq 0$ is

$$\sum |\lambda_i| = b$$

We want to look at specific cases for insight into the problem.

3.3.1. b Solution. We consider the case when the sum of the eigenvalues equals to b. The minimum error in this case is

$$P_E = \frac{1-Q}{2}(1-b)$$

The problem remains to minimize this expression as a function of the variables c1 and c2. The relation between the two is that $c1\sqrt{1-c1^2} + c2\sqrt{1-c2^2} = \cos\theta$. Remembering that

$$b = \eta_1 - \eta_2 - \xi[p\eta_1 c_1^2 - d\eta_2 c_2^2 + \frac{\eta_1(1-p) - \eta_2(1-d)}{2}]$$

where

$$\xi = \frac{Q}{p\eta_1 c_1^2 + d\eta_2 c_2^2 + \frac{1-p\eta_1 - d\eta_2}{2}}$$

There are three foreseeable solutions to this problem, where we refer to the bracketed term ajoining ξ as the one of interest:

Case 1: [ ] = 0

Case 2: [ ] is maximum

Case 2: [ ] is minimum

Let's conisder them sequentially.

Case 1:

Here the solution is

$$p\eta_1 c_1^2 = d\eta_2 c_2^2 - \frac{\eta_1(1-p) - \eta_2(1-d)}{2}$$

Case 2:

To maximize the bracketed term we want to maximize c1. By the constraint $c1\sqrt{1-c1^2} + c2\sqrt{1-c2^2} = \cos\theta$ we can do this by setting c2 = 0 and solving for $c1^2$, where we get

$$c1^2 = \frac{1}{2}(1 + \sqrt{1 + 4\cos^2\theta})$$

Case 3:

To minimize the bracketed term we minimize c2. By direct analogy to the previous case we set c1 = 0 and find

$$c2^2 = \frac{1}{2}(1 + \sqrt{1 + 4\cos^2\theta})$$

A more difficult question is when which of the cases applies.

3.3.2. Case: p=d. To find the value of the purity p at which c changes signs and hence the solution changes we let c=0 and solve for the roots of p. We find only one positive root,

$$p_R = \frac{\eta_1 - \eta_2}{2\sqrt{\eta_1\eta_2 sin^2\theta + \frac{(\eta_1-\eta_2)^2}{4}}}$$

This root is always in the range $0 \leq p_R \leq 1$ for $\eta_1 \geq \frac{1}{2}$ and $0 \leq sin^2\theta \leq 1$ as is visible in graph 1.

The value of c is always positive for $sin^2\theta = 0$ and the relationship between p_R and c is linear when $sin^2\theta = 1$

To check our previous assumption let us set p=1 in the resulting expressions to get

$$b = -(\eta_2 - \eta - 1 + Q\frac{\eta_1 c_1^2 - \eta_2 c_2^2}{\eta_1 c_1^2 + \eta_2 c_2^2})$$
$$c = (\eta_1\eta_2)\frac{(Q - \eta_1 c_1^2 - \eta_2 c_2^2)(c_1 s_2 - c_2 s_1)^2}{\eta_1 c_1^2 + \eta_2 c_2^2}$$

Now $b^2 - 4c$ reduces to an expression that when optimized (setting $\eta_1 c_1^2 = \eta_2 c_2^2$) gives $(1-Q)^2 - (Q-Q_0)^2$

We can conclude that the optimal general relationship between the failure rates is $\eta_1 c_1^2 = p\eta_2 c_2^2 + (1-p)F$ where F is some unknown expression.

It is reasonable to assume that since the relative purities of both states remain the same, the optimal relationship between the failure rates should be the same. In this case we still have $\eta_1 c_1^2 = \eta_2 c_2^2$

We can now find the optimal expression for $b^2 - 4c$ in the general case:

$$\frac{1}{(2p\eta_1 c_1^2 + \frac{1-p}{2})^2}[Q^2(\frac{1-p}{2})^2(\eta_1 - \eta_2)^2 - Qp(2p\eta_1 c_1^2 + \frac{1-p}{2})(p[1-Q_0^2] -$$
$$-[\eta_1 - \eta_2]^2) - p^2(2p\eta_1 c_1^2 + \frac{1-p}{2})^2(Q_0^2 - 1)]$$

where we can also make the substitution $\eta_1 c_1^2 = \frac{\eta_1 \eta_2 sin^2\theta}{1-Q_0}$, but this appears to be the simplest form.

It can be shown that this expression also simplifies to $(1-Q)^2 - (Q-Q_0)^2$ when $p = 1$

Similarly, the optimal expression for b in the general case is

$$b = (\eta_1 - \eta_2)[\frac{Q(1-p)}{4p\eta_1 c_1^2 + 1 - p} - 1]$$

3.3.3. Case: d = 1. Stuff

3.3.3.1. $\eta_1 = \eta_2$. other stuff

3.3.4. Case: ME limit. This problem is fairly simple since there is no failure. We have no need for the transformation and immediately get

$$P_e = \frac{1}{2}(1 - Tr\,|\Lambda|)$$

finding the simplified b and c values as

$$b = \eta_2 - \eta_1$$

$$c = \frac{1}{4}[(\eta_1^2 - \eta_2^2) - (\eta_1 p - \eta_2 d)^2] - \eta_1\eta_2 pd sin^2\theta$$

for $c > 0$ we get the solution b and for $c < 0$ we get the $\sqrt{b^2 - 4c}$ solution, with no further optimization necessary. For given initial conditions, positivity doesn't change during the interpolation.

If b is the answer, we find the error rate to be $P_e = \frac{1}{2}(1-(\eta_1-\eta_2)) = \eta_2$ assuming $\eta_1 > \eta_2$ This strategy is analogous to guessing the more likely state.

Otherwise

$$P_e = \frac{1}{2}[1 - \sqrt{(\eta_1 p + \eta_2 d)^2 - 4\eta_1\eta_2 pd cos^2\theta}]$$

which reduces to the Helstrom bound when p=d=1.

The critical point c= 0 occurs when

$$cos^2\theta = \frac{-(\eta_1 - \eta_2)^2 + (\eta_1 p + \eta_2 d)^2}{4\eta_1\eta_2 pd}$$

3.3.5. Case: MC limit. Not pretty...Numerically Feasible?

3.3.6. General Optimization Arguments. To understand how to optimize the two solutions we give two arguments for optimization:

For the first, consider a fixed failure operator Π_0. This fixes the value of ξ and we notice that c is independent of c_1 and c_2, the last two mutually dependent free variables; it only depends on the total overlap of the two pure states θ. Hence the minimization of error requires minimization of b. This is achieved by the solution:

$$p\eta_1 c_1^2 - d\eta_2 c_2^2 + \frac{\eta_1(1-p) - \eta_2(1-d)}{2} = 0$$

or

$$p\eta_1 c_1^2 + \frac{\eta_1(1-p)}{2} = d\eta_2 c_2^2 + \frac{\eta_2(1-d)}{2}$$

The second argument is as follows:

We know there must be an optimal relationship between the two overlap terms, and we predict that it involves the full individual failure rates, so we can write one as a function of the other:

$$p\eta_1 c_1^2 + \frac{\eta_1(1-p)}{2} = F(d\eta_2 c_2^2 + \frac{\eta_2(1-d)}{2})$$

Now we expand the function F as a Taylor series:

$$F(d\eta_2 c_2^2 + \frac{\eta_2(1-d)}{2}) = \sum_{n=0}^{\infty} \frac{a_n}{n!}(d\eta_2 c_2^2 + \frac{\eta_2(1-d)}{2})^n$$

We can evaluate this for p=d=1 when we know the solution. Here

$$\eta_1 c_1^2 = \sum_{n=0}^{\infty} \frac{a_n}{n!}(\eta_2 c_2^2)^n = \eta_2 c_2^2$$

Hence $a_{n\neq 1} = 0$ and $a_1 = 1$

3.4. Confidence-Valued Measurements

CHAPTER 4

Cloning Pure States

In this chapter we discuss the transformation which takes an input state and ancilla together to make extra versions of the input state. The conversation began with a paper on superluminal communication [?] by making copies of quantum states. This was soundly refuted by [?, ?] who gave proofs of the no-cloning theorem for quantum states. We include a simple version here:

4.0.1. No-Cloning Theorem. Suppose we wrote the equations describing the unitary that cloned one of two input pure states ψ_i with a-priori probabilities η_i for $i = 1, 2$. If we could perform this measurement perfectly we would write

$$U|\psi_1\rangle|0\rangle = |\psi_1\rangle|\psi_1\rangle \tag{4.0.1}$$

$$U|\psi_2\rangle|0\rangle = |\psi_2\rangle|\psi_2\rangle \tag{4.0.2}$$

However, taking the inner product of the two equations we find $\langle\psi_1|\psi_2\rangle = \langle\psi_1|\psi_2\rangle^2$, restricting the functionality of this unitary to the trivial case when the two states are orthogonal and $\langle\psi_1|\psi_2\rangle = 0$. Therefore there does not exist in general a unitary to perform this ideal cloning task. As with state discrimination, we must choose a figure of merit to maximize for given initial conditions.

4.1. Deterministic Approximate Cloning

In this section we derive the 'minimum error' cloning method and make two imperfect copies of one of two input states. The input states are

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

$$|\psi_2\rangle = \cos\theta|0\rangle - \sin\theta|1\rangle,$$

and the output states are

$$|\phi_1\rangle = \cos\phi_1|0\rangle + \sin\phi_1|1\rangle,$$

$$|\phi_2\rangle = \cos\phi_2|0\rangle - \sin\phi_2|1\rangle.$$

The unitary representation of this transformation is

$$U|\psi_1\rangle = |\phi_1\rangle|\phi_1\rangle$$

$$U|\psi_2\rangle = |\phi_2\rangle|\phi_2\rangle$$

The function we want to maximize is the average fidelity

$$F = \eta_1\langle\psi_1|\phi_1\rangle^2 + \eta_2\langle\psi_1|\phi_1\rangle^2$$

The Lagrange multiplier method isn't necessary here as we can rephrase the problem in just one variable. First we rewrite the fidelity by noticing $\langle\psi_1|\phi_1\rangle = \cos(\theta - \phi - 1) = \cos(\theta - \frac{\phi_1+\phi_2}{2} - \frac{\phi_1-\phi_2}{2}) = \cos(\alpha - x)$. Similarly $\langle\psi_2|\phi_2\rangle = \cos(\theta - \phi_2) = \cos(\alpha + x)$ so that

$$F = \eta_1\cos^2(\alpha - x) + \eta_2 cos^2(\alpha + x)$$

is just a function of x. Differentiating to find the minimum we find the optimal relation between alpha and x as

$$\tan 2x = (\eta_1 - \eta_2)\tan 2\alpha$$

To make use of this relation we can rewrite the fidelity as

$$F = \eta_1(\cos(2(\alpha - x)) + 1)/2 + \eta_2(\cos(2(\alpha + x)) + 1)/2 =$$

$$\frac{1}{2}[\cos 2\alpha\cos 2x(1 + (\eta_1 - \eta_2)\tan 2\alpha\tan 2x)] + \frac{1}{2}.$$

And we can rewrite

$$\cos 2x = \frac{1}{\sqrt{1 + (\eta_1 - \eta_2)^2\tan^2 2\alpha}}$$

to eliminate x to write

$$F = \frac{1}{2}(\cos 2\alpha\sqrt{1 + (\eta_1 - \eta_2)^2 \tan^2 2\alpha}) + \frac{1}{2}.$$

Inserting the cos term into the square root we find

$$\cos^2 2\alpha + (\eta_1 - \eta_2)^2 \sin^2 2\alpha = \cos^2 2\alpha + (\eta_1 + \eta_2)^2 \sin^2 2\alpha - 4\eta_1\eta_2 \sin^2 2\alpha = 1 - 4\eta_1\eta_2 \sin^2 2\alpha,$$

leaving us with the fidelity in its final form as

$$F = \frac{1}{2}(1 + \sqrt{1 - 4\eta_1\eta_2 \sin^2 2\alpha}).$$

4.2. Probabilistic Exact Cloning

We envision a state dependent probabilistic cloner as a machine with an input port, an output port and two flags that herald the success or failure of cloning. The input $|\psi_i^m\rangle = |\psi_i\rangle^{\otimes m}$, $i = 1, 2$ (m identical copies of either $|\psi_1\rangle$ or $|\psi_2\rangle$) is fed through the input port for processing. In case of success, n perfect clones $|\psi_i^n\rangle = |\psi_i\rangle^{\otimes n}$ are delivered through the output port with probability p_i, conditioned on the input state being $|\psi_i^m\rangle$. Otherwise, the output is in a generic failure state, with a failure probability $q_i = 1 - p_i$.

For cloning, optimality is usually addressed from a Bayesian viewpoint that assumes the states to be cloned are prepared with some prior probabilities η_1 and η_2, $\eta_1 + \eta_2 = 1$. Then a natural cost function for the probabilistic cloning machine is the average failure probability,

$$Q = \eta_1 q_1 + \eta_2 q_2. \tag{4.2.1}$$

The aim is to find the optimal cloner that minimizes the cost function Q, and yields the minimum average failure probability $Q_{\min}$ for arbitrary priors η_1 and η_2.

