Optimal Measurement Tasks and Their Physical Realizations

by

Vadim Yerokhin

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

2015
Vadim Yerokhin
Some rights reserved.
This work is licensed under a Creative Commons
Attribution 4.0 United States License.
http://creativecommons.org/licenses/by/4.0/

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. János A. Bergou Chair of Examining Committee
Date	Prof. Igor L. Kuskovsky Executive Officer

Supervisory Committee:

Prof. Ed Feldman

Prof. Steven Greenbaum

Prof. Mark Hillery

Prof. Neepa T. Maitra

THE CITY UNIVERSITY OF NEW YORK

Abstract

Optimal Measurements Tasks and Their Physical Realizations

by

Vadim Yerokhin

Advisor: János 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This thesis reflects works previously published by the author and materials hitherto unpublished on the subject of quantum information theory.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.

Our interest lies in the scenario where we are given one of two quantum states prepared with a known a-priori probability. We are given full information about the states and are assigned the task of performing an optimal measurement on the incoming state. Given that none of these tasks is in general possible to perform perfectly we must choose a figure of merit to optimize, and as we shall see there is always a trade-off between competing figures of merit, such as the likelihood of getting the desired result versus the quality of the result.

For state discrimination the competing figures of merit are the success rate of the measurement, the errors involved, and the inconclusiveness. Similarly increasing the separation between states comes at a cost of less frequent successful applications of the separation protocol. For cloning, aside from successfully producing clones we are also interested in the fidelity of the clones compared to the original state, which is a measure of the quality of the clones. Because all quantum operations obey the same set of conditions for evolution one

may expect similar restrictions on disparate measurement strategies, and our work shows a deep connection between all three branches, with cloning and separation asymptotically converging to state discrimination.

Via Neumark's theorem, our description of these unitary processes can be implemented using single-photon interferometry with linear optical devices. Amazingly any quantum mechanical evolution may be decomposed as an experiment involving only lasers, beamsplitters, phase-shifters and mirrors. Such readily available tools allow for verification of the aforementioned protocols and we build upon existing results to derive explicit setups that the experimentalist may build.

Acknowledgements

I want to thank his advisor Prof. János Bergou for his guidance and incredible understanding patience. Science does not require humanity and compassion but these facets made it more meaningful and worthwhile. Further thanks are given to Andi Shehu, without whom I would likely not have joined the information theory group, and whose collaboration and support were an integral and most enjoyable part of the graduate experience. Prof. Ed Feldman, for his expansive knowledge and useful insights into the mathematics of physics. Prof. Hillery, for his help and contributions to the theory group. Significant thanks are owed to Emilio Bagan, without whom much of the present work would not be possible. His exemplary scientific mind and humble and sincere demeanor made collaboration a joy.

Thanks also to the members of the physics department at Hunter college who made the years terrifically fun; Jorge Colon, Mallory Gobet, Ugur Guney, Kay Hiranaka, Tetiana Nosach, Denis Sharoukhov, Zhenmao Wan.

Most dear thanks are given to my family who believed in me and made the journey with me. This thesis is dedicated to my grandfather Iosif Fishman and grandmother Ellita Volodarskaya for their love and guidance. My parents Janet and Igor, and brother Oleg, for their unconditional love and support.

Contents

List of Figures

CHAPTER 1

Introduction

Motivation to study the subsequent topics comes from a long list of achievements by physicists for the last hundred years. The tenet that all information is physical and therefore encoded in the states of systems allows us to consider optimal interactions with quantum systems as measurement matrices acting on density matrices representing the states' configurations. Because the state is not a directly observable quantity we have prescribed methods of interacting with it. Particularly the formulation due to Neumark [**1**] where the evolution of the state of a quantum system can be described by the action of a unitary matrix. This unitary can be decomposed into measurement operators responsible for different outcomes. We will be particularly interested in using this description to study three topics. Given one of a set of two known quantum states, but we don't know which, we will seek optimal strategies for discriminating and cloning the state, or increasing the separation between the states.

For discrimination, given one of two quantum states whose preparation and a-priori preparation probability we know, we wish to determine which state we were given. As we shall see this is a highly non-trivial task and involves choosing relevant figures of merit to optimize. We review the existing measurement strategies and apply them to different classes of states.

In general, state discrimination can be seen as a two-step measurement process: in the first we probabilistically separate the two states to make them more distinguishable, and then perform a discrimination on the resulting states. Given that separation involves a change in overlap between the initial states, we are motivated to study state separation independently of any protocol which may use it. We discover the relationship between optimal separation and the success rate of such a protocol.

Another set of protocols which depend on separation include quantum cloning. If we wish to make copies of one of two given states, we increase our chances of making good copies by first making the states more distinguishable, hence we discover that state separation plays an integral role in the solution of this general problem.

We conclude by describing linear optical experiments for actualizing these measurement tasks. Our implementations require technology readily available today and such as has been used for previous experimental tasks.

CHAPTER 2

Discrimination of Pure States

2.1. Introduction

Quantum state discrimination is the theory by which a measurement on a quantum system relates to the type of information gathered from that system. This field is important for a wide range of applications in communication, cryptography, and computation. For a recent review see [**2**]. It remains important to devise better theoretical frameworks for these processes and to realize them experimentally.

To perform a measurement on a quantum mechanical system, we must look to its state. The state of a quantum system contains all information about the system [**3**]. Given a mixture of different states, only orthogonal states can be distinguished perfectly [**4**]. For non-orthogonal states [**5**], optimum discrimination will depend on the type of information we wish to obtain and on prior information that we have. Depending on the type of results needed, several different strategies exist.

The traditional measurement minimizes the chance of incorrectly identifying the state [**6**]. It is known as the Minimum Error (ME) strategy. In this strategy the measurements aim to identify the most prominent states, either pure [**7**] or mixed [**8, 9**]. It was first experimentally tested by Barnett and Riis [**10**], and quantitatively tested for nonorthogonal states by Clarke *et al.*[**11**].

If the states to be distinguished are linearly independent it is possible to construct a measurement that always correctly identifies the input state but does not always give an answer, i.e., has a certain failure rate. This strategy is called Unambiguous state Discrimination (UD) and was introduced by Ivanovic [**12**]. The optimal solution for two pure states and arbitrary a-priori probabilities was solved by Jaeger and Shimony [**13**]. The solution was later extended to include symmetric pure states, three pure states, and other classes of

mixed states. Experimental realizations of UD were first done by Huttner *et al.* [**14**]. Clarke *et al.* considered pure trine and tetrad states in two dimensions [**15**]. UD was experimentally verified for discriminating pure and mixed quantum states by Mohseni *et al.* [**16**].

If the input states are not linearly independent it is not possible to distinguish them perfectly, but it is possible to improve on the error associated with the ME strategy. For this a more recent addition to the measurement strategy ensemble was the Maximum Confidence (MC) strategy [**17**]. Here the idea of the Bayes' rule is used to maximize the chance of the appropriate detector clicking for its desired state. The experimental realization was suggested by Herzog and Benson [**18**], and done by Steudle *et al.* [**19**].

The MC strategy was a move in the direction of Intermediate (IM) strategies, where both errors and failure are permissible for a specific measurement goal. The IM scheme was introduced by Chefles and Barnett [**20**] where they found the lower bound on the combination of error and inconclusive result for two pure states with equal prior probabilities. Later, Fiurasek and Jezek [**21**] generalized this result for mixed states.

Recently the two pure state IM problem was first solved by Sugimoto *et al.* [**22**] using Semi-Definite Programming (SDP), where a fixed error rate was used to obtain an analytical expression for the minimum failure rate. Then Bagan *et al.* [**23**] solved the problem using an operator transformation technique that reformulated the problem into a ME one with an extra optimization parameter. Herzog [**24, 25**] generalized the SDP approach to include classes of mixed states to the realm of IM analytic solutions.

In this chapter we describe the IM solution two different ways. Both solutions are important in our subsequent work. First we describe the normalization that reduces the problem [**23**] to a ME problem with an implicit degree of freedom: a fixed failure rate. The second solution uses the method of Lagrange multipliers with constrained failure rate to develop an explicit algebraic solution. The advantage of this solution is that it allows us to extract not only the total probabilities of success and failure, as were derived previously, but individual probabilities for each state. These error rates will be subsequently used in the experimental implementation.

2.1.1. State Representation. We start the formal treatment of the subject with a brief mathematical introduction to the relevant objects and historically relevant topics. We begin this by describing a pure quantum state ψ as a vector 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

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

where the unitary matrix U obeys $UU^\dagger = I$. There are several ways to view this formula. The first, due to Neumark's theorem, 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

$$U(|\psi_A\rangle \otimes |\phi_B\rangle) = \sum_i A_i |\psi_A\rangle \otimes |i_B\rangle. \tag{2.1.1}$$

This equation describes the alternative effects of the evolution between states $|\psi_A\rangle$ and $|\phi_B\rangle$ in their joint space $\mathcal{H}_1 \otimes \mathcal{H}_2$ as a Kraus operator A_i for each outcome direction $|i_B\rangle$ of the ancillary Hilbert space such as a measurement device. For example a particle striking a particular detector can be associated with a particular operator A_i.

A simpler version of this unitary evolution can be considered when we exclude the ancilla. Here 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 ρ previously described can be viewed 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 operator can be either a projector onto an eigenvector of the Hilbert space or a generalized measurement (POVM). In the latter case it must fulfill only 2 properties:

(1) $\sum \Pi_i = 1$.

(2) $\Pi_i \geq 0$. This corresponds to real, non-negative eigenvalues (measurement outcomes).

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

2.1.2. No Perfect Discrimination. Consider the problem of discriminating between two pure states $|\psi_1\rangle$ and $|\psi_2\rangle$ that are sent one at a time with a known probability distribution. By Neumark's theorem since the Hilbert space of our states is two dimensional, the measurement space is also two dimensional allowing for two orthogonal outcomes, one associated with each state. If we could always perform this perfectly then we should be able to write a unitary U such that if we apply this unitary to our incoming states, we always turn them into orthogonal vectors:

$$U|\psi_1\rangle = |1\rangle,$$

$$U|\psi_2\rangle = |2\rangle.$$

The states $|1\rangle$ and $|2\rangle$ are orthogonal, or $\langle 1|2\rangle = 0$, and each result is associated with the respective input. However since the unitary is inner-product preserving, taking the product of the first equation with the second's adjoint shows that such a unitary is 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

$$\Pi_1|\psi_2\rangle = 0,$$

$$\Pi_2|\psi_1\rangle = 0,$$

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 to the input states and their spanning the space, proved contradictory we must give up one of these two functions in order to perform a physical measurement. 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 with known probabilities η_1 and η_2 respectively, such that $\eta_1 + \eta_2 = 1$.

2.2. Minimum Error Discrimination

Historically, the first solution to this problem is due to Helstrom [**6**]. Now known as the Minimum Error (ME) strategy, the figure of merit is the average rate of mistakenly 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 again associate the outcome Π_i with state $\rho_i = |\psi_i\rangle\langle\psi_i|$. 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:

$$P_E = \frac{1}{2}[1 - \sqrt{1 - 4\eta_1\eta_2|\langle\psi_1|\psi_2\rangle|^2}]. \tag{2.2.1}$$

2.2.1. Neumark Solution. We provide a derivation of this result to demonstrate the Neumark formalism. Since both ψ_1 and ψ_2 can now evolve to either 1 or 2, we must write the unitary equations as

$$U|\psi_1\rangle = \sqrt{p_1}|1\rangle + \sqrt{r_1}|2\rangle,$$
$$U|\psi_2\rangle = \sqrt{p_2}|2\rangle + \sqrt{r_2}|1\rangle,$$

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

$$s = \langle\psi_1|\psi_2\rangle = \sqrt{p_1 r_2} + \sqrt{p_2 r_1}. \tag{2.2.2}$$

We wish to minimize the average error rate $P_E = \eta_1 r_1 + \eta_2 r_2$ subject to the constraint (2.2.2). We solve this two variable problem using the method of Lagrange multipliers. The constrained error equation can be written as

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

Our extrema will be found when all partial derivatives of this equation are zero, $\frac{\partial F_E}{\partial r_1} = \frac{\partial F_E}{\partial r_2} = 0$. The resulting equations can be written such that the left-hand-side of each is dependent on only a single variable:

$$\frac{2\eta_1}{\lambda}\sqrt{r_1(1-r_1)} = \sqrt{r_1 r_2} - \sqrt{(1-r_1)(1-r_2)}, \tag{2.2.3}$$
$$\frac{2\eta_2}{\lambda}\sqrt{r_2(1-r_2)} = \sqrt{r_1 r_2} - \sqrt{(1-r_1)(1-r_2)}. \tag{2.2.4}$$

The right hand sides of Eq.(2.2.3) and (2.2.4) 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 2.2.2. After a lot of algebra we

get the explicit form of the individual error rates as

$$r_i = \frac{1}{2}\left[1 - \frac{1 - 2\eta_i s^2}{\sqrt{1 - 4\eta_1\eta_2 s^2}}\right]. \tag{2.2.5}$$

Inserting r_1 and r_2 into $P_e = \eta_1 r_1 + \eta_2 r_2$, the Helstrom bound is retrieved.