In our formulation, similar to that in [?], the Hilbert space $\mathcal{H}^{\otimes m}$ of the original m copies is supplemented by an ancillary space $\mathcal{H}^{\otimes(n-m)} \otimes \mathcal{H}_F$ that accommodates the additional $n - m$ clones and the success/failure flags. Next, we introduce a unitary transformation U

via

$$U|\psi_i^m\rangle|0\rangle = \sqrt{p_i}|\psi_i^n\rangle|\alpha_i\rangle + \sqrt{q_i}|\Phi^n\rangle|\alpha_0\rangle, \quad i = 1, 2. \tag{4.2.2}$$

Here the ancillas are initialized in a reference state $|0\rangle$. The states of the flag associated with successful cloning, $|\alpha_i\rangle$, are orthogonal to the state associated with failure, $|\alpha_0\rangle$, and $|\Phi^n\rangle$ is a generic failure state in $\mathcal{H}^{\otimes n}$, the same for both inputs for optimality.

We note that, although $|\alpha_1\rangle = |\alpha_2\rangle$ for optimal cloning, we consider a more general scenario allowing for $|\alpha_1\rangle \neq |\alpha_2\rangle$, to include "cloning by discrimination". In this protocol we employ UD to identify the input state and then prepare clones of the identified state. For UD the success flag states must be fully distinguishable, so $\langle\alpha_1|\alpha_2\rangle = 0$. Further, an even more general scenario could allow for two different failure states, $|\Phi_i^n\rangle$ $(i = 1, 2)$, in Eqs. (4.2.2). This, however, would necessarily be sub-optimal since we could probabilistically determine whether we received $|\psi_1\rangle$ or $|\psi_2\rangle$ by applying UD to the states $|\Phi_i^n\rangle$. In case of success, we could prepare n copies of the state, increasing the overall success rate of the cloning strategy.

To proceed, we now take the inner product of equation i with itself in (4.2.2), yielding that the probabilities are normalized, $p_i + q_i = 1$. Taking the inner product of equation (4.2.2) with $i = 1$ and $i = 2$ yields the main constraint,

$$s^m = \sqrt{p_1 p_2}\, s^n \alpha + \sqrt{q_1 q_2}, \tag{4.2.3}$$

which is a consequence of the unitarity of U. Here we introduced the notation $s \equiv \langle\psi_1|\psi_2\rangle$ and $\alpha \equiv \langle\alpha_1|\alpha_2\rangle$. The overlaps s and α can be chosen to be real without any loss of generality, so $0 \leq s, \alpha \leq 1$. Clearly, $\alpha = 1$ for optimal cloning, while $\alpha = 0$ for cloning by discrimination. If Eq. (5.0.17) is satisfied, it is not hard to prove that U has a unitary extension on the whole space.

4.2.1. Geometric description of optimality. Before developing the analytical theory of optimizing (minimizing) Q, we present a complete geometric picture that visually solves the optimization problem and serves as guide for the subsequent calculations. Such

a geometric viewpoint proved particularly useful in obtaining a visual solution to optimization under highly nonlinear constraints [?]. For the geometrization the following features of (5.0.17) turn out to be important.

(1) For fixed s, n and m (5.0.17) defines a class of smooth curves on the unit square $0 \leq q_i \leq 1$ (e.g., solid, dashed or dotted curves in Fig. 5.0.1).

(2) All these curves meet at their endpoints, $(1, s^{2m})$ and $(s^{2m}, 1)$.

(3) At the endpoints the curves become tangent to the vertical and horizontal lines $q_1 = 1$ and $q_2 = 1$, respectively, provided $\alpha \neq 0$.

(4) For $\alpha = 0$ the curve defined by Eq. (5.0.17) is the hyperbola $q_1 q_2 = s^{2m}$ (dashed line in Fig. 5.0.1).

(5) The curve corresponding to a particular value of α and the segments joining its end points with the vertex $(1, 1)$ form the boundary of the set S_α (any of the gray regions in Fig. 5.0.1), where

$$S_\alpha = \{(q_1, q_2) : \sqrt{p_1 p_2}\, s^n \alpha + \sqrt{q_1 q_2} - s^m \geq 0\}. \tag{4.2.4}$$

The sets satisfy $S_\alpha \subset S_{\alpha'}$ if $\alpha < \alpha'$.

(6) The sets S_α are convex if $\alpha \geq 0$. In particular S_1 is convex.

These considerations lead to the emergence of a geometrical picture of the optimization problem which we display in Fig. 5.0.1.

Eq. (5.0.14) defines a straight segment on the unit square $0 \leq q_i \leq 1$ with a normal vector in the first quadrant parallel to (η_1, η_2). For fixed priors, the average failure probability Q is proportional to the distance from this segment to the origin $(0, 0)$. Since S_1 is convex and the stretch of its boundary given by Eq. (5.0.17) with $\alpha = 1$ is smooth, a unique point (q_1, q_2) of tangency with the segment (5.0.14) exists for any value of the priors and finite n. It gives $Q_{\min}$ and defines the optimal cloning strategy.

We note that the inclusion hierarchy of the sets S_α provides a simple geometrical proof that $\alpha = 1$, i.e., $|\alpha_1\rangle = |\alpha_2\rangle$, is the optimal choice for cloning. For "cloning by discrimination", on the other hand, $\langle\alpha_1|\alpha_2\rangle = \alpha = 0$. From points (4) and (5) above, it follows that for any

finite n this protocol is strictly suboptimal, i.e., $Q_{\min} < Q_{\rm UD}$, noticing that the failure rate of cloning by discrimination is that of UD, $Q_{\rm UD}$. This is intuitively obvious since in cloning one is asking for less than in UD; the identity of the input states is not revealed for any finite n. However, optimal cloning and UD become one and the same in the limit $n \to \infty$, when $s^n \to 0$ and the curve (5.0.17) collapses to the hyperbola $q_1 q_2 = s^{2m}$, as it does for $\alpha = 0$. We return to this point below.

4.2.2. Parametrization. A more quantitative analysis requires finding a convenient parametrization of the constraint (5.0.17). To this end, we write $\sqrt{q_i} = \sin\theta_i$ for $0 \le \theta_i \le \pi/2$. By further introducing the variables $x = \cos(\theta_1 + \theta_2)$ and $y = \cos(\theta_1 - \theta_2)$ we manage to linearize the curve (5.0.17), which now is the straight segment $2s^m = (1+s^n)y - (1-s^n)x$ with $|x| \le y \le 1$. Its guiding vector is readily seen to be $(1+s^n, 1-s^n)$, so the segment's parametric equation can be written as

$$x = \frac{1-(1+s^n)t}{s^{n-m}}, \qquad y = \frac{1-(1-s^n)t}{s^{n-m}}, \tag{4.2.5}$$

where we have rescaled the guiding vector so that the final expressions are simplest. Because of the symmetry of this procedure, the parameters x and y are invariant under $q_1 \leftrightarrow q_2$ (equivalently, under $\theta_1 \leftrightarrow \theta_2$). Thus, the two mirror halves of the curve (5.0.17) under this transformation are mapped onto the same straight line (5.1.2). By expressing q_i as a function of t only half of the original curve is recovered. The other half is trivially obtained by applying $q_1 \leftrightarrow q_2$. After putting the various pieces together one can easily get rid of the trigonometric functions and express Eq. (5.0.17) in parametric form as

$$q_i = \frac{1 - xy - (-1)^i\sqrt{1-x^2}\sqrt{1-y^2}}{2}, \quad i = 1, 2. \tag{4.2.6}$$

Fig. 5.1.1 shows examples of the unitary curve (5.0.17).

For $n > 2$ the curves closely approximate $q_1 q_2 = s^{2m}$ (dashed lines) for small and moderate values of s, while for s close to 1 the hyperbolas remain closer to the vertex $(1, 1)$, but retain the same end points. In the limit $n \to \infty$ all these curves become hyperbolic.

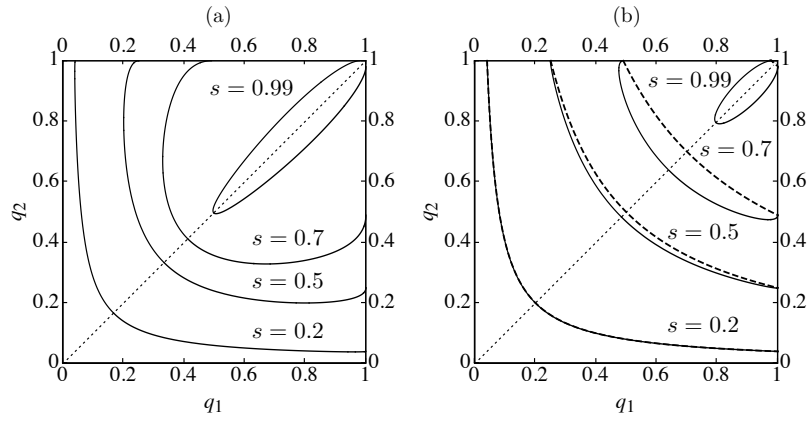

FIGURE 4.2.1. Unitarity curves, q_2 vs. q_1, from Eq. (5.0.17) for different values of s and for (a) $m = 1$, $n = 2$ and (b) $m = 1$, $n = 5$. The curves are symmetric under mirror reflection along the dotted line $q_1 = q_2$, i.e., under the transformation $q_1 \leftrightarrow q_2$. The dashed lines in (b) are the hyperbolae $q_1 q_2 = s^{2m}$.

4.2.3. Optimization. We now return to finding the minimum, $Q_{\min}$, of the average failure probability Q. Despite its apparent simplicity, this involves solving a high-order equation without a simple form. Instead, we will derive a parametric equation for $Q_{\min}(\eta_1)$. Along with the complete description of the unitary curve (5.0.17), this provides a complete solution of the problem in parametric form.

With no loss of generality we may assume $\eta_1 \le \eta_2$ or, equivalently, $0 \le \eta_1 \le 1/2$. Then the slope of the straight line (5.0.14), $-\eta_1/\eta_2$, satisfies $-1 \le -\eta_1/\eta_2 \le 0$. Hence, it can only become tangent to the lower half of the unitarity curve (5.0.17) (see Fig. 5.1.1). Increasing q_1, the slope of this lower half increases monotonically from -1 at $q_1 = q_2$, to 0 before we reach the line $q_1 = 1$. This follows from the properties (a)–(f) above and can be checked using Eq. (5.1.1). The values of t at which the slope is -1 and 0 are, respectively,

$$t_{-1} = \frac{1 - s^{n-m}}{1 - s^n}, \quad t_0 = \frac{1 - s^{2(n-m)}}{1 - s^{2n}}. \tag{4.2.7}$$

For any point $(q_1(t), q_2(t))$ with $t \in [t_{-1}, t_0]$ there is a line $Q = \eta_1 q_1 + \eta_2 q_2$ that is tangent to it, starting with $\eta_1 = \eta_2 = 1/2$ for $t = t_{-1}$ up to $\eta_1 = 0$, $\eta_2 = 1$ for $t = t_0$.

This observation enables us to derive the desired parametric expression for the optimality curve $Q_{\min}(\eta_1)$ as follows. For a given t in the range above, a necessary condition for tangency

is $\eta_1 q_1' + \eta_2 q_2' = 0$, where $q_i' = dq_i/dt$. In this equation we can solve for η_1 (or η_2) using that $\eta_1 + \eta_2 = 1$. By substituting q_1 and q_2 in Eq. (5.0.14) with (5.1.1) we enforce contact with the unitarity curve and obtain the expression of $Q_{\min}$. The final result can be cast as:

$$\eta_1 = \frac{q_2'}{q_2' - q_1'}, \quad Q_{\min} = \frac{q_2' q_1 - q_1' q_2}{q_2' - q_1'}, \quad t_{-1} \leq t \leq t_0, \tag{4.2.8}$$

where t_{-1}, t_0 and q_i are given in Eqs. (5.1.3) and (5.1.1). The expressions for the derivatives q_i' are

$$q_i' = \frac{\sqrt{q_i(1-q_i)}}{s^{n-m}} \left\{ \frac{1+s^n}{\sqrt{1-x^2}} - (-1)^i \frac{1-s^n}{\sqrt{1-y^2}} \right\}. \tag{4.2.9}$$

4.2.4. Relation of perfect cloning with discrimination. Fig. 5.3.1 shows plots of the curves $Q_{\min}(\eta_1)$ for $m = 1$ input copies and (a) $n = 2$ or (b) $n = 5$ clones, as in the previous figure. We see that $Q_{\min}$ is an increasing function of η_1 in the given range $[0, 1/2]$. The values of $Q_{\min}$ at the end points of this range follow by substituting t_0 and t_{-1}, Eq. (5.1.3), into Eq. (5.1.1). They are given by

$$Q_0 = q_2(t_0) = \frac{s^{2m} - s^{2n}}{1 - s^{2n}}, \quad Q_{-1} = \frac{s^m - s^n}{1 - s^n}, \tag{4.2.10}$$

where $Q_{\min} = Q_{-1}$ holds for equal priors and $Q_{\min} = Q_0$ for $\eta_1 \to 0$ (i.e., $\eta_2 \to 1$). The

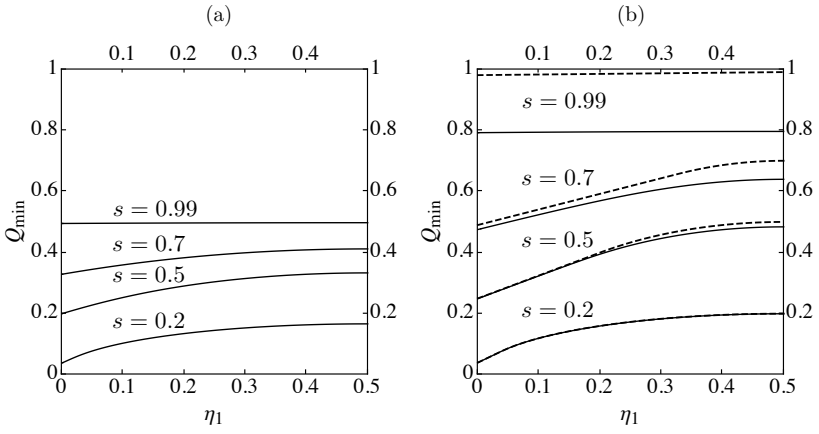

FIGURE 4.2.2. Minimum cloning failure probability $Q_{\min}$ vs. η_1 (solid lines) and UD failure probability $Q_{\rm UD}$ vs. η_1 (dashed lines) for the same values of m, n and s used in the previous figure.

dashed lines in Fig. 5.3.1 (b) depict the well known optimal UD solution [**?**]:

$$Q_{\rm UD} = \begin{cases} 2\sqrt{\eta_1\eta_2}\, s^m, & \dfrac{s^{2m}}{1+s^{2m}} \le \eta_1 \le \dfrac{1}{2}; \\ \eta_1 + s^{2m}\eta_2, & 0 \le \eta_1 \le \dfrac{s^{2m}}{1+s^{2m}}. \end{cases} \tag{4.2.11}$$

It is apparent from these plots that the optimal cloning protocol performs strictly better than cloning by discrimination, as was geometricaly proved in Figs. 5.0.1 and 5.1.1. However, the difference in performance decreases with increasing number of clones. In Fig. 5.3.1 (b), for only $n = 5$, a difference is hardly noticeable for $s \le 0.5$. For $s > 0.5$ the convergence is slower but in the limit $n \to \infty$ there is perfect agreement for any $s < 1$.