2.3. Unambiguous Discrimination

It was noticed by Ivanovic [**12**] that we may completely eliminate error from the measurement results by giving up the constraint that our two measurement operators span the whole space when $\Pi_1 + \Pi_2 = I$. If we give up this condition then we need 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 $\Pi_0 + \Pi_1 + \Pi_2 = I$. Our measurement operators for two pure states can be written as

$$\begin{aligned} \Pi_1 &= \frac{p_1}{|\langle\psi_1|\psi_2^\perp\rangle|^2}|\psi_2^\perp\rangle\langle\psi_2^\perp|, \\ \Pi_2 &= \frac{p_2}{|\langle\psi_2|\psi_1^\perp\rangle|^2}|\psi_1^\perp\rangle\langle\psi_1^\perp|. \end{aligned} \tag{2.3.1}$$

We can now determine Π_0 as a function of the success probabilities p_i by writing it as $\Pi_0 = I - \Pi_1 - \Pi_2$. The eigenvalues of Π_0 must be non-negative, giving us the inequality constraint between the individual failure rates as

$$q_1 q_2 \geq |\langle\psi_1|\psi_2\rangle|^2. \tag{2.3.2}$$

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. (2.3.2), giving us the minimum failure rate at

$$Q \equiv Q_0 = 2\sqrt{\eta_1\eta_2}|\langle\psi_1|\psi_2\rangle|. \tag{2.3.3}$$

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 orthogonally to the less likely state. Therefore the total UD solution is

$$Q_{\mathrm{c}} = \begin{cases} \eta_1 + \eta_2 \cos^2\theta, \text{if } \eta_1 < \dfrac{\cos^2\theta}{1+\cos^2\theta} \equiv \eta_1^{(l)}, \\ \eta_2 + \eta_1 \cos^2\theta, \text{if } \eta_1 > \dfrac{1}{1+\cos^2\theta} \equiv \eta_1^{(r)}, \\ 2\sqrt{\eta_1\eta_2}\cos\theta \equiv Q_0, \text{if } \eta_1^{(l)} \leq \eta_1 \leq \eta_1^{(r)} \ . \end{cases} \tag{2.3.4}$$

2.4. 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. By minimizing the failure rate for a given rate of error we may more accurately discriminate states in the lab. On the other hand, success rates may significantly increase for small changes in the error rate, thereby motivating a small error rate in permissible scenarioss. We notice from the optimized intermediate discrimination curve depicted in Fig. 2.4.1 that the trade-off for marginally increased detection accuracy comes at the cost of significant inconclusive results. This motivates an intermediate strategy where we wish to perform the best measurement for a fixed rate of failure. By varying the failure rate from 0 to Q_c (Q_0 in this case) we should recover both the ME and UD results in a solution that minimizes error for a particular failure rate.

2.4.1. Operator Transformation. We review the solution from [**23**] involving a transformation that eliminates the failure operator from the discrimination problem. We are given two pure states $\rho_1 = |\psi_1\rangle\langle\psi_1|$ and $\rho_2 = |\psi_2\rangle\langle\psi_2|$ with a-priori probabilities η_1 and η_2 respectively. These two probabilities add to one as usual. 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 $Q = tr[\Pi_0(\eta_1\rho_1 + \eta_2\rho_2)]$. This is analogous to minimizing the error rate since $P_s + P_e + Q = 1$, since measurement operators Π_i span

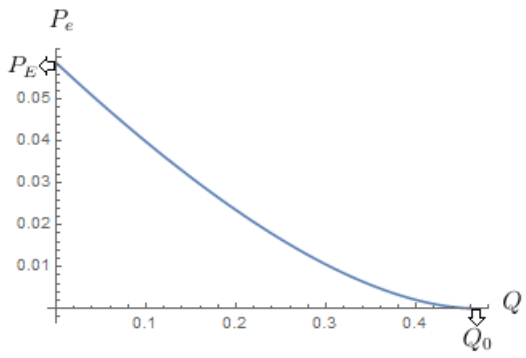


FIGURE 2.4.1. Optimal Error rate P_e vs Q interpolates between the ME limit P_E and UD limit Q_0 for values $0 \leq Q \leq Q_0$ when $Q_0 = .47$.

the Hilbert space: $\Pi_1 + \Pi_2 + \Pi_0 = I$. The transformation we implement is

$$\Omega^{-1/2}[\Pi_1 + \Pi_2]\Omega^{-1/2} = I, \tag{2.4.1}$$

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

$$P_e = \frac{1}{2}(1 - Q - \sqrt{(1-Q)^2 - (Q - Q_0)^2}) \tag{2.4.2}$$

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 for the POVM regime. 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. Fig. (2.4.1) reflects this solution.

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

require an extra ancillary degree of freedom if we want to measure in the basis of our original qubit. This dimension we add by extending our Hilbert space with a direct sum extension.

$$U|\psi_1\rangle = \sqrt{p_1}|1\rangle + \sqrt{r_1}|2\rangle + \sqrt{q_1}|0\rangle \tag{2.4.3}$$
$$U|\psi_2\rangle = \sqrt{r_2}|1\rangle + \sqrt{p_2}|2\rangle + \sqrt{q_2}|0\rangle$$

Here p_i is the probability that the state i is correctly identified when it is sent into the measurement 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.

2.4.2.1. *Equal Priors Solution.* The equal priors solution is always a special case and here it allows for a beautifully quick solution. The important insight is that in this case the outcome probabilities for both states must be identical due to symmetry, so $p_i = P_s$, $r_i = P_e$, and $q_i = Q$. Taking the inner product of the two equations in Eq. (2.4.4) we get the unitarity constraint $s = \sqrt{p_1 r_2} + \sqrt{p_2 r_1} + \sqrt{q_1 q_2}$ that simplifies to $s = \sqrt{p(1-p-Q)} + Q$ But this is just a quadratic equation for p, giving us the larger root as

$$p = \frac{1}{2}[1 - Q + \sqrt{(1-Q)^2 - (s-Q)^2}]. \tag{2.4.4}$$

2.4.3. Full Solution. We use the Lagrange multiplier method to minimize P_e subject to the overlap constraint

$$s = \sqrt{(1 - r_1 - q_1)r_2} + \sqrt{(1 - r_2 - q_2)r_1} + \sqrt{q_1 q_2}, \tag{2.4.5}$$

where we've used $p_i + r_i + q_i = 1$. Our constrained function is

$$F = \eta_1 r_1 + \eta_2 r_2 + \lambda(s - \text{OverlapConstraint}).$$

We first find the partial derivatives $\partial F/\partial r_i = 0$, then solve for the error rates r_i in terms of the Lagrange multiplier λ. Substituting the error rates into the overlap constraint allows us to solve for λ, which gives us the individual error rates after a lot of algebra as

$$r_1 = \frac{1}{2}[\alpha_1 - \frac{[2\eta_2\omega - \alpha_1(1-Q)]}{\sqrt{(1-Q)^2 - 4\omega\eta_1\eta_2}}], \tag{2.4.6}$$

$$r_2 = \frac{1}{2}[\alpha_2 - \frac{[2\eta_1\omega - \alpha_2(1-Q)]}{\sqrt{(1-Q)^2 - 4\omega\eta_1\eta_2}}], \tag{2.4.7}$$

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. We put these into the expression of minimum error and do a final optimization that amounts again to distributing the failure rate among the states. Before optimization this reads

$$P_E = \eta_1 r_1 + \eta_2 r_2 = \frac{1}{2}[(1-Q) - \sqrt{(1-Q)^2 - 4\eta_1\eta_2(s - \sqrt{q_1 q_2})^2}]$$

We're left to minimize the expression under the square root subject to constraint $Q = \eta_1 q_1 + \eta_2 q_2$, leaving us with a single variable function that is easily minimized to find the optimal condition $\eta_1 q_1 = \eta_2 q_2$ and gives us the optimal interpolative solution Eq. (2.4.2):

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

This result only makes sense if none of these probabilities become negative. One can check that for $\eta_1 \leq 1/2$ (i.e., $\eta_1 \leq \eta_2$) p_1 and r_2 vanish at the threshold value

$$Q_{\mathrm{b}} = \frac{2\eta_1\eta_2(1-s^2)}{1-Q_0}. \tag{2.4.8}$$

Moreover, r_1 and r_2 also vanish at the UD value Q_c from 2.3.4 as they should. We thus note that the solution given by the first two equations cannot be correct if

$$Q_{\mathrm{b}} < Q_c,$$

since for Q in the allowed range $Q_{\rm b} < Q < Q_c$ both p_1 and r_2 become negative and a three-outcome measurement does not exist. The conditions under which this happens are discussed in [**23**] and have nothing to do with the approach used (Neumark or POVM).

In the regime $Q_{\rm b} < Q < Q_c$, $(p_1 = r_2 = 0)$, the unitarity constraint becomes

$$s = \sqrt{r_1}\sqrt{p_2} + \sqrt{q_1}\sqrt{q_2}, \tag{2.4.9}$$

but $r_1 = 1 - q_1$ and $p_2 = 1 - q_2$, so we can write

$$s = \sqrt{1 - q_1}\sqrt{1 - q_2} + \sqrt{q_1}\sqrt{q_2}.$$

We also have

$$\frac{P_{\rm s}}{\eta_2} = p_2 = 1 - q_2; \quad \frac{P_{\rm e}}{\eta_1} = r_1 = 1 - q_1.$$

We can solve for $\sqrt{P_{\rm s}/\eta_2}$ to obtain

$$\sqrt{\frac{P_{\rm s}}{\eta_2}} = \sqrt{\frac{P_{\rm e}}{\eta_1}}s - \sqrt{1 - \frac{P_{\rm e}}{\eta_1}}\sqrt{1 - s^2}.$$

Therefore,

$$1 - Q - P_{\rm e} = \eta_2 \left(\sqrt{\frac{P_{\rm e}}{\eta_1}}s - \sqrt{1 - \frac{P_{\rm e}}{\eta_1}}\sqrt{1 - s^2} \right)^2, \tag{2.4.10}$$

which is the solution given in [**23**]. These bounds are depicted in Fig. (2.4.2).

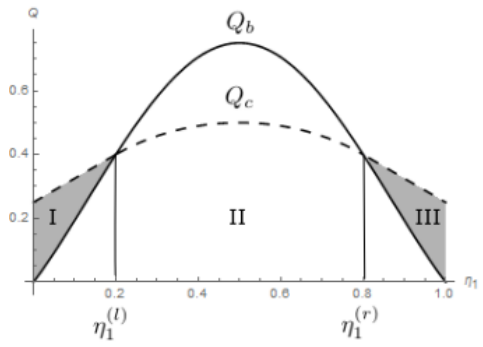


FIGURE 2.4.2. Failure rate Q vs a-priori probability η_1, with the maximum failure Q_c Eq.(2.3.4) and boundary $Q_{\rm b}$ Eq. (2.4.8) between the two projective measurement regimes I and III, and the POVM regime II.