The convergence of the optimal cloning failure probability, $Q_{\rm min}$, to that of cloning by discrimination, $Q_{\rm UD}$ in Eq. (5.0.21) follows from our geometrical approach. Recall that in the limit $n \to \infty$ (or equivalently $\alpha \to 0$) the right hand side of Eq. (5.0.17) describes hyperbolas that we can write as $q_2 = s^{2m}/q_1$ (dashed lines in Figs. 5.0.1 and 5.1.1). Their slopes are in the range $[-1, -s^{2n}]$. A unique point of tangency with the line (5.0.14) can only exists if the slope of this line, $-\eta_1/\eta_2$, is within this same range. This gives the η_1 interval in the first line of (5.0.21), and one can easily obtain the corresponding expression for $Q_{\rm UD}$. If the slope of the line (5.0.14) is outside the range, tangency is not possible, and the optimal line merely touches the end points of the hyperbolas, so the expression of $Q_{\rm UD}$ becomes the second line of Eq. (5.0.21). In geometrical terms, the straight line (5.0.14) pivots on the end points as we vary η_1. Furthermore, note that for the second line in (5.0.21) we have $p_1 = 1 - q_1 = 0$, which leads to a 2-outcome projective measurement, as only one success flag state ($|\alpha_2\rangle$) is needed in Eqs. (4.2.2).

Interestingly, in this limit a phenomenon analogous to a second order phase transition takes place. Our geometrical approach shows that the average failure probability $Q_{\rm min}(\eta_1)$ is an infinitely differentiable function of η_1 for finite n. However, as n goes to infinity (corresponding to $\alpha = 0$) the limiting function $Q_{\rm UD}(\eta_1)$ has a discontinuous second derivative. Moreover, the symmetry $q_1 \leftrightarrow q_2$ breaks in the "phase" corresponding to the first line in Eq. (5.0.21). A similar phenomenon arises in UD of more than two pure states [**?**].

It has been argued above that cloning by discrimination is strictly suboptimal (unless $n \to \infty$). One could likewise wonder if discrimination by cloning can be optimal. On heuristic grounds, one should not expect this to be so, as cloning involves a measurement and some information can be drawn from the observed outcome. However, the equal-prior and the $\eta_1 \to 0$ cases provide remarkable exceptions. For both we may write the total failure rate as $Q_{\rm C} + (1 - Q_{\rm C})Q_{\rm UD}$, where C stands for cloning. For $\eta_1 = \eta_2 = 1/2$, Eq. (5.1.6) implies $Q_{\rm C} = Q_{-1}$, in which case the produced n-clone states are equally likely. The UD of these states fails with probability s^n, as follows from Eq. (5.0.21) applied to n copies. The total failure rate is then s^m, which is the optimal UD failure rate of the original input states, Eq. (5.0.21). If $\eta_1 \to 0$ then only $|\psi_2^n\rangle$ is produced with non-vanishing probability and $Q_{\rm C} = Q_0$. Failure in the second step (UD) is given by the top line in Eq. (5.0.21) applied to n copies. The total failure rate is s^{2m}, also achieving optimality.

Using our main result in Eqs. (5.1.4), and (5.1.1) one can check that these are the only cases where discrimination by cloning is optimal. These are also the only cases where no information gain can be drawn from the cloning measurement. This hints at how special these cases are and justifies the need of the derived solution for arbitrary priors to have a full account of two-state cloning.

4.3. Hybrid Cloning: Interpolation between exact and approximate cloning

In this section we seek to interpolate between probabilistic exact cloning and approximate cloning machines using our results from state separation. Exact cloning machines produce perfect clones while allowing for some inconclusive outcomes. Approximate cloning machines produce copies on demand which resemble the input states by maximizing the fidelity. One can imagine a scheme where fidelity can be higher then maximum fidelity in the approximate cloning machine while it allows for a fixed rate of inconclusive outcomes, FRIO. This scheme should reproduce exact cloning and approximate cloning machines by setting FRIO to Q_o and zero respectively. Chefles and Barnett [**?**] solved the problem for the case when the input states are prepared with equal a priori probabilities. We extend the solution to the more

general case when the states are prepared with different priors. Such a solution is possible due to our recent work on making N perfect clones from M copies of one of two known pure states with minimum failure probability in the general case where the known states have arbitrary priori probabilities.

4.3.1. Equal priors. The solution to the interpolation of cloning for equal a priori probabilities has been derived by Chefles et al [**?**]. The authors develop a scheme which, depending on the fidelity of the clones, can interpolate between exact cloning with inconclusive results in one extreme and optimal approximate cloning on the other extreme. In our work this scheme has been generalized for the case when the input states are prepared with different a priori probabilities. First we show the derivation of the equal priors as it will help to better understand the general case.

For $\eta_1 = \eta_2 = 1/2$, the output states are symmetric, $\phi_1 = \phi_2 = \phi$, and the optimal global fidelity, F_{MN}, in Eq.(**??**) reduces to:

$$\begin{aligned} F_{MN} &= \frac{1}{2}\left[1 + \sqrt{1 - \sin^2(2\theta - 2\phi)}\right], \\ &= \frac{1}{2}\left[1 + \cos^2(2\theta - 2\phi)\right]. \end{aligned} \tag{4.3.1}$$

Duan and Guo [**?**] showed that the maximum success probability of obtaining N exact clones from M given copies of non-orthogonal quantum states $\{|\psi_1\rangle, |\psi_2\rangle\}$, which are prepared with equal a priori probabilities, is:

$$P_{MN} = \frac{1 - s^M}{1 - s^N}, \tag{4.3.2}$$

where s is the overlap of the input states $s = \langle\psi_1|\psi_2\rangle$. The success rate for 1 to 2 cloning, $M = 1, N = 2$, reduces to:

$$P_{12} = \frac{1}{1+s}. \tag{4.3.3}$$

The interpolation takes us from optimal exact cloning to maximum fidelity. Given a set K of two non-orthogonal quantum states, $\{|\psi_1\rangle, |\psi_2\rangle\}$ the goal is to make N clones $\{|\phi_1\rangle, |\phi_2\rangle\}$, which are similar to the input states but not perfect. The Neumark setup is:

$$U|\psi_1\rangle^{\otimes M}|i\rangle = \sqrt{p}|\phi_1\rangle^{\oplus N}|1\rangle + \sqrt{q}|f\rangle|0\rangle \tag{4.3.4}$$

$$U|\psi_2\rangle^{\oplus M}|i\rangle = \sqrt{p}|\phi_2\rangle^{\oplus N}|1\rangle + \sqrt{q}|f\rangle|0\rangle \tag{4.3.5}$$

The input states are prepared with equal a priori probabilities. A click in the $|1\rangle$ direction means that we succeed in making the clones and the probability of success is p. A click in the $|0\rangle$ direction means that we failed to create a clone with a probability q. The inner product or (4.3.4) and (4.3.5) gives the constraint:

$$s^M = ps'^N + q \tag{4.3.6}$$

Using the unitarity condition $p + q = 1$, the average rate of successfully making a clone is:

$$p = \frac{1 - s^M}{1 - s'^N} \tag{4.3.7}$$

s' is the overlap of the clones $s' = \langle\phi_1|\phi_2\rangle$. If the final states are orthogonal, $s' = 0$ then the state separation reaches the IDP limit and $P_S = P_{IDP} = 1 - |\langle\psi_1|\psi_2\rangle|^M$.

First we express the overlap of the output states in terms on the success rates and the overlap of input states, $\cos 2\theta = |\langle\psi_1|\psi_2\rangle|^N$

$$|\langle\phi_1|\phi_2\rangle|^N = 1 - \frac{1 - |\langle\psi_1|\psi_2\rangle|^M}{P_S} \tag{4.3.8}$$

$$\cos^N(\phi_1 + \phi_2) = 1 - \frac{P_{IDP}}{P_S} \tag{4.3.9}$$

The exact clones live in an N dimensional space $|\psi_{1,2}^N\rangle = \cos\theta|1\rangle \pm \sin\theta|0\rangle$. The approximate clones can be expressed as $|\phi_{1,2}\rangle = \cos\phi_1|1\rangle \pm \sin\phi_1|0\rangle$.

The fidelity rate for equal priors is:

$$F_{MN} = \frac{1}{2}\left[1 + \cos(2\theta - (\phi_1 + \phi_2)\right], \tag{4.3.10}$$

and we want to use the relationship in (4.3.9). Let us expand the cosine term

$\cos(2\theta - (\phi_1 + \phi_2)) = \cos 2\theta \cos(\phi_1 + \phi_2) + \sin 2\theta \sin(\phi_1 + \phi_2)$.

The fidelity becomes:

$$F_{MN} = \frac{1}{2}\left[1 + |\langle\psi_1^N|\psi_2^N\rangle|\left(1 - \frac{P_{IDP}}{P_S}\right) + \frac{1}{P_S}(\left(1 - |\langle\psi_1^N|\psi_2^N\rangle|^2\right)\left(P_S^2 - (P_S - P_{IDP})^2\right)^{1/2}\right]$$

As $N \to \infty$, $|\langle\psi_1|\psi_2\rangle|^N \to 0$ and F_{MN} reduces to

$$F_{MN} = \frac{1}{2}[1 + \frac{1}{P_S}\sqrt{P_S^2 - (P_S - P_{IDP})^2}].$$

We can also express the fidelity in terms of fixed failure rate $Q = 1 - P_S$ which serves as the parameter by which we are interpolating and the optimal failure rate $Q_o = |\langle\psi_1|\psi_2\rangle|$

$$\begin{aligned} F_{MN} &= \frac{1}{2}\left[1 + |\langle\psi_1^N|\psi_2^N\rangle|(1 - \frac{1-Q_o}{1-Q}) + \frac{1}{1-Q}(\left(1 - |\langle\psi_1^N|\psi_2^N\rangle|^2\right)\left((1-Q)^2 - (Q-Q_o)^2\right)^{1/2}\right], \\ &= \frac{1}{2(1-Q)}\left[(1-Q) + Q_o^N(Q_o - Q) + \sqrt{(1-Q_o^{2N})\left[(1-Q)^2 - (Q-Q_o)^2\right]}\right]. \end{aligned}$$

In the limit $N \to \infty$, $|\langle\psi_1|\psi_2\rangle|^N \to 0$

$$\begin{aligned} F_{MN} &= \frac{1}{2}\left[1 + \frac{1}{1-Q}\sqrt{(1-Q)^2 - (Q-Q_o)^2}\right], \\ (1-Q)F_{MN} &= \frac{1}{2}\left[(1-Q) + \sqrt{(1-Q)^2 - (Q-Q_o)^2}\right]. \end{aligned}$$

$(1-Q)F_{MN} = P_{succes}$, the probability of successfully identifying a state.

$$P_{success} = \frac{1}{2}[(1-Q) + \sqrt{(1-Q)^2 - (Q-Q_o)^2}]$$

(This is a different success rate then the P_S defined above, the P_S was defined as the rate of successfully carrying out a state separation.)

This formula describes the relationship between the discrimination of states with a fixed rate of inconclusive outcome. When $Q = 0$ it reaches the Helstrom bound of minimum error and when $Q = Q_o$ it reaches the IDP limit in UD.

4.3.2. General case. We would like to generalize the above results for the case when the incoming states are prepared with different prior probabilities.