CHAPTER 3

Discrimination of Mixed States

3.1. Introduction

Here we consider the problem of discriminating between two mixed states,

$$\rho_1 = \sum_i r_i |r_i\rangle\langle r_i|, \tag{3.1.1}$$

$$\rho_2 = \sum_i s_i |s_i\rangle\langle s_i|, \tag{3.1.2}$$

where their eigenvalues are normalized such that $\sum_i r_i = \sum_i s_i = 1$. Again we allow for each to occur with different a-priori probability η_1 or η_2 with $\eta_1 + \eta_2 = 1$. Since the two states aren't necessarily diagonalizable in the same basis, the measurement operators must be projectors or POVM's on the joint Hilbert space. If the two states have non-zero kernels, then unambiguous discrimination is possible by projecting onto those spaces. Otherwise we must use a figure of merit such as error or confidence to optimize our measurement.

We begin this chapter with a brief treatment of minimum error discrimination. Then we discuss the discrimination of mixed states through the perspectives developed in the preceding one, particularly by applying the operator transformation solution to several classes of mixed states. For the first, the two mixed states lie in a Jordan basis with respect to one another, thereby creating a structure of independent two-dimensional subspaces. Each subspace is similar to the original problem. This structure simplifies the mixed-state problem because we can apply the previous two-dimensional solution to each subspace in our Hilbert space to find the optimal error rate for a fixed failure probability in the subspaces. The solution is a normalized version of the known result. However an extra optimization is required to find the optimal distribution of failure over the subspaces and this final optimization leads to an interesting threshold structure. When the failure rate is decreased

from the UD limit, we reach a critical value when one subspace's failure rate reaches zero and it is excluded from the optimization. Subsequently all remaining subspaces are removed one after the other as we continue to decrease the failure rate, requiring a re-optimization at each epoch. As we reach the ME boundary only a single subspace has any failure result. We thoroughly analyze this threshold structure, considering points of continuity, intersection, and limits to the regime. The other class of mixed states we consider are qubits. Again a subsequent optimization is required after the initial transformation is applied. This optimization is non-trivial. While numerical results can be obtained for all ranges of solutions, only in particular cases can we obtain analytic results.

We also consider applying the maximum confidence (MC) framework to high dimensional mixed states to generalize the role of detection operators. Instead of associating a single detection operator for a given state, we add the differentiating factor of confidence to a class of operators for each state. The advantage of this formulation is that each measurement operator is extracting information at the optimal level of confidence from the corresponding eigenvector in the confidence space of the density matrices. This allows us to expand the success rate of the UD regime in higher dimensional spaces and have higher confidence levels for some measurement results over any interpolative strategy of the type described earlier.

3.1.1. Minimum Error Discrimination. The first solution was also provided by Helstrom[**6**], but we provide an alternative derivation [**27**] using the POVM decomposition of the measurement space. We consider two detectors Π_1 and Π_2 such that $\Pi_1 + \Pi_2 = 1$ with $\Pi_1 = \sum_1^d |\lambda_i\rangle\langle\lambda_i|$, $\Pi_2 = \sum_d^N |\lambda_i\rangle\langle\lambda_i|$ where the $|\lambda_i\rangle$ form a basis on the total space of both states of dimension N and is ordered such that the first d vectors are more likely to contain ρ_1.

Particularly, the error probability takes the form

$$P_E = \frac{1}{2}[1 - Tr|\eta_2\rho_2 - \eta_1\rho_1|] \tag{3.1.3}$$

$$= \frac{1}{2}[1 - Tr|\Lambda|], \tag{3.1.4}$$

where $\Lambda = \eta_2\rho_2 - \eta_1\rho_1$ plays the role of determining the weight of each state in a particular eigenvector. Its diagonalization gives

$$\Lambda = \sum^{N} \lambda_i |\lambda_i\rangle\langle\lambda_i| = \sum_{1}^{d} \lambda_{i<0} |\lambda_i\rangle\langle\lambda_i| + \sum_{d+1}^{N} \lambda_{i\geq 0} |\lambda_i\rangle\langle\lambda_i|.$$

For eigenvectors belonging to the value 0, grouping those vectors with either Π_1 or Π_2 does not affect the measurement outcomes. Naturally this equation reduces to the Helstrom bound for pure states.

3.2. Mixed States with Jordan Structure

In this section we discuss the interpolative discrimination of a class of mixed states with a Jordan structure [**26**]. As before, they are to be discriminated with a-priori probabilities η_1, η_2 but now their density matrices 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 of the operator transformation to multiple subspaces is straightforward at first so we begin with the two subspaces example.

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

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

We define 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 transformation 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 coefficient

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

Proceeding similarly for the second state $|s_i\rangle$ in that subspace we get the error rate in this subspace as

$$P_{e,1} =$$

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

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 should 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. Subspaces formalism. 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 multiple fiber optic cables. Each cable contains two degrees of freedom that could be horizontal and vertical polarization.

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 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 (3.2.9) don't change. Now it is straightforward to apply the 2d solution to

each subspace. 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 generalized version of the solution we derived for one subspace in (3.2.8):

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

3.2.3. 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

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

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

$$Q_i = \frac{Q_{0,i} - \omega_i Q_0 + Q(\omega_i - Q_{0,i})}{1 - Q_0}. \tag{3.2.11}$$

Now the optimized subspace error rate is

$$P_{e,i} = \frac{1}{2}(\omega_i - Q_i - (\omega_i - Q_{0,i})\sqrt{\frac{1 + Q_0 - 2Q}{1 - Q_0}}), \tag{3.2.12}$$

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

$$P_e = \frac{1}{2}(1 - Q - \sqrt{(1 - Q)^2 - (Q - Q_0)^2}). \tag{3.2.13}$$

While this appears identical to the 2d solution (2.4.2), 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 (3.2.11) 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 (3.2.13) 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 (3.2.11) to find the critical value of Q for that subspace to be

$$Q = Q_c^i = \frac{\omega_i Q_0 - Q_{0,i}}{\omega_i - Q_{0,i}}. \tag{3.2.14}$$

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 worthwhile to consider 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

$$Q_i^{(1)} = \frac{Q_{0,i}\Lambda_{N-1} - \omega_i F_{N-1} + Q(\omega_i - Q_{0,i})}{\Lambda_{N-1} - F_{N-1}} \tag{3.2.15}$$

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

$$Q_i^{(n)} = \frac{Q_{0,i}\Lambda_{N-n} - \omega_i F_{N-n} + Q(\omega_i - Q_{0,i})}{\Lambda_{N-n} - F_{N-n}}. \tag{3.2.16}$$

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

$$Q_c^{(n)i} = \frac{\omega_i F_{N-n} - Q_{0,i}\Lambda_{N-n}}{\omega_i - Q_{0,i}}. \tag{3.2.17}$$

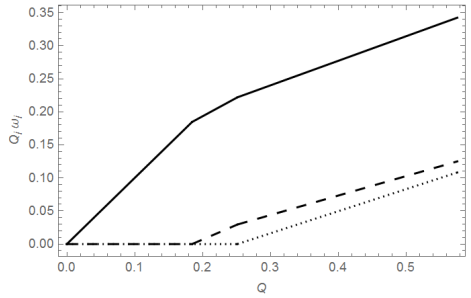
This is similar to the first set of critical points found in (3.2.14). For this to be positive (and to be a candidate for elimination) we need $\frac{F_{N-n}}{\Lambda_{N-n}} \geq \frac{Q_{0,i}}{\omega_i}$ which states that the relative UD failure rate for that subspace be smaller than average to be considered 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. We demonstrate this threshold structure in the following figure, which shows the failure rates and the error rates versus the total failure rate for all three states over all thresholds.

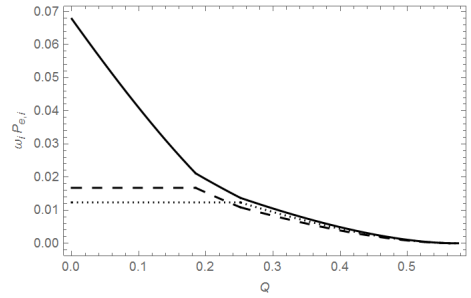
3.2.4.3. *Continuity and intersection.* It is worthwhile to demonstrate the continuity of our solutions for the Q_i's. To do this we need to show that the optimal solutions match at the boundaries where a subspace is discarded, or

$$Q_i^{(n)}(Q = Q_c^{(n)N-n}) = Q_i^{(n+1)}(Q = Q_c^{(n)N-n}), \tag{3.2.18}$$

where we have chosen to consider the n^{th} 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.



(a)



(b)

FIGURE 3.2.1. Three subspace interpolation graphs showing the minimum failure probability (a) and minimum error probability (b) for $\eta_1 = 1/2$, $r_{1,2,3} = 5/8, 1/8, 1/4$, $s_{1,2,3} = 3/8, 3/8, 1/4$, and $\theta_{1,2,3} = \pi/4, \pi/3, \pi/3$.

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 $\omega_1 > \frac{Q_{0,1}}{Q_0}$.

We show an example of such intersection for both the error and failure graphs below:

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

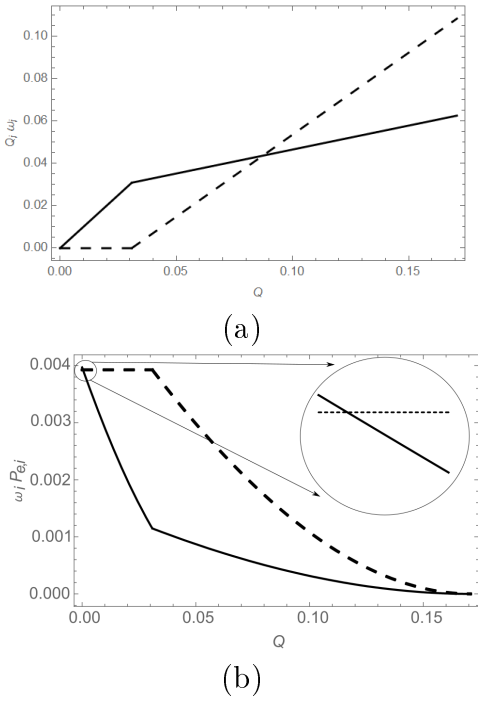


FIGURE 3.2.2. Two subspace interpolation graphs showing the minimum failure probability (a) and minimum error probability (b) for $\eta_1 = 1/2$, $r_{1,2} = 3/4, 1/4$, $s_{1,2} = 3/4, 1/4$, and $\cos\theta_{1,2} = 1/(4\sqrt{3}), 1/4$.

domain (SSD) where we use projective measurements. Here we ignore the less likely state and focus on optimally discriminating the other one in that subspace.

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

$$Q = Q_c = \frac{2\eta_1\eta_2 sin^2\theta}{1 - Q_0}. \tag{3.2.19}$$

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

$$Q_i = Q_i^{cc} = \frac{2\eta_{1,i}\eta_{2,i} sin^2\theta_i}{\omega_i - Q_{0,i}} \tag{3.2.20}$$

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

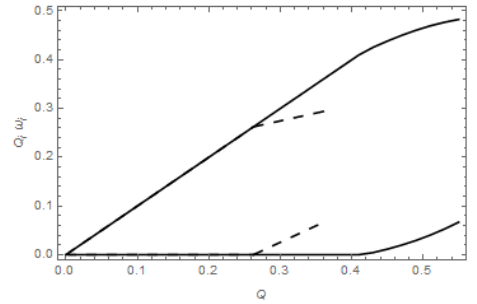
$$\frac{\omega_i}{1 + \cos^2\theta_i} \leq \eta_{1,i}, \tag{3.2.21}$$

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

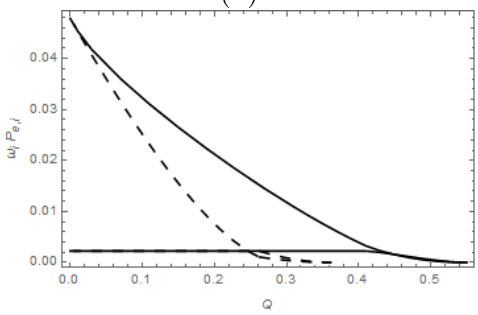
$$\frac{\omega_i \cos^2\theta_i}{1 + \cos^2\theta_i} \geq \eta_{1,i} \tag{3.2.22}$$

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 normalized requirements for a measurement to be projective are identical to the single-subspace case, which is also the condition for the projective regime in the UD case (5.2.16). We provide examples



(a)



(b)

FIGURE 3.2.3. Individual failure rates (a) and individual error rates (b) vs total failure rate Q for two subspaces. The solid lines represent the incorrect solution if there were no projective (single-state) regime. The dashed lines represent the correct values from numerical optimization. It is clear that they diverge at a different threshold from the normal threshold, and total permissible failure rate is lower than would be otherwise. Also interesting is that only subspace one reaches the projective regime but both failure rates become non-linear as a result. Plotted for $r_1 = s_2 = 9/10$, $s_1 = r_2 = 1/10$, $\theta_1 = \pi/16$, $\theta_2 = 3\pi/7$.

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. We want to implement the same operator transformation as in the previous section. Since the states are mixed our equations take the

form

$$\widetilde{\rho_1} = \frac{\Omega^{1/2}\rho_1\Omega^{1/2}}{Tr(\Omega\rho_1)}, \tag{3.3.1}$$

$$\widetilde{\eta_1} = \frac{\eta_1 Tr(\Omega\rho_1)}{1-Q}, \tag{3.3.2}$$

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

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

We rewrite the term $(c_1 s_2 - c_2 s_1)^2 = sin^2\theta$ and 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

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

At this point the optimization remains to choose the relationship between the two failure rates, or, equivalently, of c_1 and c_2. Since they are constrained by the overlap of the two states, we can easily rewrite this as a function of a single variable and optimize numerically. However the resulting equation is not tractable algebraically for the general case. We look at the minimum error case for insight into the problem.

3.3.1. 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

$$\begin{aligned} b &= \eta_2 - \eta_1, \\ c &= \frac{1}{4}[(\eta_1^2 - \eta_2^2) - (\eta_1 p - \eta_2 d)^2] - \eta_1\eta_2 pd sin^2\theta. \end{aligned}$$

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

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

This solution may not seem remarkable but it is very closely related to the subspace ME solution: For subspaces where $\rho_1 = \sum r_i|r_i\rangle\langle r_i|$ and $\rho_2 = \sum s_i|s_i\rangle\langle s_i|$ with $\langle r_i|s_j\rangle = \delta_{ij} cos\theta_i$ we got the error rate with failure to be

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

where $\eta_i = \eta_1 r_i + \eta_2 s_i$ and $Q_{0,i} = 2\sqrt{\eta_1 r_i \eta_2 s_i} cos\theta_i$. Setting $Q_i = 0$ we get

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

Let's look at the square root in this expression:

$$\eta_i^2 - Q_{0,i}^2 = (\eta_1 r_i + \eta_2 s_i)^2 - 4\eta_1 r_i \eta_2 s_i cos^2\theta_i = (\eta_1 r_i - \eta_2 s_i)^2 + 4\eta_1 r_i \eta_2 s_i sin^2\theta_i.$$

If we simply call $r_i = p$ and $s_i = d$ for some subspace then we reproduce the square-root term of the previous solution! This leads us to believe that the problems are very closely related.

3.4. Maximum Confidence Measurements

We begin from the work of Croke *et al.* [**17**] and Herzog [**28**] on maximum confidence. There two mixed states were measured using 3 POVM's including one failure operator. The final successful measurement Π_i was a normalized, shortened projector onto the eigenket corresponding to the largest eigenvalue of the normalized density matrix $\widetilde{\rho}_i$. This allows us

to maximize the confidence given by

$$C_i = \frac{\eta_i Tr(\rho_i \Pi_i)}{Tr(\rho \Pi_i)}. \tag{3.4.1}$$

Mathematically, given ρ_1 and ρ_2, the density matrices of the two states to be distinguished and their a-priori probabilities η_1 and η_2, the MC formalism tells us to find:

$$\rho = \eta_1 \rho_1 + \eta_2 \rho_2, \tag{3.4.2}$$

$$\widetilde{\rho_i} = \eta_i \rho^{-\frac{1}{2}} \rho_i \rho^{-\frac{1}{2}}, \tag{3.4.3}$$

$$\widetilde{\Pi_i} = \frac{\rho^{\frac{1}{2}} \Pi_i \rho^{\frac{1}{2}}}{Tr \rho \Pi_i}, \tag{3.4.4}$$

where $C_i = Tr \widetilde{\rho_i} \widetilde{\Pi_i}$ is the ith confidence of the measurement. The normalized density matrix for the first state is

$$\widetilde{\rho_1} = v_{max} |v_1\rangle\langle v_1| + v_{min} |v_N\rangle\langle v_N| + \sum_{n=2}^{N-1} v_n |v_n\rangle\langle v_n|.$$

Conveniently $\widetilde{\rho_1} = I - \widetilde{\rho_2}$, and we assume for clarity that no eigenvalues other than $1/2$ are degenerate. The measurement operators take the form

$$\Pi_1 = c_1 \rho^{-\frac{1}{2}} |v_1\rangle\langle v_1| \rho^{-\frac{1}{2}}, \tag{3.4.5}$$

$$\Pi_2 = c_2 \rho^{-\frac{1}{2}} |v_N\rangle\langle v_N| \rho^{-\frac{1}{2}}, \tag{3.4.6}$$

$$\Pi_0 = I - \Pi_1 - \Pi_2. \tag{3.4.7}$$

Hence we have that the rest of the spectrum of eigenvalues are tossed into the failure operator. Finally, the probability of failure is minimized while maintaining the rate of confidence.

The new idea is to leave the Π_1 and Π_2, and reduce the failure by a two-fold strategy: First by creating distinct measurements for the remainder of the non-equal eigenvalue spectrum of $\widetilde{\rho_1}$, while assigning the equal eigenvalues to the failure operator. And second by checking the confidence of the measurement that completes the space, which we can assign on a ket-by-ket basis to operators detecting either ρ_1 or ρ_2. There would be up to as many

distinct operators D+P+2 as $\lambda_{max-(P-1)} \neq 1/2$ and $\lambda_{min+(D-1)} \neq 1/2$ for $D, P < N/2$. The decreasing difference between the eigenvalues corresponds to decreasing confidence, which can be interpreted as a ladder of confidence-valued POVM's. Hence we start by decomposing the normalized density matrices as:

$$\widetilde{\rho_1} = \sum_{\lambda \neq 1/2}^{D+P} \lambda_i |v_i\rangle\langle v_i| + 1/2 \sum_{n=D+P+1}^{N} |v_n\rangle\langle v_n|, \tag{3.4.8}$$

with the condition that $\lambda_i \geq \lambda_{i+1}$. Our new POVM's would retain two measurements of the previous scheme (those for the extremum eigenvalues) with the additional ones for other eigenvalues:

$$\begin{aligned}
\Pi_1^{C_1} &= c_1 \rho^{-\frac{1}{2}} |v_1\rangle\langle v_1| \rho^{-\frac{1}{2}}, \\
&\vdots \\
\Pi_1^{C_P} &= c_P \rho^{-\frac{1}{2}} |v_P\rangle\langle v_P| \rho^{-\frac{1}{2}}, \\
\Pi_2^{C_{P+1}} &= c_{P+1} \rho^{-\frac{1}{2}} |v_{P+1}\rangle\langle v_{P+1}| \rho^{-\frac{1}{2}}, \\
&\vdots \\
\Pi_2^{C_{P+D}} &= c_{P+D} \rho^{-\frac{1}{2}} |v_{P+D}\rangle\langle v_{P+D}| \rho^{-\frac{1}{2}}, \\
\Pi_0 &= c_{P+D+1} \rho^{-\frac{1}{2}} \sum_{n=D+P+1}^{N} |v_n\rangle\langle v_n| \rho^{-\frac{1}{2}},
\end{aligned}$$

It follows that the total success rate for the measurement is now

$$P_s = \eta_1 Tr \rho_1 \sum_{n=1}^{P} c_n \Pi_1^{C_n} + \eta_2 Tr \rho_2 \sum_{n=1}^{D} c_{P+n} \Pi_2^{C_{P+n}}. \tag{3.4.9}$$

The failure operator Π_0 is hence only necessary for the range of the spectrum in which both states are equally likely, and the new measurements for ρ_1 and ρ_2 replace what would otherwise have been less-than-maximum-confidence spaces sorted into the failure operator.

We demonstrate this formalism with a few trivial examples. Suppose we wish to discriminate between two mixed states in 3 dimensions with 1 dimensional kernels. Suppose

$$\rho_1 = \frac{5}{9}(|v_1\rangle\langle v_1| + \frac{4}{5}|v_2\rangle\langle v_2|),$$
$$\rho_2 = \frac{5}{6}(|v_3\rangle\langle v_3| + \frac{1}{5}|v_2\rangle\langle v_2|),$$

with equal a-priori probabilities $\eta_1 = \eta_2$. We solve the problem for maximum confidence measurements:

$$\begin{aligned}
\rho &= \eta_1\rho_1 + \eta_2\rho_2 = \frac{1}{36}(10|v_1\rangle\langle v_1| + 11|v_2\rangle\langle v_2| + 15|v_3\rangle\langle v_3|), \\
\rho^{1/2} &= \frac{1}{6}(\sqrt{10}|v_1\rangle\langle v_1| + \sqrt{11}|v_2\rangle\langle v_2| + \sqrt{15}|v_3\rangle\langle v_3|), \\
\rho^{-1/2} &= 6(\frac{1}{\sqrt{10}}|v_1\rangle\langle v_1| + \frac{1}{\sqrt{11}}|v_2\rangle\langle v_2| + \frac{1}{\sqrt{15}}|v_3\rangle\langle v_3|), \\
\widetilde{\rho_1} &= (|v_1\rangle\langle v_1| + \frac{8}{11}|v_2\rangle\langle v_2|), \\
\widetilde{\rho_2} &= (|v_3\rangle\langle v_3| + \frac{3}{11}|v_2\rangle\langle v_2|).
\end{aligned}$$

If we were to use the original MC framework the measurements would take the form

$$\Pi_1 = |v_1\rangle\langle v_1|, \quad \Pi_2 = |v_3\rangle\langle v_3|, \quad \Pi_0 = |v_2\rangle\langle v_2|.$$

$C_1 = C_2 = 1$ implies unambiguous discrimination. We attain the success probability of the measurement as

$$P_s = \eta_1 Tr(\rho_1\Pi_1) + \eta_2 Tr(\rho_2\Pi_2) = \frac{1}{2}(\frac{5}{9} + \frac{5}{6}) = \frac{25}{36},$$

with failure probability equaling the difference from unity,

$$Q = 1 - P_s = \frac{11}{36}.$$

Since the middle eigenvalue isn't $\frac{1}{2}$, if we assign Confidence-Valued POVM's to each eigenvalue we get:

$$\Pi_1 = |v_1\rangle\langle v_1|, \quad \Pi_1^{\frac{8}{11}} = |v_2\rangle\langle v_2|, \quad \Pi_2 = |v_3\rangle\langle v_3|.$$

We have no failure operator and thus have $Q = 0$, instead creating a success probability associated only with the measurement along $|v_2\rangle$ as

$$P_s' = \eta_1 Tr(\rho_1 \Pi_1^{\frac{8}{11}}) = \frac{1}{2}(\frac{4}{9}) = \frac{8}{36}.$$

Again $C_1 = C_3 = 1$ for the projectors onto the kernels while $C_2 = \frac{8}{11}$ is the confidence of the newly-interpreted operator so that the total success of the new measurement strategy is

$$P_s^{tot} = P_s + P_s' = \frac{33}{36}.$$

Our strategy has lost none of the advantages of unambiguous discrimination and has significantly improved the probability of success. By allowing for more than a single measurement to detect one particle, we have a meaningful strategy for handling higher-dimensional discrimination problems.