- Step 1: State Separation

Optimally separate the incoming states $\left\{\left|\psi_1^M\right\rangle, \left|\psi_2^M\right\rangle\right\}$ with a fixed rate of inconclusive results q_i, then prepare states $\left\{\left|\psi_1^N\right\rangle, \left|\psi_2^N\right\rangle\right\}$ with the corresponding success probabilities.

$$\begin{aligned} U|\psi_1^M\rangle|0\rangle &= \sqrt{p_1}|\phi_1\rangle|1\rangle + \sqrt{q_1}|\Phi\rangle|2\rangle, \\ U|\psi_2^M\rangle|0\rangle &= \sqrt{p_2}|\phi_2\rangle|1\rangle + \sqrt{q_2}|\Phi\rangle|2\rangle, \end{aligned} \tag{4.3.11}$$

The incoming states are separated with a success probability p_i and failed to separate the states with a failure probability q_i. The inner product of the two equations gives the unitarity constraint

$$s = \sqrt{p_1 p_2}s' + \sqrt{q_1 q_2} \tag{4.3.12}$$

- Step 2: Optimize Fidelity

The fidelity for state $|\psi_1\rangle$ is: $F_1 = \left|\left\langle\psi_1^N \mid \phi_1\right\rangle\right|^2$. Similarly the fidelity for state $|\psi_2\rangle$is $\left|\left\langle\psi_2^N \mid \phi_2\right\rangle\right|^2$. The overall fidelity is

$$F = \frac{\eta_1 p_1 F_1 + \eta_2 p_2 F_2}{\eta_1 p_1 + \eta_2 p_2} = \frac{\eta_1 p_1 F_1 + \eta_2 p_2 F_2}{1 - Q} = \tilde{\eta}_1 F_1 + \tilde{\eta}_2 F_2,$$

where the normalized a priori probabilities are $\tilde{\eta}_i = \frac{\eta_i p_i}{1-Q}$. The average fidelity is the same as calculated in (**??**) with the new normalized probabilities:

$$\begin{aligned} F_{MN} &= \frac{1}{2}\left[1 + \sqrt{1 - 4\tilde{\eta}_1\tilde{\eta}_2 \sin^2\left(2\theta - (\phi_1 + \phi_2)\right)}\right], \\ &= \frac{1}{2(1-Q)}\left[(1-Q) + \sqrt{(1-Q)^2 - 4\eta_1\eta_2 p_1 p_2 \sin^2\left(2\theta - (\phi_1 + \phi_2)\right)}\right] \end{aligned} \quad (4.3.13)$$

It can be seen that in the limit $N \to \infty$, expanding the sin term as we did in the previous section, the FRIO [**?**] results are recovered. It again shows a close relationship between fidelity and state discrimination.

Solving the problem of hybrid cloning however requires one last optimization, that of the second term under the square root

$$\begin{aligned} \Lambda &= \sqrt{p_1 p_2}\sin\left(2\theta - (\phi_1 + \phi_2)\right), \\ &= \sqrt{p_1 p_2}\sqrt{1 - s^{2n}}s' - \sqrt{p_1 p_2(1 - s'^2)}s^n \\ &= \sqrt{1 - s^{2n}}\left(s - \sqrt{q_1 q_2}\right) - s^n\sqrt{1 - (q_1 + q_2) + q_1 q_2 - \left(s - \sqrt{q_1 q_2}\right)^2}, \\ &= \sqrt{1 - s^{2n}}\left(s - u\right) - s^n\sqrt{1 - s^2 - 2v + 2sv}. \end{aligned}$$

Here $u \equiv \sqrt{q_1 q_2}, \quad v \equiv \frac{1}{2}(q_1 + q_2)$ and we used the constraint from the unitarity in (**??**) to replace $\sqrt{p_1 p_2}s' = s - \sqrt{q_1 q_2}$.

CHAPTER 5

State Separation

We can always imagine that a probabilistic quantum transformation is carried out by a machine with an input port, an output port and two flags that herald the success or failure of the transformation. The input $|\psi_i\rangle$, $i = 1, 2$ is fed through the input port for processing. In case of success, states $|\psi_i'\rangle$, with the desired degree of separation, are delivered through the output port with conditioned probability p_i. Otherwise, the output is in a failure state. Conditioned on the input state being $|\psi_i\rangle$, the failure probability is $q_i = 1 - p_i$.

We address optimality from a Bayesian viewpoint that assumes the states to be transformed are given with some *a priori* probabilities η_1 and η_2, $\eta_1 + \eta_2 = 1$. Then a natural cost function for our probabilistic machines is given by the average failure probability

$$Q = \eta_1 q_1 + \eta_2 q_2. \tag{5.0.14}$$

If $|\psi_i\rangle$ and the corresponding transformed states $|\psi_i'\rangle$ are given, the optimal machine is one that minimizes the cost function Q. In this case our aim is to find that optimal machine and the minimum average failure probability $Q_{\min}$ for arbitrary priors η_1 and η_2.

A different way of approaching optimality may consist in finding the machine (or machines) that achieves the highest degree of separation, namely, minimizes de overlap $s' := |\langle\psi_1'|\psi_2'\rangle|$ for given initial states $|\psi_i\rangle$, subject to the condition that the average probability Q does not exceed some given value, $Q_{\max}$. In this case we could further assume that either the initial overlap $s := |\langle\psi_1|\psi_2\rangle|$ is given, in which case one can compute the tradeoff curve $s'_{\min}(Q_{\max})$, or else assume that $Q_{\max}$ is fixed and compute the curve $s'_{\min}(s)$. It is easy to see that $s'_{\min}(Q_{\max})$ and $Q_{\max}(s'_{\min})$ are just inverses of each otrher.

Whether we approach optimality one way or another depends merely on the problem at hand. Hence, e.g., for perfect cloning from one initial copies of either $|\psi_1\rangle$ or $|\psi_2\rangle$ to n

final copies (i.e., $|\psi_i'\rangle = |\psi_i\rangle^{\otimes n}$), the former approach is most suitable since the final overlap is fixed, $s' = s^n$, and so is the degree of separation attained by the cloner. So, in [?] the solution was given in terms of $Q_{\min}$ as a function of the prior probability η_1. However, one may need to know what is the maximum number of clones that can be produced if the failure rate cannot exceed $Q_{\max}$, in which case one takes the latter approach, and compute $n_{\max} = \log[s'(Q_{\max})]/\log s$.

The machine that carries the probabilistic transformation is usually described by two Kraus operators A_{succ}, A_{fail}, so that $A_{\text{succ}}^\dagger A_{\text{succ}} + A_{\text{fail}}^\dagger A_{\text{fail}} = [?]$. We can think of A_{succ} and A_{fail} as measurement operators. The transformation is successfully applied if the outcome of such (generalized) measurement is "succ", and fails otherwise. Neumark's theorem provides an alternative approach that turns out to be more convenient for our analysis. Additional details on this method can be found in [?]. In this formulation, the Hilbert space $\mathcal{H}$ of the original states is supplemented with an ancillary space $\mathcal{H}_{\text{extra}} \otimes \mathcal{H}_F$ that accommodates both the required extra-dimensions (if necessary) as well as the success/failure flags. Then, a unitary transformation U (time evolution) from $\mathcal{H} \otimes \mathcal{H}_{\text{extra}} \otimes \mathcal{H}_F$ onto $\mathcal{H}' \otimes \mathcal{H}_F$ is defined through [?]

$$U|\psi_1\rangle|0\rangle = \sqrt{p_1}|\psi_1'\rangle|\alpha_1\rangle + \sqrt{q}_1|\phi\rangle|\alpha_0\rangle, \tag{5.0.15}$$

$$U|\psi_2\rangle|0\rangle = \sqrt{p_2}|\psi_2'\rangle|\alpha_2\rangle + \sqrt{q}_2|\phi\rangle|\alpha_0\rangle. \tag{5.0.16}$$

Here the ancillas are initialized in a reference state $|0\rangle$. The states of the flag associated with successful transformation $|\alpha_i\rangle$ are constrained to be orthogonal to the state $|\alpha_0\rangle$ that signals failure. Upon performing a projective measurement on the flag space $\mathcal{H}_F$, the final state delivered through the outport port of our probabilistic machine is either $|\psi_i'\rangle$, in case of success, or $|\phi\rangle$ in case of failure. So, the outcome of this measurement tells us if the machine has succeeded or failed in delivering the right transformed state. On general grounds, optimality requires $|\alpha_1\rangle = |\alpha_2\rangle$. Here we choose to consider a more general setup where these two states are different to include state discrimination, for which the success flag states must be fully distinguishable, so $\langle\alpha_1|\alpha_2\rangle = 0$. Likewise, we could consider a more

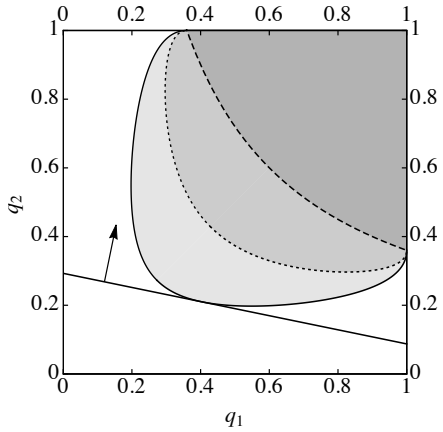

FIGURE 5.0.1. Unitarity curves in Eq. (5.0.17) and the associated sets S_β in Eq. (5.0.18) for $\beta = 0.45$ (solid/light gray), $\beta = 0.30$ (dotted/medium gray), and $\beta = 0$ (dashed/dark gray). The figure also shows the optimal straight segment $Q = \eta_1 q_1 + \eta_2 q_2$ and its normal vector (η_1, η_2). Plotted for $s = 0.6$, $\eta_1 = 0.17$, $\eta_2 = 0.83$ and $Q = 0.24$.

general setup with two failure states $|\phi_1\rangle$ and $|\phi_2\rangle$ in Eqs. (5.0.15) and (5.0.16). This is necessarily sub-optimal since we could probabilistically determine whether we received $|\psi_1\rangle$ or $|\psi_2\rangle$ by applying unambiguous discrimination to the failure states $|\phi_i\rangle$. Sometimes we would be certain of the input state, in which case we could prepare $|\psi_1'\rangle$ or $|\psi_2'\rangle$ accordingly, thereby increasing the overall success rate.

Taking the inner product of Eqs. (5.0.15) and (5.0.16) with themselves shows that our probabilities are normalized: $p_i + q_i = 1$. Similarly, by taking the product of Eq. (5.0.15) with Eq. (5.0.16), we find the unitarity constraint,

$$s = \sqrt{p_1 p_2}\,\beta + \sqrt{q_1 q_2}, \tag{5.0.17}$$

where $\beta = s'|\langle\alpha_1|\alpha_2\rangle|$. Without any loss of generality, in deriving Eq. (5.0.17) we have chosen $\langle\psi_1|\psi_2\rangle$, $\langle\psi_1'|\psi_2'\rangle$ and $\langle\alpha_1|\alpha_2\rangle$ to be real and positive. We note that $0 \le \beta \le s$, and $\beta = 0$ for both full separation ($s' = 0$) and unambiguous discrimination ($|\langle\alpha_1|\alpha_2\rangle| = 0$), whereas for optimal separation $|\langle\alpha_1|\alpha_2\rangle| = 1$. If Eq. (5.0.17) is satisfied, it is not hard to prove that U has a unitary extension on the whole Hilbert space and the Kraus operators, A_{succ}, A_{fail}, can be obtained by tracing out the ancillary degrees of freedom .

For convenient referencing, we gather in a lemma all the features of the unitary constraint that we will need. Points (a)–(d) are straightforward, so only (e) and (f) are proven below.

LEMMA 1. *(a) For fixed s, Eq. (5.0.17) defines a class of smooth curves on the unit square $0 < q_i < 1$ (e.g., solid, dashed or dotted curves in Fig. 5.0.1). (b) All these curves meet at their endpoints, $(1, s^2)$ and $(s^2, 1)$. (c) At the endpoints the curves become tangent to the vertical and horizontal lines $q_1 = 1$ and $q_2 = 1$ respectively, provided β is not zero. (d) For $\beta = 0$ the curve is an arc of the hyperbola $q_1 q_2 = s^2$ (dashed line in Fig. 5.0.1). (e) Each of these curves and the segments joining their end points with the vertex $(1, 1)$ enclose the sets (any of the gray regions in Fig. 5.0.1)*

$$S_\beta = \{(q_1, q_2) \in [0,1] \times [0,1] : \sqrt{p_1 p_2}\,\beta + \sqrt{q_1 q_2} - s \geq 0\}. \tag{5.0.18}$$

They satisfy $S_\beta \subset S_{\beta'}$ if $\beta < \beta'$. (f) Moreover, the sets S_β are convex.