For two states in 7 dimensions with equal prior probabilities, suppose the density matrices are

$$\rho_1 = \frac{2}{7}(|v_1\rangle\langle v_1| + \frac{4}{5}|v_2\rangle\langle v_2| + \frac{4}{5}|v_3\rangle\langle v_3| + \frac{1}{2}|v_4\rangle\langle v_4| + \frac{1}{5}|v_5\rangle\langle v_5| + \frac{1}{5}|v_6\rangle\langle v_6|),$$

and

$$\rho_2 = \frac{2}{7}(|v_7\rangle\langle v_7| + \frac{1}{5}|v_2\rangle\langle v_2| + \frac{1}{5}|v_3\rangle\langle v_3| + \frac{1}{2}|v_4\rangle\langle v_4| + \frac{4}{5}|v_5\rangle\langle v_5| + \frac{4}{5}|v_6\rangle\langle v_6|).$$

Following our strategy we write:

$$\rho = \frac{1}{7}I_7, \quad \rho^{\frac{1}{2}} = \frac{1}{\sqrt{7}}I_7, \quad \rho^{-\frac{1}{2}} = \sqrt{7}I_7,$$

$$\Pi_1 = |v_1\rangle\langle v_1|, \quad \Pi_1^{\frac{4}{5}} = |v_2\rangle\langle v_2|, \quad \Pi_1^{\frac{4}{5}} = |v_3\rangle\langle v_3|, \quad \Pi^{\frac{1}{2}} = |v_4\rangle\langle v_4|,$$

$$\Pi_2^{\frac{4}{5}} = |v_5\rangle\langle v_5|, \quad \Pi_2^{\frac{4}{5}} = |v_6\rangle\langle v_6|, \quad \Pi_2 = |v_7\rangle\langle v_7|.$$

Now $P_s = \eta_1 Tr(\rho_1\Pi_1) + \eta_2 Tr(\rho_2\Pi_2) = \frac{10}{35}$ is the success probability of the UD measurements and $Q_{UD} = \frac{25}{35}$. The extra conclusive measurements give us an additional success rate as $P_s' = \eta_1 Tr(\rho_1\Pi_1^{\frac{4}{5}}) + \eta_2 Tr(\rho_2\Pi_2^{\frac{4}{5}}) = \frac{16}{35}$, with total success probability being the sum of the

individual rates and a failure rate significantly lower than Q_{UD}:

$$P_s^{tot} = P_s + P_s' = \frac{26}{35}, \quad P_e' = \frac{4}{35}, \quad Q' = \frac{5}{35}.$$

Now we apply the theory to existing situations, beginning with the simplest, two symmetric pure states in 2d. First we show the example of an attempt to improve the success probability for UD measurements, where the input states are $|\psi_1\rangle = cos\theta|0\rangle + sin\theta|1\rangle$, $|\psi_2\rangle = cos\theta|0\rangle - sin\theta|1\rangle$. The well known measurements take the form

$$\Pi_1 = a_1|\phi_1\rangle\langle\phi_1| \tag{3.4.10}$$

$$\Pi_2 = a_2|\phi_2\rangle\langle\phi_2| \tag{3.4.11}$$

$$\Pi_0 = 1 - \Pi_1 - \Pi_2 \tag{3.4.12}$$

where

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

$$|\phi_1\rangle = \sin\theta|0\rangle - cos\theta|1\rangle$$

and

$$a_1 = \frac{1 - \sqrt{\frac{\eta_2}{\eta_1}}\cos 2\theta}{\sin^2 2\theta} \tag{3.4.13}$$

$$a_2 = \frac{1 - \sqrt{\frac{\eta_1}{\eta_2}}\cos 2\theta}{\sin^2 2\theta} \tag{3.4.14}$$

for unequal a-priori probabilities. Now our success and failure rates are

$$P_s = \eta_1 Tr\rho_1\Pi_1 + \eta_2 Tr\rho_2\Pi_2 = 1 - 2\sqrt{\eta_1\eta_2}\cos 2\theta,$$

$$Q = 2\sqrt{\eta_1\eta_2}\cos 2\theta.$$

Interestingly, even when the a-priori probabilities are unequal and $\eta_1 > \eta_2$, the optimal failure operator still clicks equally often for both states:

$$q_1 = \eta_1 Tr\rho_1\Pi_0 = \eta_1(Tr\rho 1 - Tr\rho_1\Pi_1 - Tr\rho_1\Pi_2) = \eta_1(1 - 4a_1\cos^2\theta\sin^2\theta) = \sqrt{\eta_1\eta_2}cos2\theta.$$

Therefore associating it with a measurement for either state we only muddle the success and error probabilities. This result suggests than we investigate the case of two pure states in 2 dimensions in which the failure rate for each state is different. In fact, we know that the optimal relation between the failure rates for intermediate pure state discrimination is $\eta_1 q_1 = \eta_2 q_2$, hence the inconclusive operator really is true to its name in this case. However the projective measurement domain for two pure states allows for a different optimality condition. In this regime $\eta_1 q_1 \neq \eta_2 q_2$ so we can assign the final measurement to measure the more likely state.

For the case of 2 mixed states in two dimensions we can write each density operator as

$$\rho_i = \frac{1}{2}\begin{pmatrix} 1 + r_i \cos 2\theta_i & r_i \sin 2\theta_i \\ r_i \sin 2\theta_i & 1 + r_i \cos 2\theta_i \end{pmatrix}. \tag{3.4.15}$$

Using $\eta_1 = \eta_2$ and the formalism of operator transformation described earlier in the chapter we can write the failure rates as

$$q_1 = \eta_1 Tr \Pi_0 \rho_1 = \frac{1}{4}\xi(1 + r_1 \cos 2\theta_1)$$

and

$$q_2 = \eta_2 Tr \Pi_0 \rho_2 = \frac{1}{4}\xi(1 + r_2 \cos 2\theta_2)$$

are the failure rates for each state. For a reassignment of the failure operator to make sense, the condition $q_1 > q_2$ or $r_1 cos 2\theta_1 > r_2 cos 2\theta_2$ must hold. This would give us new success, error and failure probabilities $P_s' = q_1$, $P_e' = q_2$, $Q' = 0$. If we consider the confidence associated with state one and Π_0 in two dimensions, we find

$$C_0 = \frac{\eta_1 Tr \rho_1 \Pi_0}{Tr \rho \Pi_0} = \frac{q_1}{q_1 + q_2} \tag{3.4.16}$$

and the individual failure rates can be written very symmetrically as

$$q_1 = \eta_1 Tr \rho_1 \Pi_0 = \eta_1 (1 - Tr[\rho_1(\Pi_1 + \Pi_2)]),$$

$$q_2 = \eta_2 Tr \rho_2 \Pi_0 = \eta_2 (1 - Tr[\rho_2(\Pi_1 + \Pi_2)]).$$

The condition for $q_1 > q_2$ is

$$\eta_1 - \eta_2 > Tr[(\Pi_1 + \Pi_2)(\eta_1\rho_1 - \eta_2\rho_2)] = -Tr[(\Pi_1 + \Pi_2)\Lambda]. \tag{3.4.17}$$

In the Maximum Confidence framework this is

$$\eta_1 - \eta_2 > -Tr[\rho^{-1}(\alpha_1\widetilde{\Pi_1} + \alpha_2\widetilde{\Pi_2})\Lambda]. \tag{3.4.18}$$

Finally we look at the last example from **[28]** for a case with higher-dimensional operators: 2 mixed states in d dimensions, d even, with equal a-priori probabilities. These can be described as

$$\rho_i = \frac{2p}{d}\sum_{k=1}^{d/2} |r_k^{(i)}\rangle\langle r_k^{(i)}| + (1-p)\frac{I_d}{d} \tag{3.4.19}$$

where $|r_k^{(i)}\rangle = c_k|0\rangle \pm s_k|1\rangle$. Herzog provides us with the MC operators

$$\Pi_1^{MC} = \sum_{k=1}^{m} \frac{|v_k^{\gamma_{max}}\rangle\langle v_k^{\gamma_{max}}|}{1 + p\cos\gamma^{max}}, \tag{3.4.20}$$

$$\Pi_2^{MC} = \sum_{k=1}^{m} \frac{|w_k^{\gamma_{max}}\rangle\langle w_k^{\gamma_{max}}|}{1 + p\cos\gamma^{max}}, \tag{3.4.21}$$

where

$$|v_k^{(\gamma_k)}/w_k^{(\gamma_k)}\rangle = \sqrt{\frac{1 - p\cos\gamma_k}{2}}|0\rangle_k \pm \sqrt{\frac{1 + p\cos\gamma_k}{2}}|1\rangle_k.$$

This gives the failure rate as

$$Q^{MC} = 1 - \frac{2m}{d}p\cos\gamma^{max} \tag{3.4.22}$$

and confidences

$$C_1 = C_2 = \frac{1}{2} + \frac{p\sin\gamma_{max}}{2\sqrt{1 - p^2\cos^2\gamma_k}}.$$

She also mentions that an alternative, average, measurement where all the dimensions contribute to conclusive results can be implemented:

$$\Pi_1^{avg} = \sum_{k=1}^{d/2} \frac{|v_k^{\gamma_k}\rangle\langle v_k^{\gamma_k}|}{1 + p\cos\gamma^k}, \tag{3.4.23}$$

$$\Pi_2^{avg} = \sum_{k=1}^{d/2} \frac{|w_k^{\gamma_k}\rangle\langle w_k^{\gamma_k}|}{1 + p\cos\gamma^k}, \tag{3.4.24}$$

with failure rate

$$Q^{avg} = \frac{2p}{d} \sum_{1}^{d/2} \cos\gamma^k \tag{3.4.25}$$

and necessarily smaller confidences

$$C_1^{avg} = C_2^{avg} = \frac{1}{2} + \frac{p\sum_{k=1}^{d/2} sin\gamma_k \sqrt{\frac{1-p\cos\gamma_k}{1+p\cos\gamma_k}}}{2\sum_{k=1}^{d/2}(1 - p\cos\gamma_k)}.$$

Now following our new strategy of multiple measurement operators for each state:

$$\Pi_1^{C_j} = \sum_{k=1}^{m_j} \frac{|v_k^{\gamma_k}\rangle\langle v_k^{\gamma_k}|}{1 + p\cos\gamma^k}, \tag{3.4.26}$$

$$\Pi_2^{C_j} = \sum_{k=1}^{m_j} \frac{|w_k^{\gamma_k}\rangle\langle w_k^{\gamma_k}|}{1 + p\cos\gamma^k}. \tag{3.4.27}$$

where m_i denotes the size of the degenerate space associated with γ_i and confidences

$$C_j = \frac{1}{2} + \frac{p\sin\gamma^j}{2\sqrt{1 - p^2\cos^2\gamma^j}}. \tag{3.4.28}$$

Clearly

$$\sum_{i=1}^{d/2} \Pi_k^{C_i} = \Pi_k^{avg}$$

so while our measurement operators span the same space in this case, they provide us with more information than their averages. This information can be used to increase the success rate of the measurement.

In conclusion, we found an effective intermediate strategy that maximizes the confidence and then success rate of the measurement by utilizing more confidences for each state in

a high-dimensional space. We have also opened the door for more optimization searches. For example, the solution set to the 2 states in 2d problem for which $q_1 > q_2$ implies three operators Π_1, Π_1', Π_2 will achieve higher rates of success than nearby equal-failure solutions, and so begs for an optimization condition on the competing values of total confidence and success in a measurement.

For the experimental implementation of multiple measurements for each state we may borrow the explanation of the additional measurement for failure. The measurement process is a two-step one. The first, probabilistic step attempts to transform the states into $\widetilde{\rho_i}$. This first measurement has associated outcomes λ_{fail} and $\lambda_{success}$ and associated POM elements Π_{succ} and Π_{fail}. A successful measurement transforms the states ρ_1 and ρ_2 into $\widetilde{\rho_1}$ and $\widetilde{\rho_2}$ respectively, which are then discriminated by $\widetilde{\Pi_1}$ and $\widetilde{\Pi_2}$. For multiple measurement outcomes this extends in the same manner, giving us at most as many measurement operators as degrees of freedom in our new normalized space.