Proof. (e) The curve (5.0.17) is readily seen to be part of the boundary of S_β. Assume that $\beta \leq \beta'$ and $(q_1, q_2) \in S_\beta$. Then

$$\sqrt{p_1 p_2}\,\beta' + \sqrt{q_1 q_2} - s \geq \sqrt{p_1 p_2}\,\beta + \sqrt{q_1 q_2} - s \geq 0, \tag{5.0.19}$$

and thus $(q_1, q_2) \in S_{\beta'}$. (f) To prove convexity let us assume that (q_1, q_2) and (q_1', q_2') belong to S_β. We define $\bar{q}_i = \lambda q_i + (1-\lambda) q_i'$, where $0 \leq \lambda \leq 1$. It follows that $\bar{p}_i := 1 - \bar{q}_i = \lambda p_i + (1-\lambda) p_i'$. Since $f(x, y) = \sqrt{xy}$ is a concave function in the unit square $\{(x, y) \,|\, 0 \leq x, y \leq 1\}$, we have $\sqrt{\bar{q}_1 \bar{q}_2} \geq \lambda \sqrt{q_1 q_2} + (1 - \lambda)\sqrt{q_1' q_2'}$ and, since $\beta \geq 0$, $\sqrt{\bar{p}_1 \bar{p}_2}\,\beta \geq \lambda \sqrt{p_1 p_2}\,\beta + (1 - \lambda)\sqrt{p_1' p_2'}\,\beta$. Then

$$\sqrt{\bar{p}_1 \bar{p}_2}\,\beta + \sqrt{\bar{q}_1 \bar{q}_2} - s \geq \lambda\,(\sqrt{p_1 p_2}\beta + \sqrt{q_1 q_2} - s) + (1 - \lambda)\left(\sqrt{p_1' p_2'}\beta + \sqrt{q_1' q_2'} - s\right) \geq 0. \tag{5.0.20}$$

Thus, $(\bar{q}_1, \bar{q}_2) \in S_\beta$, which proves the convexity of S_β for $\beta \geq 0$. ■

Now that we have characterized the geometry of the unitarity constraint, a geometrical picture of the optimization problem emerges (See Fig. 5.0.1). Eq. (5.0.14) defines a straight segment on the square $0 \leq q_i \leq 1$ with a normal vector in the first quadrant parallel to (η_1, η_2). For fixed *a priori* probabilities, the average failure probability Q is proportional to the distance from this segment to the origin $(0, 0)$. The intersection of such a straight segment with the boundary of S_β provides an admissible unitary transformation U and its corresponding failure probability Q. Since S_β is convex and the stretch of its boundary given by Eq. (5.0.17) is smooth, the optimal transformation, for which Q is minimal, is defined by the unique point (q_1, q_2) of tangency with the segment (5.0.14) that exists for any value of the priors and for $\beta > 0$. So, this tangency point determines the minimum failure probability $Q_{\min}$ and defines the optimal separation strategy through Eqs. (5.0.15) and (5.0.16).

For $\beta = 0$ (full separation/unambiguous discrimination), the right hand side of Eq. (5.0.17) describes hyperbolas that we can write as $q_2 = s^2/q_1$ (dashed lines in Figs. 5.0.1). Their slopes, $q_2' = -s^2/q_1^2$, are in the range $[-s^{-2}, -s^2]$. A unique point of tangency with the line (5.0.14) can only exists if the slope of this line, $-\eta_1/\eta_2$, is within this same range, namely if $s^2/(1+s^2) \leq \eta_1 \leq 1/(1+s^2)$. The tangency point is then seen to be $(q_1, q_2) = \sqrt{\eta_1\eta_2}\, s\, (\eta_1^{-1}, \eta_2^{-1})$. This leads to an minimum average failure probability given by $Q_{\min} = Q_{\text{UD}} := 2\sqrt{\eta_1\eta_2}s$, where the subscript UD stands for unambiguous discrimination. If the slope is outside the range tangency is not possible, and then the optimal line merely touches the end points of the hyperbolas. For $\eta_1 < s^2/(1+s^2)$, the straight segment (5.0.14) pivots on the lower end point, $(1, s^2)$, as we vary η_1 and we have the minimum average failure probability as $Q_{\text{UD}} = \eta_1 + \eta_2 s^2$. Likewise, for $\eta_1 > 1/(1+s^2)$, the pivoting point is the upper end point of the hyperbola, $(s^2, 1)$, which leads to $Q_{\text{UD}} = \eta_1 s^2 + \eta_2$. The above can be summarized by

$$Q_{\text{UD}} = \begin{cases} 2\sqrt{\eta_1\eta_2}\, s, & \dfrac{s^2}{1+s^2} \leq \eta_1 \leq \dfrac{1}{1+s^2}; \\ \eta_1 + s^2\eta_2, & 0 \leq \eta_1 \leq \dfrac{s^2}{1+s^2}; \\ \eta_1 s^2 + \eta_2, & \dfrac{1}{1+s^2} \leq \eta_1 \leq 1. \end{cases} \tag{5.0.21}$$

We immediately recognize that this expression is the average failure probability of unambiguously discriminating the input states $|\psi_1\rangle$ and $|\psi_2\rangle$, as it was anticipated.

Furthermore, we note that for the second (third) line in (5.0.21) we have $p_1 = 1 - q_1 = 0$ ($p_2 = 1 - q_2 = 0$), which leads to a 2-outcome projective measurement, as only the success flag state $|\alpha_2\rangle$ ($|\alpha_1\rangle$) is needed in Eqs. (5.0.15) and (5.0.16). The solution in the first line of Eq. (5.0.21) is manifestly symmetric under the exchange of the input states, i.e., under $\eta_1 \leftrightarrow \eta_2$. However, this symmetry is lost in the other lines. Instead, the effect of swapping the states turns the solution in the second line of Eq. (5.0.21) into the solution in the third line. One can also check that $Q_{\rm UD}$ is a twice differentiable function of η_1 (or η_2), with discontinuous second derivative at $\eta_1 = s^2/(1+s^2)$ and $\eta_1 = 1/(1+s^2)$. Our geometrical approach shows that the average failure probability $Q_{\rm min}$ is an infinitely differentiable function of η_1 for $\beta > 0$, since according to our lemma, the boundary curve (5.0.17) merges smoothly into the lines $q_1 = 1$ and $q_2 = 1$. So, it turns out that at $\beta = 0$ a phenomenon similar to a second order symmetry breaking phase transition takes place. A similar phenomenon was observed in unambiguous discrimination of more than two pure states [?].

Our lemma can likewise be used to address optimality for given priors η_1 and η_2 and average failure probability not exceeding $Q_{\rm max}$, with $0 \leq Q_{\rm max} < Q_{\rm UD}$. First, since the unitary curve is a function of $\beta = s'\langle\alpha_1|\alpha_2\rangle$, we set $|\alpha_1\rangle = |\alpha_2\rangle$, i.e., $\langle\alpha_1|\alpha_2\rangle = 1$, to ensure the minimum value of $s' > 0$ for a given β. Then, it follows from the lemma that the minimum final overlap $s' > 0$ (the maximum degree of separation attainable), which we call $s'_{\rm min}$, is that for which the segment (5.0.14), with $Q = Q_{\rm max}$, and the boundary of $S_{\beta=s'}$ become tangent. Setting the margin $Q_{\rm max}$ in the range $[Q_{\rm UD}, 1]$ leads, obviously, to the trivial solution $s'_{\rm min} = 0$, for such margin would allow full separation using unambiguous discrimination with a failure rate of exactly $Q = Q_{\rm UD}$, below the given margin.

In summary, our lemma provides the solution to optimal state separation from a geometrical viewpoint by showing that it is a convex optimization problem, for which a unique solution exists. Unfortunately, a closed form for this solution does not exist for arbitrary prior probabilities, since finding the tangency point of the segment in Eq. (5.0.14) with the

curve in Eq. (5.0.17) requires solving a six degree polynomial equation, as one can easily check. In the next sections, we give an analytic solution to state separation in parametric form. This solution contains all the information one may need in a simple and straightforward fashion. In particular, it enables us to easily draw plots of the relevant quantities for the various cases we will consider.

5.1. Minimum failure probability for a fixed degree of separation

When the overlap of the final states is fixed, as in perfect cloning, we argued above that a natural problem consists in deriving the minimum failure rate of the optimal protocol, $Q_{\min}$, as a function of one of the priors, say η_1. In this section we address this problem by following the method employed in our derivation for cloning in [?]. All the expressions below can be obtained from their analogs in [?] with the simple replacements $s^m \to s$ and $s^n \to s'$, starting with the symmetric parametrization of the curve (5.0.17). Its lower half (for which $q_2 \leq q_1$) is parametrized as

$$q_i = \frac{1 - xy - (-1)^i \sqrt{1-x^2}\sqrt{1-y^2}}{2}, \quad i = 1, 2, \tag{5.1.1}$$

where

$$x = \frac{1 - (1+s')t}{s'/s}, \qquad y = \frac{1 - (1-s')t}{s'/s}. \tag{5.1.2}$$

This parametrization arises from a change of variables that linearizes the unitary constraint, which proved very convenient in [?], where the advantages of its highly symmetric form were also apparent. The upper half of the curve (5.0.17) can be obtained by applying the transformation $q_1 \leftrightarrow q_2$. However, without any loss of generality, we can assume that $0 \leq \eta_1 \leq 1/2$ (thus, $1/2 \leq \eta_2 \leq 1$), so only the lower half given by Eq. (5.1.1) can actually become tangent to the straight segment in Eq. (5.0.14).

Fig. 5.1.1 (a) shows plots of the unitarity curve [Eq. (5.1.1) plus the reflection $q_1 \leftrightarrow q_2$] for $s = 0.6$ and $s' = 0.05$, 0.3, 0.5, 0.59. For $s' = 0.59$, very close to the value of s (small separation), the vertex of the curve approaches the origin, which becomes a singular point in the limit $s' \to s$. As s' decreases (increasing separation), the curves approach the hyperbola

$q_1 q_2 = s^2$. It is apparent from the figure that the curves merge smoothly onto the lines $q_1 = 1$ and $q_2 = 1$ for the larger values of s'. It becomes less obvious for small values of s', such as $s' = 0.05$. However a blowup of Fig. 5.1.1 (a) would reveal that this is so. A cusp at $(s^2, 1)$ and $(1, s^2)$ arises only for $s' \to 0$.

It follows from our lemma, and it can be checked using Eq. (5.1.1), that the slope of the lower half of the unitarity curve increases monotonically as we move away from the line $q_1 = q_2$, where it has the value -1, and vanishes before we reach the line $q_1 = 1$. The values of t at which the slope is -1 and 0 are, respectively,

$$t_{-1} = \frac{1 - s'/s}{1 - s'}, \quad t_0 = \frac{1 - s'^2/s^2}{1 - s'^2}. \tag{5.1.3}$$

So, there is a straight segment (5.0.14), with slope $-\eta_1/\eta_2$, that is tangent to each point $(q_1(t), q_2(t))$, $t \in [t_{-1}, t_0]$, of the unitarity curve parametrized by Eq. (5.1.1). Since the slope of this curve is $q_2'(t)/q_1'(t)$, where the prime stands for derivative with respect to t, a parametric expression for η_1 can be obtained from the equal slope condition $-\eta_1/\eta_2 = q_2'(t)/q_1'(t)$. The parametric expression for $Q_{\min}$ follows from imposing that $(q_1(t), q_2(t))$ must be a point of the straight segment (5.0.14), so $Q_{\min} = \eta_1 q_1(t) + \eta_2 q_2(t)$. The final result can be cast as

$$\eta_1 = \frac{q_2'}{q_2' - q_1'}, \quad Q_{\min} = \frac{q_2' q_1 - q_1' q_2}{q_2' - q_1'}, \quad t_{-1} \le t \le t_0, \tag{5.1.4}$$

where we have dropped the argument of $q_i(t)$ and $q_i'(t)$ to simplify the equation. Further, one can check that the derivatives of $q_i(t)$ can be written as

$$q_i' = \frac{\sqrt{q_i(1 - q_i)}}{s'/s} \left\{ \frac{1 + s'}{\sqrt{1 - x^2}} - (-1)^i \frac{1 - s'}{\sqrt{1 - y^2}} \right\}. \tag{5.1.5}$$

Eq. (5.1.4) gives $Q_{\min}(\eta_1)$ in parametric form for $0 < s' < s$. The solution for $s' = 0$ was already derived in the previous section and for $s' = s$ we have the trivial solution $Q_{\min} = 0$. These special cases can also be derived from Eq. (5.1.4) by carefully taking the corresponding limits. The values of $Q_{\min}$ at the end points of this range follow by substituting t_0 and t_{-1},

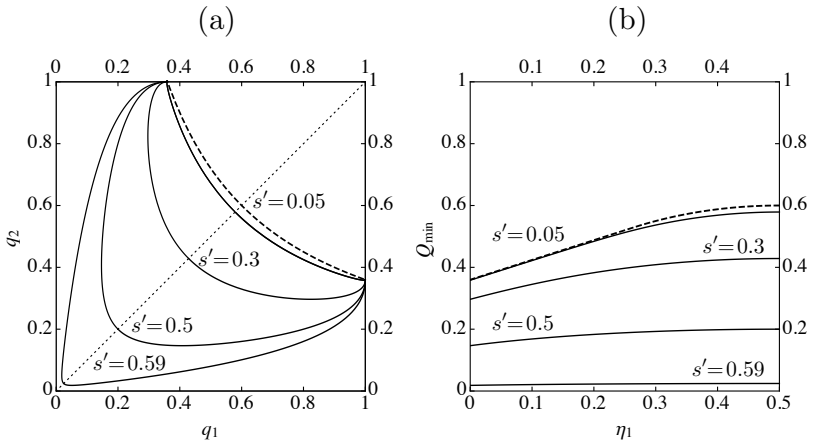

FIGURE 5.1.1. (a) Unitarity curves for different values of s'. The curves are symmetric under mirror reflexion along the (dotted) straight line $q_1 = q_2$, i.e., under the transformation $q_1 \leftrightarrow q_2$. (b) Minimum separation failure probability $Q_{\min}$ vs. η_1 (solid lines), for the same values of s' used in (a). The dashed lines corresponds to full separation/unambiguous discrimination ($s' = 0$) in both figures.

Eq. (5.1.3), into Eq. (5.1.1). They are given by

$$Q_0 = q_2(t_0) = \frac{s^2 - s'^2}{1 - s'^2}, \quad Q_{-1} = \frac{s - s'}{1 - s'}, \tag{5.1.6}$$

where $Q_{\min} = Q_{-1}$ holds for equal priors and $Q_{\min} = Q_0$ for $\eta_1 \to 0$ (i.e., $\eta_2 \to 1$).

Fig. 5.1.1 (b) shows plots of the curves $(\eta_1, Q_{\min})$ for the same values of s and s' as the ones given above (solid lines). We see that $Q_{\min}$ is an increasing function of η_1 in the given range $[0, 1/2]$, as one should expect. The figure also shows the failure rate for unambiguous discrimination (dashed line), which coincides with $Q_{\min}$ for $s' = 0$. From the plots, it is clear that $Q_{\min}$ is a decreasing function of s', again as it should be.

5.2. State Separation: unequal priors

5.3. Maximum separation

In this section, we assume η_1, η_2 are fixed given quantities and we focus on the relationship among the initial overlap, the final overlap and the maximum allowed failure rate. To find the explicit form of these relationships, we will need to develop a new geometric view of both the unitarity constraint, Eq. (5.0.17), and $Q = \eta_1 q_1 + \eta_2 q_2$. We aim at a geometric representation

simple enough to grasp visually the solution and yet powerful enough to provide this solution analytically. We show below that the unitary curve and the straight segment of the previous sections can be mapped into conic curves, in particular into families of parabolas and ellipses respectively. This is arguably the simplest extension to our geometric description of state separation. The desired transformation is defined in terms of the new variables u and v as

$$u = \sqrt{q_1 q_2}; \quad v = \frac{q_1 + q_2}{2}. \tag{5.3.1}$$

They are just the geometric and arithmetic means of the failure probabilities, q_1 and q_2. Under this transformation the unitary constraint becomes a parabola that can be conveniently written as

$$v = \frac{1+u^2}{2} - \frac{(u-s)^2}{2s'^2}. \tag{5.3.2}$$

From this expression, one can immediately check that as s varies we obtain a family of parabolas whose envelope is yet another parabola, $v = (1+u^2)/2$, independently of s'. As s' decreases from its maximum value $s' = s$, the parabolas in Eq. (5.3.2) become thinner. For $s' = 0$ they degenerate into the vertical segment $u = s$, $0 \leq v \leq (1+s^2)/2$. These features are illustrated in Fig. 5.3.1 (a).

Under the same transformation, Eq. (5.3.1), the line $Q = \eta_1 q_1 + \eta_2 q_2$ becomes an ellipse, which is most easily expressed parametrically in terms of the polar angle θ, measured relative to the axis $v = 0$ from the center of the ellipse. It is given by

$$\begin{aligned} u &= \frac{Q}{\sqrt{1-\Delta^2}} \cos\theta, \\ v &= \frac{Q}{1-\Delta^2} + \frac{Q\Delta}{1-\Delta^2} \sin\theta, \end{aligned} \tag{5.3.3}$$

where we have defined $\Delta = \eta_2 - \eta_1$. It is clear from this expression that the eccentricity of the ellipse is only a function of the priors. For equal priors, $\Delta = 0$, the ellipse degenerates into the horizontal segment $v = Q$, $0 \leq u \leq Q$, whereas for $Q = 0$ it collapses into the origin $(u, v) = (0, 0)$. As one increases Q, a family of similar ellipses is obtained. As they increase in size, their center moves up along the axis $u = 0$. The line $u = v$ is the envelope of this

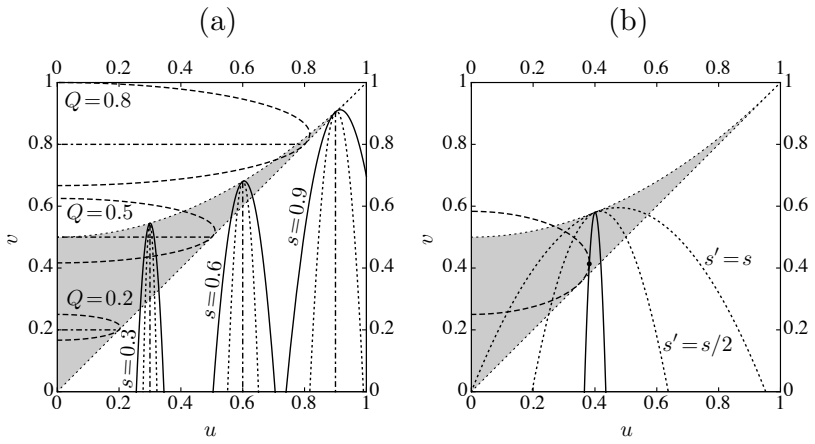

FIGURE 5.3.1. (a) Unitarity parabolas, Eq. (5.3.2), for different values of s, $s' = s/7$ (solid lines) and $s' = s/14$ (dotted lines). The dashed lines are the ellipses in Eq. (5.3.3) for various values of the failure rate Q. The top boundary line to the gray region, given by $v = (1 + u^2)/2$, is the envelope of the solid and dotted parabolas. The bottom boundary line, i.e., the straight line $v = u$, is the envelope of the family of ellipses (dashed lines). The geometric solution to optimal separation falls in the gray region. In this figure $\eta_1 = 0.4$. The degenerate curves for $s' = 0$ (dot-dashed vertical line) and $\Delta = 0$ (dot-dashed horizontal line) are also shown. (b) Optimal (solid) and suboptimal (dotted) parabolas. The tangency point is also displayed. In this figure $\eta_1 = 0.3$, $s = 0.4$ and $Q = Q_{\max} = 0.35$. The optimal (minimum) value of s', which gives the solid parabola, turns out to be $s' = 0.032$.

family, as one can easily check using Eq. (5.3.3). Fig. 5.3.1 (a) also illustrates these features.

In terms of this conic geometry, optimality is again given by a tangency point, this time between ellipses and parabolas. Because of the features of these families of conics, these points of tangency necessarily lie in the region between their envelopes, which is the gray area in Fig. 5.3.1. Fig. 5.3.1 (b) illustrates optimality. Given a maximum failure rate $Q_{\max}$ and some initial overlap s ($Q_{\max} = 0.35$ and $s = 0.4$ in the example considered in the figure), we plot the corresponding ellipse defined by Eq. (5.3.3) (dashed line). Among the various parabolas, characterized by the final overlap s' (the figure shows two of them, for $s' = s$ and $s' = s/2$), the one that minimizes s' (solid line) has a unique point of tangency with the ellipse, thus giving us the solution, $s'_{\min}$. To keep the notation simple we will drop the subscript "min" wherever no confusion arises.

To find the condition that gives the tangency point, we first note that the slopes of the ellipse and the parabolas are given respectively by

$$\begin{aligned} \frac{dv}{du} &= \frac{v'}{u'} = -\frac{\Delta}{\sqrt{1-\Delta^2}}\cot\theta, \\ \frac{dv}{du} &= u - \frac{u-s}{s'^2}, \end{aligned} \tag{5.3.4}$$

where the prime stands for derivative with respect to the polar angle θ. The right hand side of these two equations must be equal at the tangency point. Moreover, the tangency point must belong to both the ellipse and the optimal parabola. Hence

$$\begin{aligned} Q\frac{1+\Delta\sin\theta}{1-\Delta^2} &= \frac{1}{2}+\frac{Q^2\cos^2\theta}{2(1-\Delta^2)}-\frac{1}{2s'^2}\left(\frac{Q\cos\theta}{\sqrt{1-\Delta^2}}-s\right)^2, \\ \frac{\Delta\cot\theta}{\sqrt{1-\Delta^2}} &= \frac{1-s'^2}{s'^2}\frac{Q\cos\theta}{\sqrt{1-\Delta^2}}-\frac{s}{s'^2}. \end{aligned} \tag{5.3.5}$$

where to obtain the first (second) equation we have simply substituted Eq. (5.3.3) into Eq. (5.3.2) [Eq. (5.3.4)]. Ideally, we would like to solve this system of equations by eliminating θ, which would lead to a closed expression relating s, s' and Q. Unfortunately, this involves solving a high degree polynomial equation in $\cos\theta$. Instead, we look at it as a system of two equations with two unknowns, s and s' (or Q and s') and keep θ as a parameter describing the curve (s,s') [or (Q,s')] in parametric form. After some algebra, we obtain the simple expressions:

$$s' = -\frac{\sqrt{(1-Q)^2-(\Delta+Q\sin\theta)^2}}{\Delta+Q\sin\theta}\tan\theta, \tag{5.3.6}$$

$$s = \frac{Q\Delta(1+\sin^2\theta)-(1-\Delta^2-2Q)\sin\theta}{\sqrt{1-\Delta^2}(\Delta+Q\sin\theta)\cos\theta}. \tag{5.3.7}$$

The range of values of the parameter θ in this equation is $-\arcsin\Delta \le \theta \le \theta_{\max}$, where

$$\theta_{\max} = \begin{cases} 0 & \text{if} \quad Q \le 1-\Delta, \\ \arcsin\dfrac{1-Q-\Delta}{Q} & \text{if} \quad Q \ge 1-\Delta. \end{cases} \tag{5.3.8}$$

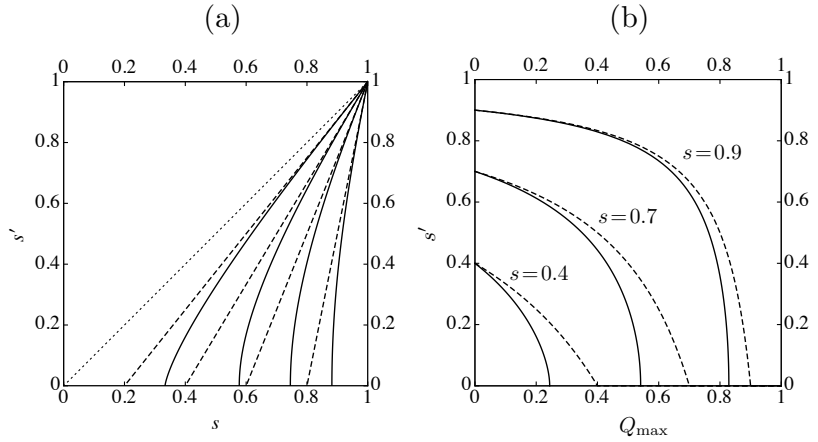

FIGURE 5.3.2. (a) Plots of s' vs. s for $\eta_1 = 0.1$ (solid lines) and $\eta_1 = 0.5$ (straight dashed lines) and for values of the failure rate. From left to right $Q_{\max} = 0.2$, 0.4, 0.6, 0.8. The dotted line is the (trivial) curve for $Q_{\max} = 0$, which is the straight line $s' = s$. (b) Minimum final overlap vs. maximum failure probability for various values of the initial overlap and the same two values of η_1 used in (a).

One can easily check the given minimum value of θ by substituting in Eqs. (5.3.6) and (5.3.7) to obtain $s' = s = 1$, as it should be. Likewise, one can check that for $\theta = \theta_{\max}$ one has $s' = 0$. The two cases in Eq. (5.3.8) reveal the appearance of the phase transition in the limit $s' \to 0$ that we discussed in previous sections. If $Q \geq 1 - \Delta$, substituting the second line of Eq. (5.3.8) in Eq. (5.3.5) we obtain $s = [(2Q + \Delta - 1)/(1 + \Delta)]^{1/2}$. Solving for Q, we find that $Q = \eta_1 + s^2\eta_2$. This means that the condition $Q \geq 1 - \Delta$ is equivalent to $\eta_1 + s^2\eta_2 \geq 1 - \Delta$, which can be immediately seen to give $\eta_1 \leq s^2/(1 + s^2)$. So we obtain the second line in Eq. (5.0.21), corresponding to the "symmetry-broken phase". If $Q \leq 1 - \Delta$, namely, if $s^2/(1 + s^2) \leq \eta_1$, we have instead $s = Q/\sqrt{1 - \Delta^2}$. This equation can be written as $Q = 2\sqrt{\eta_1\eta_2}s$. So, Eq. (5.3.8) has the same content as Eq. (5.0.21). Recall that we are assuming $\eta_1 \leq 1/2 \leq 1/(1 + s^2)$. The third line in Eq. (5.0.21) never applies under this assumption.

Eqs. (5.3.6) and (5.3.7) are plotted in Fig. 5.3.2 (a) for two possible priors: $\eta_1 = 0.1$ (solid lines) and $\eta_1 = 0.5$, i.e., for equal priors (dashed lines). From left to right, the maximum allowed failure rate $Q_{\max}$ is 0.2, 0.4, 0.6 and 0.8. We see that for small values of the initial

overlap, s, one can attain full separation ($s' = 0$). Past the critical value,

$$s_{\rm cr} = \begin{cases} \dfrac{Q_{\max}}{2\sqrt{\eta_1 \eta_2}} & \text{if} \quad Q \leq 2\eta_1, \\ \sqrt{\dfrac{Q_{\max} - \eta_1}{\eta_2}} & \text{if} \quad Q \geq 2\eta_1, \end{cases} \tag{5.3.9}$$

full separation is no longer possible and s' increases (quite abruptly for small η_1). In the region $s < s_{\rm cr}$, the margin $Q_{\max}$ is not necessarily saturated, since the unambiguous discrimination failure probability $Q_{\rm UD}$ is smaller than $Q_{\max}$. For $s \geq s_{\rm cr}$ we necessarily have to saturate the margin, $Q = Q_{\max}$. For equal priors (dashed lines) one can obtain the curves in explicit form from Eq. (5.0.17) using that $q_1 = q_2 = Q$:

$$s' = \begin{cases} 0 & \text{if} \quad s \leq Q_{\max}, \\ \dfrac{s - Q_{\max}}{1 - Q_{\max}} & \text{if} \quad s \geq Q_{\max}. \end{cases} \tag{5.3.10}$$

This expression could also be obtained by carefully taking the limit $\Delta \to 0$ in Eqs. (5.3.6) through (5.3.8). The figure clearly shows that separation becomes less demanding as we move away from the equal prior case. For $Q_{\max} = 0$, i.e., in the deterministic limit, we recover the trivial solution $s' = s$ (dotted line).