CHAPTER 4

Pure State Separation

4.1. Introduction

All intermediate discrimination strategies may be considered as two-step processes. In the first step we probabilistically separate the two states. If this step is successful, we perform a projective measurement on the more separate states to optimally discriminate. We consider the first step by itself for several reasons: any probabilistic transformation acting on a two-state family can be characterized by the change in separation of the two states. Whether the map is physical for a given failure rate is entirely dependent on the state separation problem. As we shall see, there are important connections between state separation and other measurement tasks. Chefles and Barnett [**29**] first considered state separation for equal prior probabilities and we extend their work to the general a-priori region using two distinct geometric methods [**30**]. For the first we parametrize the relevant equations so that the optimal point is described by an intersection of two conic sections: a parabola and an ellipse. This parametrization allows us several different perspectives on the problem. For example, if we wish to maximize separation for a given failure rate, or to see the tradeoff between separation and failure, the parametrized system of equations of the conic sections allows us to solve for any two variables in terms of the other parameters of the problem. The second solution will be described in the subsequent chapter as it pertains to quantum cloning, and it is straightforward to generalize the parametrization devised for perfect probabilistic cloning to describe the optimal rate of failure for a fixed separation.

We can start by imagining 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

$$Q = \eta_1 q_1 + \eta_2 q_2. \tag{4.1.1}$$

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 other.

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. One option is to find the solution in terms of $Q_{\min}$ as a function of the prior probability η_1. However, one may need to know 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 $A_{\text{succ}}^\dagger A_{\text{succ}} + A_{\text{fail}}^\dagger A_{\text{fail}} = I$. 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 [**2**]. 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 [**31**]

$$U|\psi_1\rangle|0\rangle = \sqrt{p_1}|\psi_1'\rangle|\alpha_1\rangle + \sqrt{q}_1|\phi\rangle|\alpha_0\rangle, \tag{4.1.2}$$

$$U|\psi_2\rangle|0\rangle = \sqrt{p_2}|\psi_2'\rangle|\alpha_2\rangle + \sqrt{q}_2|\phi\rangle|\alpha_0\rangle. \tag{4.1.3}$$

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 output 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 general setup with two failure states $|\phi_1\rangle$ and $|\phi_2\rangle$ in Eqs. (4.1.2) and (4.1.3). 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. (4.1.2) and (4.1.3) with themselves shows that our probabilities are normalized: $p_i + q_i = 1$. Similarly, by taking the product of Eq. (4.1.2) with

Eq. (4.1.3), we find the unitarity constraint,

$$s = \sqrt{p_1 p_2}\,\beta + \sqrt{q_1 q_2}, \tag{4.1.4}$$

where $\beta = s'|\langle\alpha_1|\alpha_2\rangle|$. Without any loss of generality, in deriving Eq. (4.1.4) 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 $0 \leq \beta \leq s$, 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. (5.2.3) 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 .

4.2. 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. (4.1.4), 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

$$u = \sqrt{q_1 q_2}; \quad v = \frac{q_1 + q_2}{2}. \tag{4.2.1}$$

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

$$v = \frac{1 + u^2}{2} - \frac{(u - s)^2}{2s'^2}. \tag{4.2.2}$$

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. (4.2.2) become thinner. For $s' = 0$ they degenerate into the vertical segment $u = s$, $0 \leq v \leq (1 + s^2)/2$.

Under the same transformation, Eq. (4.2.1), 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

$$\begin{aligned} u &= \frac{Q}{\sqrt{1-\Delta^2}} \cos\theta, \\ v &= \frac{Q}{1-\Delta^2} + \frac{Q\Delta}{1-\Delta^2} \sin\theta, \end{aligned} \tag{4.2.3}$$

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 family, as one can easily check using Eq. (4.2.3). Fig. 4.2.1 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. 4.2.1. This figure also illustrates optimality. Given a maximum failure rate $Q_{\max}$ and some initial overlap s ($Q_{\max} = 0.45$ and $s = 0.6$ in the example considered in the figure), we plot the corresponding ellipse defined by Eq. (4.2.3) Among the various parabolas, characterized by the final overlap s' the one that minimizes s' 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.

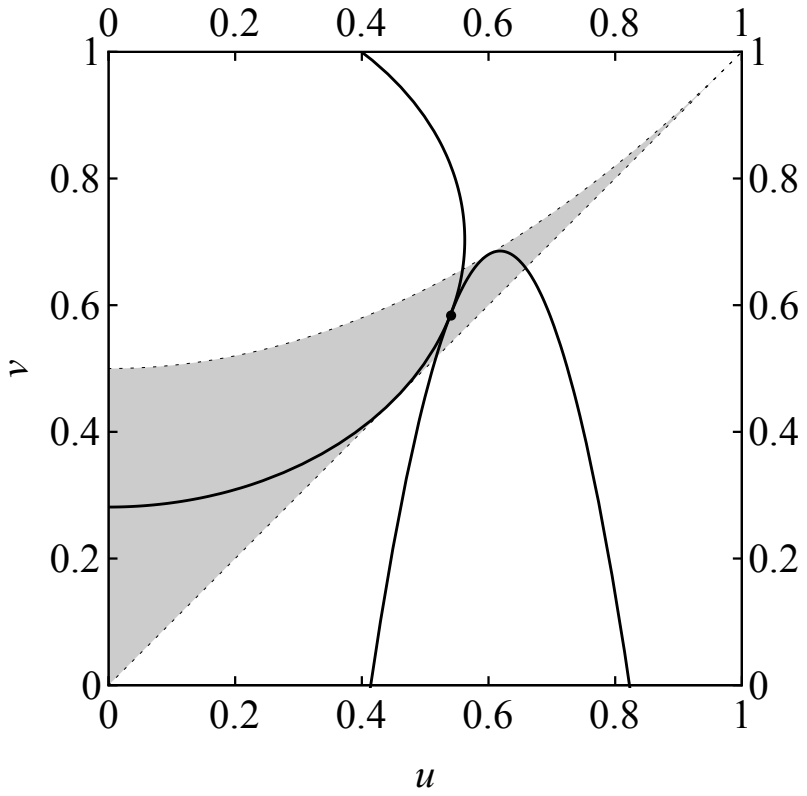


FIGURE 4.2.1. The optimal separation point as the intersection of the parabola Eq. (4.2.2) and ellipse $Q = \eta_1 q_1 + \eta_2 q_2$. In this figure $\eta_1 = 0.2$, $s = 0.6$ and $Q = Q_{\max} = 0.45$. The optimal (minimum) value of s', which gives the solid parabola, turns out to be $s' = 0.17$.

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

$$\begin{aligned} \frac{dv}{du} &= \frac{v'}{u'} = -\frac{\Delta}{\sqrt{1-\Delta^2}} \cot\theta, \\ \frac{dv}{du} &= u - \frac{u-s}{s'^2}, \end{aligned} \tag{4.2.4}$$

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

$$\begin{aligned} Q\frac{1+\Delta\sin\theta}{1-\Delta^2} &= \frac{1}{2}+\frac{Q^2\cos^2\theta}{2(1-\Delta^2)}-\frac{1}{2s'^2}\left(\frac{Q\cos\theta}{\sqrt{1-\Delta^2}}-s\right)^2, \\ \frac{\Delta\cot\theta}{\sqrt{1-\Delta^2}} &= \frac{1-s'^2}{s'^2}\frac{Q\cos\theta}{\sqrt{1-\Delta^2}}-\frac{s}{s'^2}. \end{aligned} \tag{4.2.5}$$

where to obtain the first (second) equation we have simply substituted Eq. (4.2.3) into Eq. (4.2.2) [Eq. (4.2.4)]. 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:

$$s' = -\frac{\sqrt{(1-Q)^2-(\Delta+Q\sin\theta)^2}}{\Delta+Q\sin\theta}\tan\theta, \tag{4.2.6}$$

$$s = \frac{Q\Delta(1+\sin^2\theta)-(1-\Delta^2-2Q)\sin\theta}{\sqrt{1-\Delta^2}(\Delta+Q\sin\theta)\cos\theta}. \tag{4.2.7}$$

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

$$\theta_{\max}=\begin{cases} 0 & \text{if} \quad Q\le 1-\Delta, \\ \arcsin\dfrac{1-Q-\Delta}{Q} & \text{if} \quad Q\ge 1-\Delta. \end{cases} \tag{4.2.8}$$

One can easily check the given minimum value of θ by substituting in Eqs. (4.2.6) and (4.2.7) 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. (4.2.8) reveal the appearance of the phase transition in the limit $s'\to 0$ that we discussed in previous sections. If $Q\ge 1-\Delta$, substituting the second line of Eq. (4.2.8) in Eq. (4.2.5) 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 $Q\ge 1-\Delta$ is equivalent to $\eta_1+s^2\eta_2\ge 1-\Delta$, which can be immediately seen to give $\eta_1\le s^2/(1+s^2)$. So we obtain the second line in Eq. (2.3.4), corresponding to the "symmetry-broken phase". If $Q\le 1-\Delta$,

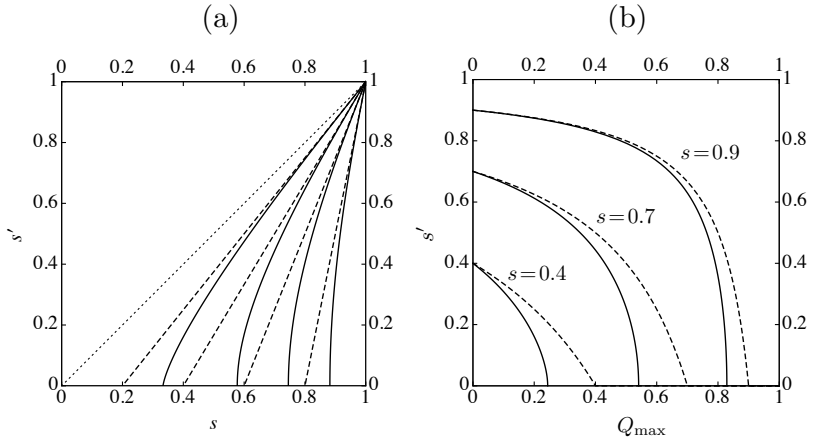


FIGURE 4.2.2. (a) Plots of s' vs. s for $\eta_1 = 0.1$ (solid lines) and $\eta_1 = 0.5$ (straight dashed lines) and for values of the failure rate. From left to right $Q_{\max} = 0.2$, 0.4, 0.6, 0.8. The dotted line is the (trivial) curve for $Q_{\max} = 0$, which is the straight line $s' = s$. (b) Minimum final overlap vs. maximum failure probability for various values of the initial overlap and the same two values of η_1 used in (a).

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. (4.2.8) has the same content as Eq. (2.3.4). Recall that we are assuming $\eta_1 \leq 1/2 \leq 1/(1+s^2)$. The first line in Eq. (2.3.4) never applies under this assumption.

Eqs. (4.2.6) and (4.2.7) are plotted in Fig. 4.2.2 (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,

$$s_{\rm cr} = \begin{cases} \dfrac{Q_{\max}}{2\sqrt{\eta_1\eta_2}} & \text{if} \quad Q \leq 2\eta_1, \\ \sqrt{\dfrac{Q_{\max}-\eta_1}{\eta_2}} & \text{if} \quad Q \geq 2\eta_1, \end{cases} \tag{4.2.9}$$

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. (4.1.4) using that $q_1 = q_2 = Q$:

$$s' = \begin{cases} 0 & \text{if} \quad s \leq Q_{\max}, \\ \dfrac{s - Q_{\max}}{1 - Q_{\max}} & \text{if} \quad s \geq Q_{\max}. \end{cases} \tag{4.2.10}$$

This expression could also be obtained by carefully taking the limit $\Delta \to 0$ in Eqs. (4.2.6) through (4.2.8). 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).