5.4. Tradeoff between Maximum separation and failure rate

By solving the system Eq. (5.3.5) for Q and s' , we obtain a parametric expression for the tradeoff curve (Q, s') in terms of the polar angle θ:

$$\begin{aligned} s'^2 &= \sqrt{1-\Delta^2}\left(\frac{\sin\theta}{\Delta+\sin\theta}\right)^2 \\ &\times \frac{\sqrt{1-\Delta^2}(1+s^2)\cos\theta - 2s(1+\Delta\sin\theta)}{\cos\theta}, \end{aligned} \tag{5.4.1}$$

$$Q = \frac{s\sqrt{1-\Delta^2}+\Delta\, s'^2\cot\theta}{(1-s'^2)\cos\theta}. \tag{5.4.2}$$

Note that Eq. (5.4.1) is an expression for the square of the final overlap. To keep the formula for Q, Eq. (5.4.2), short, we use s'^2 as a shorthand for Eq. (5.4.1). The range of θ in Eqs. (5.4.1) and (5.4.2) is:

$$-\arctan\frac{s\Delta}{\sqrt{1-\Delta^2}} \le \theta \le \theta_{\max},$$

where the upper limit of the interval can be written as

$$\theta_{\max} = \begin{cases} 0 & \text{if} \quad \eta_1 \ge \dfrac{s^2}{1+s^2}, \\ -\arccos\dfrac{2s\sqrt{1-\Delta^2}}{1-\Delta+s^2(1+\Delta)} & \text{if} \quad \eta_1 \le \dfrac{s^2}{1+s^2}. \end{cases} \tag{5.4.3}$$

The lower limit in the range of allowed θ can be derived from Eqs. (5.4.1) and (5.4.2) by imposing that $Q = 0$ at $s' = s$. The upper limit can be derived from Eq. (5.4.1) by imposing $s' = 0$. Once again, we see that a second order phase transition occurs in the limit of full separation: by substituting the first (second) line of Eq. (5.4.3) in Eq. (5.4.2) we obtain $Q = s\sqrt{1-\Delta^2}$ ($Q = [1-\Delta+s^2(1+\Delta)]/2$), which is the first (second) case in Eq. (5.0.21).

Fig. 5.3.2 (b) shows various plots of the separation vs. failure-rate tradeoff curve. As in Fig. 5.3.1, the plots are for $\eta_1 = 0.1$ (solid lines) and for equal priors, $\eta_1 = \eta_2 = 0.5$ (dashed lines). For equal priors, there is the explicit formula for the curves given in Eq. (5.3.10). Again, we see that as η_1 gets smaller, departing from the equal prior value 1/2, the states

can be separated more for the same maximum rate of failure. As $Q_{\max}$ increases, the minimum overlap gets smaller, as it should. When the margin $Q_{\max}$ reaches the unambiguous discrimination value Q_{UD} we have $s' = 0$, attaining full separation. Larger values of $Q_{\max}$ are rather meaningless in this context, since they will never be saturated by an optimal protocol, which requires a failure rate of only $Q = Q_{\mathrm{UD}}$ ($< Q_{\max}$) to fully separate the input states.

CHAPTER 6

Linear Optical Experimental Realizations

6.1. Reck-Zeilinger Algorithm and Single-Photon Interferometry

A significant motivation to solving information theory problems using Neumark's theorem is that solutions lend themselves to linear optical implementation. This implementation requires sources of individual photons, beamsplitters, mirrors, and photodetectors. Our representation of the

The three income ports, labeled 1, 2, 3 in the figure, are in a superposition of zero and/or one photons, corresponding to the orthogonal states $|0\rangle$, i.e., the vacuum state, $|1\rangle = a_1^\dagger|0\rangle$, $|2\rangle = a_2^\dagger|0\rangle$ and $|3\rangle = a_3^\dagger|0\rangle$, where $a_i^\dagger$ is the creation operator of the electromagnetic field in port i, $i = 1, 2, 3$. Similarly, for the output ports we have $|1'\rangle = a_1'^\dagger|0\rangle$, $|2'\rangle = a_2'^\dagger|0\rangle$ and $|3'\rangle = a_3'^\dagger|0\rangle$.

6.2. Implementation of Pure State Interpolative Discrimination

The main reason to seek a solution using the Neumark setup is because it lends itself into an optical implementation. This implementation, as we will see, can be carried out using only linear optical elements (beamsplitters and a mirror). The possible states are represented by single photons and several photodetectors will carry out the measurement process at the output.

We use a strategy similar to that developed by J.A Bergou *et al.* [**?**], and seek a unitary transformation that transforms the states as in (2.1) and (2.2), with the qubits in the states $|\psi_1\rangle = |1\rangle$, $|\psi_2\rangle = \cos\theta|1\rangle + \sin\theta|2\rangle$ and assume the ancilla space is empty for the initial

preparation, i.e., $\alpha_1 = \alpha_2 = 0$:

$$U|1\rangle = \sqrt{p_1}|1\rangle + \sqrt{r_1}|2\rangle + \sqrt{q_1}|3\rangle \tag{6.2.1}$$

$$U(\cos\theta|1\rangle + \sin\theta|2\rangle) = \sqrt{r_2}|1\rangle + \sqrt{p_2}|2\rangle + \sqrt{q_2}|3\rangle \tag{6.2.2}$$

From these two equations we can read out six of nine elements of the three by three Unitary matrix, e.g., $\langle 1|U|1\rangle = \sqrt{p_1}$. The rest can be calculated from the conditions of the unitarity, $U^T U = I$. They are, up to phase,

$$U = \begin{pmatrix} \sqrt{p_1} & \frac{\sqrt{r_2}-\sqrt{p_1}\cos\theta}{\sin\theta} & \pm\frac{\sqrt{\sin^2\theta - p_1 - r_2 + 2\sqrt{p_1 r_2}\cos\theta}}{\sin\theta} \\ \sqrt{r_1} & \frac{\sqrt{p_2}-\sqrt{r_1}\cos\theta}{\sin\theta} & \pm\frac{\sqrt{\sin^2\theta - r_1 - p_2 + 2\sqrt{p_2 r_1}\cos\theta}}{\sin\theta} \\ \sqrt{q_1} & \frac{\sqrt{q_2}-\sqrt{q_1}\cos\theta}{\sin\theta} & \pm\frac{\sqrt{sin^2\theta - q_1 - q_2 + 2\sqrt{q_1 q_2}\cos\theta}}{\sin\theta} \end{pmatrix}. \tag{6.2.3}$$

It is worth mentioning that all equations in this section referencing r_i and p_i are using the optimal values (2.9) and (2.10) derived in the previous section.

Now that we have a full unitary matrix we want to express it in terms of linear optical devices. M. Reck *et al.* [**?**], prove that any discrete finite-dimensional unitary operator can be constructed using optical devices. They derive an algorithm which gives the exact ordering of the beamsplitters and phase shifters. In our work we use the simplified version of the Reck algorithm given by Y. Sun *et al.* [**?**] : the operator U is decomposed into beamsplitters

in the order of $U = M_1 \cdot M_2 \cdot M_3$, and no phase shifters are needed:

$$M_1 = \begin{pmatrix} \sin\omega_1 & \cos\omega_1 & 0 \\ \cos\omega_1 & -\sin\omega_1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$M_2 = \begin{pmatrix} \sin\omega_2 & 0 & \cos\omega_2 \\ 0 & 1 & 0 \\ \cos\omega_2 & 0 & -\sin\omega_2 \end{pmatrix},$$

$$M_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sin\omega_3 & \cos\omega_3 \\ 0 & \cos\omega_3 & -\sin\omega_3 \end{pmatrix},$$

where the coefficients of reflectivity and transmittance are given by $\sqrt{R_i} = \sin\omega_i$ and $\sqrt{T_i} = \cos\omega_i$.

All of the beamsplitter coefficients can be derived up to a phase by using just U_{31}, U_{32}, U_{21}. The sign of the coefficients is chosen by matching all the elements from the two matrices. The coefficients are:

$$\cos\omega_1 = \sqrt{\frac{r_1}{1-q_1}},$$
$$\sin\omega_1 = \sqrt{\frac{p_1}{1-q_1}},$$
$$\cos\omega_2 = \sqrt{q_1},$$
$$\sin\omega_2 = \sqrt{1-q_1},$$
$$\cos\omega_3 = -\frac{1}{\sqrt{1-q_1}}[\frac{\sqrt{q_2}-\sqrt{q_1}\cos\theta}{\sin\theta}],$$
$$\sin\omega_3 = \frac{\sqrt{\sin^2\theta - q_1 - q_2 + 2\sqrt{q_1 q_2}\cos\theta}}{\sqrt{1-q_1}\sin\theta}.$$

All the terms can be expressed in terms of the fixed failure rate and fixed a-priori probabilities. Using the optimal relationship between the individual failure rates $\eta_1 q_1 = \eta_2 q_2 = Q/2$,$q_1 = Q/2\eta_1, q_2 = Q/2\eta_2$ and the optimal expressions of success and error rates the beamsplitters are

$$M_1 = \begin{pmatrix} \sqrt{\frac{p_1}{1-Q/2\eta_1}} & \sqrt{\frac{r_1}{1-Q/2\eta_1}} & 0 \\ \sqrt{\frac{r_1}{1-Q/2\eta_1}} & -\sqrt{\frac{p_1}{1-Q/2\eta_1}} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (6.2.4)$$

$$M_2 = \begin{pmatrix} \sqrt{1-Q/2\eta_1} & 0 & \sqrt{Q/2\eta_1} \\ 0 & 1 & 0 \\ \sqrt{Q/2\eta_1} & 0 & -\sqrt{1-Q/2\eta_1} \end{pmatrix}, \quad (6.2.5)$$

$$M_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{1-Q_o^2/4\eta_1\eta_2 - Q/(2\eta_1\eta_2) + QQ_o/(2\eta_1\eta_2)}}{\sqrt{(1-Q/2\eta_1)(1-Q_o^2/4\eta_1\eta_2)}} & -\frac{\sqrt{Q/2\eta_2}-Q_o/2\eta_1\sqrt{Q/2\eta_2}}{\sqrt{(1-Q/2\eta_1)(1-Q_o^2/4\eta_1\eta_2)}} \\ 0 & -\frac{\sqrt{Q/2\eta_2}-Q_o/2\eta_1\sqrt{Q/2\eta_2}}{\sqrt{(1-Q/2\eta_1)(1-Q_o^2/4\eta_1\eta_2)}} & -\frac{\sqrt{1-Q_o^2/4\eta_1\eta_2 - Q/(2\eta_1\eta_2) + QQ_o/(2\eta_1\eta_2)}}{\sqrt{(1-Q/2\eta_1)(1-Q_o^2/4\eta_1\eta_2)}} \end{pmatrix}. \quad (6.2.6)$$

An attractive simplification can be acheived by setting $\eta_1 = \eta_2$, the equal apriori condition. In this case our final unitary matrix can be expressed as

$$U = \begin{pmatrix} \sqrt{p} & \frac{\sqrt{r}-\sqrt{p}Q_o}{\sqrt{1-Q_o^2}} & \sqrt{\frac{Q}{1+Q_o}} \\ \sqrt{r} & \frac{[\sqrt{p}-\sqrt{r}Q_o]}{\sqrt{1-Q_o^2}} & \sqrt{\frac{Q}{1+Q_o}} \\ \sqrt{Q} & \sqrt{\frac{Q(1-Q_o)}{1+Q_o}} & -\frac{\sqrt{p}+\sqrt{r}}{\sqrt{1+Q_o}} \end{pmatrix} \tag{6.2.7}$$

By choosing the FRIO this matrix minimizes the error rate and maximizes the rate of success. Hence, by setting the FRIO to zero we obtain the setup to the minimum error problem on the other hand setting the error rate to zero gives the setup of the optimal unambiguous discrimination where the optimal inconclusive rate is the $Q_o = s$. This simplifies the works of the experimentalists because now they only need one setup and are not restrained to the extreme points.

The unitary transformation for the Helstrom bound $Q = 0$ implementation can be written using the relations $r = (\frac{\sqrt{r}-\sqrt{p}Q_o}{\sqrt{1-Q_o^2}})^2$, $p = (\frac{\sqrt{p}-\sqrt{r}Q_o}{\sqrt{1-Q_o^2}})^2$, as

$$U_{ME} = \begin{pmatrix} \sqrt{p} & \sqrt{r} & 0 \\ \sqrt{r} & -\sqrt{p} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{6.2.8}$$

Clearly only the M_1 beamsplitter is necessary to implement the ME state discrimination.