4.3. Tradeoff Between Maximum Separation and Failure Rate

By solving the system Eq. (4.2.5) for Q and s' , we obtain a parametric expression for the tradeoff curve (Q, s') in terms of the polar angle θ:

$$\begin{aligned} s'^2 &= \sqrt{1-\Delta^2}\left(\frac{\sin\theta}{\Delta+\sin\theta}\right)^2 \\ &\times \frac{\sqrt{1-\Delta^2}(1+s^2)\cos\theta - 2s(1+\Delta\sin\theta)}{\cos\theta}, \end{aligned} \tag{4.3.1}$$

$$Q = \frac{s\sqrt{1-\Delta^2} + \Delta\, s'^2 \cot\theta}{(1-s'^2)\cos\theta}. \tag{4.3.2}$$

Note that Eq. (4.3.1) is an expression for the square of the final overlap. To keep the formula for Q, Eq. (4.3.2), short, we use s'^2 as a shorthand for Eq. (4.3.1). The range of θ in Eqs. (4.3.1) and (4.3.2) is:

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

where the upper limit of the interval can be written as

$$\theta_{\max} = \begin{cases} 0 & \text{if} \quad \eta_1 \geq \dfrac{s^2}{1+s^2}, \\ -\arccos\dfrac{2s\sqrt{1-\Delta^2}}{1-\Delta+s^2(1+\Delta)} & \text{if} \quad \eta_1 \leq \dfrac{s^2}{1+s^2}. \end{cases} \tag{4.3.3}$$

The lower limit in the range of allowed θ can be derived from Eqs. (4.3.1) and (4.3.2) by imposing that $Q = 0$ at $s' = s$. The upper limit can be derived from Eq. (4.3.1) 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. (4.3.3) in Eq. (4.3.2) we obtain $Q = s\sqrt{1-\Delta^2}$ ($Q = [1 - \Delta + s^2(1+\Delta)]/2$), which is the third (second) case in Eq. (2.3.4).

Fig. 4.2.2 (b) shows various plots of the separation vs. failure-rate tradeoff curve. As in Fig. 5.2.3, 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. (4.2.10). 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 $Q_{\rm UD}$ 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 $Q = Q_{\rm UD}$ ($< Q_{\max}$) to fully separate the input states.

CHAPTER 5

Cloning Pure States

It is a logical conclusion that if perfect quantum state discrimination is impossible, then so too should perfect quantum cloning be impossible. Were it possible, we would clone infinite copies and then extract total information from any system. The original proofs of the no-cloning theorem are by Wooters and Dieks [**32, 33**], but we demonstrate a simple proof of the principle using Neumark's theorem shortly. Strangely, the impossibility of cloning leads to advantages to quantum based communications protocols of practical relevance over classical ones. A common example is provable security in quantum key distribution, with recent works showing more applications [**34, 35**]. Developments in cloning, including deterministic but approximate cloning [**36, 37, 38, 39, 40, 41**] and probabilistically perfect cloning [**31, 42, 43**] provide anchors for better understanding quantum theory as a whole, such as the relationship between the no-cloning and no-signaling theorems [**44**], and fundamental limits on quantum measurements [**45, 46, 47, 48**]. In particular, cloning's relationship with state discrimination will be central to this discussion. For reviews citing recent developments, applications and experiments related to cloning see [**49, 50**].

When knowledge of the state preparation is available, perfect cloning is probabilistically possible. With the first result in probabilistic cloning, Duan and Guo [**31**] considered the problem of producing perfect clones of linearly independent pure states and focused on the two state case. They found the maximum average success rate when both states are equally likely, the case of equal priors, and set this success probability as lower bound for arbitrary prior probabilities. While other work has been done on this problem, there has until now been no general solution. In our first result on cloning we obtain the general analytic solution [**51**] and, with its help, examine the relationship of cloning to state discrimination. Subsequently we solve the approximate probabilistic cloning problem for the same initial

conditions. In order to optimize the conditioned global fidelity of the clone states we must find the relationship between this fidelity of the clones and the rate of successful cloning.

5.0.1. No-Cloning Theorem. Suppose we wrote the equations describing the unitary that cloned one of two known 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

$$U|\psi_1\rangle|0\rangle = |\psi_1\rangle|\psi_1\rangle \tag{5.0.4}$$

$$U|\psi_2\rangle|0\rangle = |\psi_2\rangle|\psi_2\rangle \tag{5.0.5}$$

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.

5.1. 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

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

5.2. Probabilistic Exact Cloning

On the opposite side of the cloning spectrum we consider the machines that make perfect clones of known states. 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,

$$Q = \eta_1 q_1 + \eta_2 q_2. \tag{5.2.1}$$

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 [**31**], 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

$$U|\psi_i^m\rangle|0\rangle = \sqrt{p_i}|\psi_i^n\rangle|1\rangle + \sqrt{q_i}\rangle|0\rangle, \quad i = 1, 2. \tag{5.2.2}$$

Here the ancillas are initialized in a reference state $|0\rangle$. The states of the flag associated with successful cloning, $|1\rangle$, is orthogonal to the state associated with failure, $|0\rangle$. To proceed, we take the inner product of equation i with itself in (5.2.2), yielding that the probabilities are normalized, $p_i + q_i = 1$. Taking the inner product of equation $i = 1$ with $i = 2$ in (5.2.2) yields the main constraint,

$$s^m = \sqrt{p_1 p_2}\, s^n + \sqrt{q_1 q_2}, \tag{5.2.3}$$

which is a consequence of the unitarity of U. Here we used the notation $s \equiv \langle\psi_1|\psi_2\rangle$.

5.2.1. Geometric description of optimality. Before developing the analytical theory of optimizing (minimizing) Q, we present a geometric picture, similar in spirit to that in [**52**], that visually solves the optimization problem and serves as guide for the subsequent calculations. The curve and straight line in Fig. 5.2.1. are Eq. (5.2.3) and Eq.(5.2.1) respectively. The curves' intersection implies a solution point, but it is easy to see that since a smaller y-intercept for any fixed slope line is always more optimal (it gives a smaller Q), the intersection must be in the form of a tangency given that the curve has a unique tangent at every point. This point gives $Q_{\min}$ and defines the optimal cloning strategy.

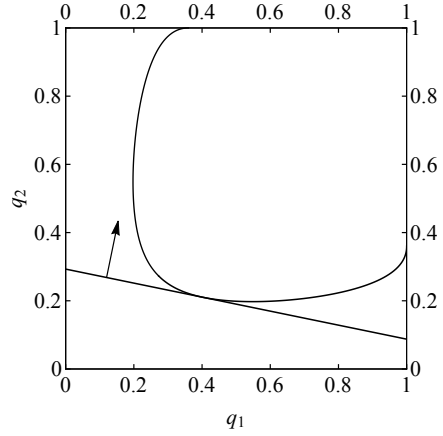


FIGURE 5.2.1. Unitarity curve, q_2 vs. q_1, from Eq. (5.2.3). The figure also shows the optimal straight segment $Q = \eta_1 q_1 + \eta_2 q_2$ and its normal vector (η_1, η_2). For this plots $s = 0.6$, $m = 1$, and $n = 2$.

5.2.2. Parametrization. A more quantitative analysis requires finding a convenient parametrization of the constraint (5.2.3). To this end, we set $\alpha = 1$ and write $\sqrt{q_i} = \sin\theta_i$ for $0 \leq \theta_i \leq \pi/2$. By further introducing the variables $x = \cos(\theta_1 + \theta_2)$ and $y = \cos(\theta_1 - \theta_2)$ we manage to linearize the curve (5.2.3), which now is the straight segment

$$2s^m = (1 + s^n)y - (1 - s^n)x \tag{5.2.4}$$

with $|x| \leq y \leq 1$. Its guiding vector is readily seen to be $(1 + s^n, 1 - s^n)$, so the segment's parametric equation can be written as

$$x = \frac{1 - (1 + s^n)t}{s^{n-m}}, \qquad y = \frac{1 - (1 - s^n)t}{s^{n-m}}, \tag{5.2.5}$$

where we have rescaled the guiding vector so that Eqs. (5.2.10) and (5.2.14) below are simplest. The explicit success probabilities in terms of trigonometric functions and the parameter t are

$$\sqrt{p_1} = \cos[\frac{1}{2}\arccos\frac{1-(1+\mathrm{s^n})\mathrm{t}}{\mathrm{s^{n-m}}} + \frac{1}{2}\arccos\frac{1-(1-\mathrm{s^n})\mathrm{t}}{\mathrm{s^{n-m}}} \tag{5.2.6}$$

$$\sqrt{p_2} = \cos[\frac{1}{2}\arccos\frac{1-(1+\mathrm{s^n})\mathrm{t}}{\mathrm{s^{n-m}}} - \frac{1}{2}\arccos\frac{1-(1-\mathrm{s^n})\mathrm{t}}{\mathrm{s^{n-m}}}] \tag{5.2.7}$$

which we may write more succinctly as

The bounds for t for this solution are

$$\frac{1-s^{n-m}}{1-s^n} \leq t \leq \min\frac{1}{1-\mathrm{s^n}}, \frac{1+\mathrm{s^{n=m}}}{1+\mathrm{s^n}}. \tag{5.2.8}$$

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 (5.2.3) under this transformation are mapped onto the same straight line (5.2.5). 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. (5.2.3) in parametric form as

$$q_i = \frac{1-xy-(-1)^i\sqrt{1-x^2}\sqrt{1-y^2}}{2}, \quad i = 1, 2. \tag{5.2.9}$$

Fig. 5.2.2 shows examples of the unitary curve (5.2.3).

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.

5.2.3. 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 $Q_{\min}(\eta_1)$.

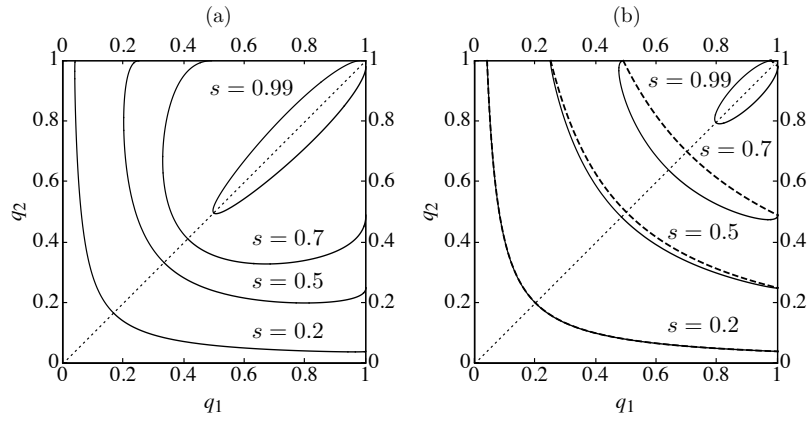


FIGURE 5.2.2. Unitarity curves, q_2 vs. q_1, from Eq. (5.2.3) for different values of s and for (a) $m = 1$, $n = 2$ and (b) $m = 1$, $n = 5$. The curves are symmetric under mirror reflection along the dotted line $q_1 = q_2$, i.e., under the transformation $q_1 \leftrightarrow q_2$. The dashed lines in (b) are the hyperbolae $q_1 q_2 = s^{2m}$.

Along with the complete description of the unitary curve (5.2.3), this provides a complete solution of the problem in parametric form.

With no loss of generality we may assume $\eta_1 \leq \eta_2$ or, equivalently, $0 \leq \eta_1 \leq 1/2$. Then the slope of the straight line (5.2.1), $-\eta_1/\eta_2$, satisfies $-1 \leq -\eta_1/\eta_2 \leq 0$. Hence, it can only become tangent to the lower half of the unitarity curve (5.2.3) (see Fig. 5.2.2). 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 $q_1 = 1$ (assuming n is finite). This follows from the properties (a)–(f) above and can be checked using Eq. (5.2.9). The values of t at which the slope is -1 and 0 are, respectively,

$$t_{-1} = \frac{1 - s^{n-m}}{1 - s^n}, \quad t_0 = \frac{1 - s^{2(n-m)}}{1 - s^{2n}}. \tag{5.2.10}$$

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 $t = t_0$. It is convenient to represent this class of lines with slope $-\frac{\eta_1}{\eta_2}$ that represent the failure rate Q by using its normal vector $\vec{n} = (\eta_1, \eta_2)$. From this set of admissible solutions, the minimum Q is the is the line in this class that intersects the unitarity curve once, and hence is tangent to it. This

tangency can also be represented by a normal vector,

$$\vec{t} = (\frac{-dq_2}{dt}, \frac{dq_1}{dt}) \tag{5.2.11}$$

For optimality we know the two tangent vectors must be parallel: $\hat{n} = \hat{t}$.