On the other extreme, the unitary transformation for the optimal UD bound ($P_E = 0$) implementation becomes:

$$U_{UD} = \begin{pmatrix} \sqrt{p} & -\frac{\sqrt{p}Q_o}{\sqrt{1-Q_o^2}} & \sqrt{\frac{Q_0}{1+Q_o}} \\ 0 & \frac{\sqrt{p}}{\sqrt{1-Q_o^2}} & \sqrt{\frac{Q_o}{1+Q_o}} \\ \sqrt{Q_0} & \sqrt{\frac{Q_o(1-Q_o)}{1+Q_o}} & -\sqrt{\frac{1-Q_o}{1+Q_o}} \end{pmatrix}. \tag{6.2.9}$$

All three beamsplitters are still necessary for a general UD measurement. This is because the measurement is essentially two-step: in the first step we attempt to orthogonalize the states, and upon succeeding we perform a projective measurement.

6.3. Implementation of State Separation

In this section we derive a similar implementation of optimal separation, sketched in Fig. 6.3.1.The implementation uses only two beam splitters, M1 and M2. The measurements are carried out by three photodetectors.

We can choose several different representations for our input states. The first is such that one state is parallel to a basis vector, allowing for the unitary transformation to be represented, using Eqs. (5.0.15) and (5.0.16), as

$$U|1\rangle = \sqrt{p_1}|1'\rangle + \sqrt{q_1}|3'\rangle, \tag{6.3.1}$$

$$U\left(s|1\rangle + \sqrt{1-s^2}|2\rangle\right) = \sqrt{p_1}\left(s'|1'\rangle + \sqrt{1-s'^2}|2'\rangle\right) + \sqrt{q_1}|3'\rangle, \tag{6.3.2}$$

which corresponds to the choice: $|\psi_1\rangle|0\rangle = |1\rangle$, $|\psi_2\rangle|0\rangle = s|1\rangle + \sqrt{1-s^2}|2\rangle$, $|\psi_1'\rangle|\alpha_1\rangle = |1'\rangle$, $|\psi_2'\rangle|\alpha_2\rangle = s'|1'\rangle + \sqrt{1-s'^2}|2'\rangle$ and $|\phi\rangle|\alpha_0\rangle = |3'\rangle$. The third rail always

So, input port 3 is always in the vacuum state. The detection of a photon in the output port $3'$ signals that separation failed. The state $|\psi_2\rangle$ can be produced in a standard way by sending a photon into a beam splitter with suitable transmission and reflection coefficients.

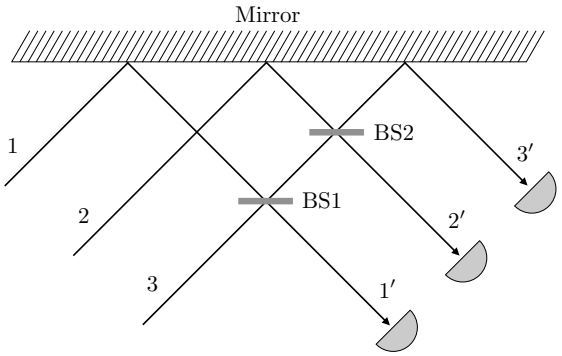

FIGURE 6.3.1. Six-port linear optics implementation of a proof-of-principle separation protocol. The transmission (reflexion) coefficients of the beam-splitters, BS1 and BS2 are given by the (off-)diagonal entries of the matrices in Eqs. (6.3.4) and (6.3.5), respectively. The input states are feed through ports 1 and 2 as a superposition of zero and one photons in each port. The separated states are output through ports $1'$ and $2'$. Port 3 is in the vacuum state. A click in the photodetector placed in port $3'$ signals failure.

For simplicity, we consider equal prior probabilities $\eta_1 = \eta_2 = 1/2$, but the same setup can be used in the general case. As mentioned above, for equal priors we must have $q_1 = q_2 = Q$ and $p_1 = p_2 = 1 - Q$ and the unitarity condition Eq. (5.0.17) can be solved explicitly. The solution is given by $Q = Q_{-1}$ in Eq. (5.1.6). Substituting in Eqs. (6.3.1) and (6.3.2) we obtain two columns of the matrix of the unitary transformation U in the basis introduced above. The remaining column can be easily obtained imposing unitarity. After some algebra we have

$$[U] = \begin{pmatrix} \sqrt{\frac{1-s}{1-s'}} & -\frac{s-s'}{\sqrt{(1-s')(1+s)}} & -\sqrt{\frac{(1+s')(s-s')}{(1-s')(1+s)}} \\ 0 & \sqrt{\frac{1+s'}{1+s}} & -\sqrt{\frac{s-s'}{1+s}} \\ \sqrt{\frac{s-s'}{1-s'}} & \sqrt{\frac{(1-s)(s-s')}{(1+s)(1-s')}} & \sqrt{\frac{(1-s)(1+s')}{(1-s')(1+s)}} \end{pmatrix}. \tag{6.3.3}$$

Using [**?, ?**] we can write U as the product $U = M_1 M_2$, where the matrices of M_1 and M_2 are

$$[M_1] = \begin{pmatrix} \sqrt{\frac{s-s'}{1-s'}} & 0 & \sqrt{\frac{s-s'}{1-s'}} \\ 0 & 1 & 0 \\ \sqrt{\frac{s-s'}{1-s'}} & 0 & -\sqrt{\frac{s-s'}{1-s'}} \end{pmatrix}, \tag{6.3.4}$$

$$[M_2] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{\frac{1+s'}{1+s}} & -\sqrt{\frac{s-s'}{1+s}} \\ 0 & -\sqrt{\frac{s-s'}{1+s}} & -\sqrt{\frac{1+s'}{1+s}} \end{pmatrix}. \tag{6.3.5}$$

We immediately recognize that the transformation M_1 and M_2 can be implemented with beamsplitters, labeled in Fig. 6.3.1 by BS1 and BS2, respectively. The corresponding matrix elements provide the transmission (diagonal) and reflection (off-diagonal) coefficients of these beamsplitters.

The degree of separation attained by the protocol can be certified by statistical analysis of the photon counts in the detectors placed in the ports $1'$ and $2'$, whereas those in the detector placed in port $3'$ provide the failure rate Q.

We should stress that this is a proof-of-principle protocol. The transformation U is design with the only aim of decreasing the overlap of the initial states, and no other communication or computational task is intended to be carried out by this implementation. However, we might consider removing the detectors in $1'$ and $2'$ and feed the output states into some other optical setup for further processing. Hence, this implementation can be thought of as a separation module in a larger experimental setup.

6.4. Implementation of Probabilistic Approximate Cloning

In order to clone N copies of state $|\psi\rangle$ approximately we need N+1 ports for our interferometer. This results in very complicated applications of the R-Z algorithm when N becomes large. We therefore demonstrate the solution on $1 \rightarrow 2$ cloning with equal prior probabilities. If we perform this operation as a one-shot measurement we would need a 5x5 unitary.

However we can probabilistically optimally separate the two input states, then apply a cloning unitary to make the desired copies. Since we know the optimal relationship between the inpupt and output overlaps from the previous chapter, the first step is choosing the desired final overlap and failure rate. Given the final overlap we optimally deterministically transform these states into the clones. This reduces the complexity of the problem sinces now we are working with 4x4 and 3x3 matricies. It was realized a shorter realization of separation exists for a different choice of basis states. We demonstate in detail next.

However we can probabilistically optimally separate the two input states using the results of the previous section, then apply a cloning unitary to make the desired copies. Since we know the optimal relationship between input and ouput overlaps from the previous chapter, the first step is choosing the desired final overlap and failure rate. Given the final overlap we optimally deterministically transform these states into the clones. This reduces the complexity of the problem since now we are working with a 4x4 matrix.

6.4.1. Shorter Separation Implementation. We start by choosing the same unitarity equations as before, instead choosing $|0\rangle$ as that for failure,

$$U|\psi_1\rangle = \sqrt{p_1}|\phi_1\rangle + \sqrt{q_1}|0\rangle \tag{6.4.1}$$

$$U|\psi_2\rangle = \sqrt{p_2}|\phi_2\rangle + \sqrt{q_2}|0\rangle \tag{6.4.2}$$

but now our states are chosen symmetrically as $|\psi_1\rangle = c_1|1\rangle + s_1|2\rangle$, $|\psi_2\rangle = c_1|1\rangle - s_1|2\rangle$, $|\phi_1\rangle = c_2|1\rangle + s_2|2\rangle$, and $|\phi_2\rangle = c_2|1\rangle - s_2|2\rangle$. For general a-priori probabilities the basis of the output states could be rotated, along with a changed overlap. However for the equal-priors case this rotation is non-existant, and the symmetry between the states is preserved. By sandwiching the unitary with the basis states we are able to write the equations for six of its nine elements as

$$\langle 0|U|\psi_1\rangle = c_1U_{01} + s_1U_{02} = \sqrt{q_1}, \tag{6.4.3}$$

$$\langle 0|U|\psi_1\rangle = c_1U_{11} + s_1U_{12} = c_2\sqrt{p_1}, \tag{6.4.4}$$

$$\langle 0|U|\psi_1\rangle = c_1U_{21} + s_1U_{22} = s_2\sqrt{p_1}, \tag{6.4.5}$$

and

$$\langle 0|U|\psi_2\rangle = c_1U_{01} - s_1U_{02} = \sqrt{q_2}, \tag{6.4.6}$$

$$\langle 1|U|\psi_2\rangle = c_1U_{11} - s_1U_{12} = c_2\sqrt{p_2}, \tag{6.4.7}$$

$$\langle 2|U|\psi_2\rangle = c_1U_{21} - s_1U_{22} = -s_2\sqrt{p_2}. \tag{6.4.8}$$

We can solve the the two sets of equations pairwise for the first six matrix elements, giving

$$[U] = \begin{pmatrix} \bullet & \frac{\sqrt{q_1}+\sqrt{q_2}}{2c_1} & \frac{\sqrt{q_1}-\sqrt{q_2}}{2s_1} \\ \bullet & \frac{c_2(\sqrt{p_1}+\sqrt{p_2})}{2c_1} & \frac{c_2(\sqrt{p_1}-\sqrt{p_2})}{2s_1} \\ \bullet & \frac{s_2(\sqrt{p_1}-\sqrt{p_2})}{2c_1} & \frac{s_2(\sqrt{p_1}+\sqrt{p_2})}{2s_1} \end{pmatrix}. \tag{6.4.9}$$

Applying the unitarity constraint $U^\dagger U = I$ can us nine equations for the remaining three uknown elements. However, using that the probability of success and failure for each state

is equal in the equal prior probability case,$p_1 = p_2$ and $q_1 = q_2$, and that when $\eta_1 = \eta_2$ we have the separation solution $p = \frac{\langle\psi_1|\psi_2\rangle - 1}{\langle\phi_1|\phi_2\rangle - 1}$, we can simplifying the matrix to

$$[U] = \begin{pmatrix} \bullet & \frac{\sqrt{s_2^2 - s_1^2}}{c_1 s_2} & 0 \\ \bullet & \frac{c_2 s_1}{c_1 s_2} & 0 \\ \bullet & 0 & 1. \end{pmatrix} \tag{6.4.10}$$

This lets us choose the final elements to be consistent with the operation of only a single beam splitter. implying we can choose the remaining elements to be

$$[U] = \begin{pmatrix} \frac{c_2 s_1}{c_1 s_2} & \frac{\sqrt{s_2^2 - s_1^2}}{c_1 s_2} & 0 \\ -\frac{\sqrt{s_2^2 - s_1^2}}{c_1 s_2} & \frac{c_2 s_1}{c_1 s_2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{6.4.11}$$

6.4.2. Deterministic Cloning Implementation. The unitary transformation for the second step is

$$U|\phi_1\rangle|0\rangle = |\xi_1\rangle|\xi_1\rangle \tag{6.4.12}$$

$$U|\phi_2\rangle|0\rangle = |\xi_2\rangle|\xi_2\rangle \tag{6.4.13}$$

where $|\xi_1\rangle = c_3|1\rangle + s_3|2\rangle$ and $|\xi_1\rangle = c_3|1\rangle - s_3|2\rangle$ Following a similar procedure we choose the basis states $|00\rangle$,$|10\rangle$,$|01\rangle$, and$|11\rangle$, giving us the final unitary as

$$[U] = \begin{pmatrix} \frac{c_3^2}{c_2} & 0 & 0 & \frac{s_3^2}{c_2} \\ 0 & \frac{c_3 s_3}{s_2} & -\frac{c_3 s_3}{s_2} & 0 \\ 0 & \frac{c_3 s_3}{s_2} & \frac{c_3 s_3}{s_2} & 0 \\ \frac{s_3^2}{c_2} & 0 & 0 & \frac{c_3^2}{c_2} \end{pmatrix} \tag{6.4.14}$$

This is clearly the action of two separate beam splitters M_{14} and M_{23} such that

$$[M_{14}] = \frac{1}{c_2}\begin{pmatrix} c_3^2 & s_3^2 \\ s_3^2 & c_3^2 \end{pmatrix} [M_{23}] = \frac{c_3 s_3}{s_2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \tag{6.4.15}$$

CHAPTER 7

Multi-Step Measurements and Information Loss

7.1. Measure and Prepare

7.2. Prepare and Measure

CHAPTER 8

Appendix

8.1. Appendix 2: Lagrange Multipliers