We can express the line tangent to the unitary curve as $\eta_1 q_1' + \eta_2 q_2' = 0$, where the primed variables are derivatives with respect to t. Using the constraint $1 = \eta_1 + \eta_2$ we have

$$\eta_1 = \frac{q_2'}{q_2' - q_1'}, \quad \eta_2 = \frac{q_1'}{q_1' - q_2'}. \tag{5.2.12}$$

We could solve these equations for $t(\eta_1)$ and substitute into q_1 for the explicit expression if it were not a high order expression. Numerically this is easily achieved.

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. (5.2.1) with (5.2.9) we enforce contact with the unitarity curve and obtain the expression of $Q_{\min}$. The final result can be cast as:

$$\eta_1 = \frac{q_2'}{q_2' - q_1'}, \;\; Q_{\min} = \frac{q_2' q_1 - q_1' q_2}{q_2' - q_1'}, \;\; t_{-1} \le t \le t_0, \tag{5.2.13}$$

where t_{-1}, t_0 and q_i are given in Eqs. (5.2.10) and (5.2.9). The expressions for the derivatives q_i' are

$$q_i' = \frac{\sqrt{q_i(1-q_i)}}{s^{n-m}} \left\{ \frac{1+s^n}{\sqrt{1-x^2}} - (-1)^i \frac{1-s^n}{\sqrt{1-y^2}} \right\}. \tag{5.2.14}$$

5.2.4. Relation of perfect cloning with discrimination. Fig. 5.2.3 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. (5.2.10), into Eq. (5.2.9). They are given by

$$Q_0 = q_2(t_0) = \frac{s^{2m} - s^{2n}}{1 - s^{2n}}, \quad Q_{-1} = \frac{s^m - s^n}{1 - s^n}, \tag{5.2.15}$$

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

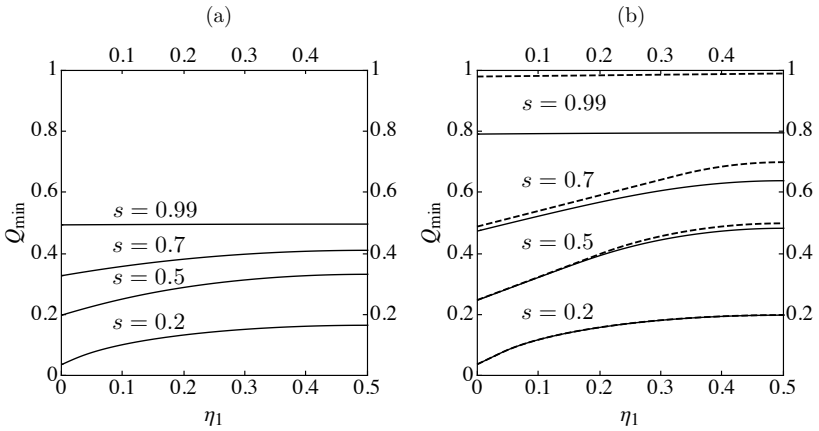


FIGURE 5.2.3. Minimum cloning failure probability $Q_{\min}$ vs. η_1 (solid lines) and UD failure probability $Q_{\rm UD}$ vs. η_1 (dashed lines) for the same values of m, n and s used in the previous figure.

dashed lines in Fig. 5.2.3 (b) depict the well known optimal UD solution [**2**]:

$$Q_{\rm UD} = \begin{cases} 2\sqrt{\eta_1\eta_2}\, s^m, & \dfrac{s^{2m}}{1+s^{2m}} \le \eta_1 \le \dfrac{1}{2}; \\ \eta_1 + s^{2m}\eta_2, & 0 \le \eta_1 \le \dfrac{s^{2m}}{1+s^{2m}}. \end{cases} \tag{5.2.16}$$

It is apparent from these plots that the optimal cloning protocol performs strictly better than cloning by discrimination, as was geometrically proved in Figs. 5.2.1 and 5.2.2. However, the difference in performance decreases with increasing number of clones. In Fig. 5.2.3 (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, $Q_{\min}$, to that of cloning by discrimination, $Q_{\rm UD}$ in Eq. (5.2.16) follows from our geometrical approach. Recall that in the limit $n \to \infty$ (or equivalently $\alpha \to 0$) the right hand side of Eq. (5.2.3) describes hyperbolas that we can write as $q_2 = s^{2m}/q_1$ (dashed lines in Figs. 5.2.1 and 5.2.2). Their slopes are in the range $[-1, -s^{2n}]$. A unique point of tangency with the line (5.2.1) 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 (5.2.16), and one can easily obtain the corresponding expression for $Q_{\rm UD}$. If the slope

of the line (5.2.1) 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. (5.2.16). In geometrical terms, the straight line (5.2.1) pivots on the end points as we vary η_1. Furthermore, note that for the second line in (5.2.16) 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. (5.2.2).

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 (or $\alpha \to 0$) 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 second line in Eq. (5.2.16). A similar phenomenon arises in UD of more than two pure states [**52**].

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. (5.2.15) 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. (5.2.16) 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. (5.2.16). 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 bottom line in Eq. (5.2.16) applied to n copies. The total failure rate is s^{2m}, also achieving optimality.

Using our main result in Eqs. (5.2.13), and (5.2.9) 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.

5.3. Interpolative Cloning

In this section [**53**] we 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 than maximum fidelity in the approximate cloning machine while it allows for a fixed rate of inconclusive outcomes. This scheme should reproduce exact cloning and approximate cloning machines by setting $Q = Q_o$ and$Q = 0$ respectively. Chefles and Barnett [**40**] 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.

5.3.1. Equal priors. First we show the derivation of the soltuion for the equal priors [**40**] 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} reduces to:

$$\begin{aligned} F_{MN} &= \frac{1}{2}\left[1 + \sqrt{1 - \sin^2(2\theta - 2\phi)}\right], \\ &= \frac{1}{2}\left[1 + \cos^2(2\theta - 2\phi)\right]. \end{aligned} \tag{5.3.1}$$

Duan and Guo [**31**] 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:

$$P_{MN} = \frac{1 - s^M}{1 - s^N}, \tag{5.3.2}$$

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:

$$P_{12} = \frac{1}{1+s}. \tag{5.3.3}$$

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:

$$U|\psi_1\rangle^{\otimes M}|i\rangle = \sqrt{p}|\phi_1\rangle^{\oplus N}|1\rangle + \sqrt{q}|f\rangle|0\rangle \tag{5.3.4}$$

$$U|\psi_2\rangle^{\oplus M}|i\rangle = \sqrt{p}|\phi_2\rangle^{\oplus N}|1\rangle + \sqrt{q}|f\rangle|0\rangle \tag{5.3.5}$$

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 (5.3.4) and (5.3.5) gives the constraint $s^M = ps'^N + q$. Using the unitarity condition $p + q = 1$, the average rate of successfully making a clone is $p = (1 - s^M)/(1 - s'^N)$. s' is again 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 $\cos 2\theta = |\langle\psi_1|\psi_2\rangle|^N$

$$|\langle\phi_1|\phi_2\rangle|^N = 1 - \frac{1 - |\langle\psi_1|\psi_2\rangle|^M}{P_S}, \tag{5.3.6}$$

$$\cos^N(\phi_1 + \phi_2) = 1 - \frac{P_{IDP}}{P_S}. \tag{5.3.7}$$

The exact clones live in an N dimensional space $|\psi^N_{1,2}\rangle = \cos\theta|1\rangle \pm \sin\theta|0\rangle$. The approximate clones can be expressed as $|\phi_{1,2}\rangle = \cos\phi_{1,2}|1\rangle \pm \sin\phi_{1,2}|0\rangle$. The fidelity rate for equal priors is:

$$F_{MN} = \frac{1}{2}\left[1 + \cos(2\theta - (\phi_1 + \phi_2)\right], \tag{5.3.8}$$

and we want to use the relationship in (5.3.7). 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

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

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 maximum failure rate is $Q_o = |\langle\psi_1|\psi_2\rangle|$. This allows us to write the fidelity expression as

$$F_{MN} = \frac{1}{2(1-Q)}\left[(1-Q) + Q_o^N(Q_o - Q) + \sqrt{(1 - Q_o^{2N})\left[(1-Q)^2 - (Q - Q_o)^2\right]}\right].$$

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

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

$(1-Q)F_{MN} = P_{success}$, 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, which was defined as the rate of successfully carrying out the 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.

5.3.2. 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.

$$\begin{aligned} U|\psi_1^M\rangle|0\rangle &= \sqrt{p_1}|\phi_1\rangle|1\rangle + \sqrt{q_1}|\Phi\rangle|2\rangle, \\ U|\psi_2^M\rangle|0\rangle &= \sqrt{p_2}|\phi_2\rangle|1\rangle + \sqrt{q_2}|\Phi\rangle|2\rangle, \end{aligned} \tag{5.3.9}$$

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

$$s = \sqrt{p_1 p_2}s' + \sqrt{q_1 q_2} \tag{5.3.10}$$

- 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 $\tilde{\eta}_i = (\eta_i p_i)/(1 - Q)$. The average fidelity is the same as calculated previously with normalized probabilities:

$$\begin{aligned} F_{MN} &= \frac{1}{2}\left[1 + \sqrt{1 - 4\tilde{\eta}_1\tilde{\eta}_2 \sin^2\left(2\theta - (\phi_1 + \phi_2)\right)}\right], \\ &= \frac{1}{2(1-Q)}\left[(1-Q) + \sqrt{(1-Q)^2 - 4\eta_1\eta_2 p_1 p_2 \sin^2\left(2\theta - (\phi_1 + \phi_2)\right)}\right] \end{aligned} \tag{5.3.11}$$

It can be seen that in the limit $N \to \infty$, expanding the sin term as we did in the previous section, the IM [**23**] 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

$$\Lambda = \sqrt{p_1 p_2}\sin\left(2\theta - (\phi_1 + \phi_2)\right),$$

where $\cos(\phi_1 + \phi_2) = s'$, the output overlap of the separated cloned states. We can write the optimal value of Λ in terms of the constrained variables of state separation $u \equiv \sqrt{q_1 q_2}$, $v \equiv \frac{1}{2}(q_1 + q_2)$, which will be discussed further in the following chapter, as

$$\Lambda_{\text{sep}} = \sqrt{1 - s^{2n}}\,(s - u) - s^n\sqrt{1 - s^2 - 2v + 2su}. \tag{5.3.12}$$

where we used the unitarity constraint from separation $\sqrt{p_1 p_2}s' = s - \sqrt{q_1 q_2}$. We can re-write our equation for Λ_{sep} in terms of v as

$$v = su + \frac{1 - s^2}{2} - \frac{1 - s^{2n}}{2s^{2n}}\left(u - s + \frac{\Lambda_{\text{sep}}}{\sqrt{1 - s^{2n}}}\right)^2. \tag{5.3.13}$$

Which is clearly the equation for a parabola in the v vs u plane. As Λ_{sep} varies, the parabolas are bounded by the straight line $v = su + (1 - s^2)/(2)$.

Furthermore, we can cast the failure constraint $Q = \eta_1 q_1 + \eta_2 q_2$ as an ellipse in the same plane with the parametrization

$$\begin{aligned} u &= \frac{Q}{\sqrt{1 - \Delta^2}}\cos\theta', \\ v &= \frac{Q}{1 - \Delta^2} + \frac{Q\Delta}{1 - \Delta^2}\sin\theta', \end{aligned} \tag{5.3.14}$$

The optimal point comes from the value of Λ_{sep} that allows the ellipse to lie tangent to the parabola. The solution when both are equal is the root of a high-order equation and cannot be written explicitly. However it can be cast parametrically as

$$\begin{aligned} Q &= \frac{(1-\Delta^2)(1-s^2) - \gamma_n\left(\Delta\cot\theta' + s\sqrt{1-\Delta^2}\right)^2}{2\left(1 + \Delta\sin\theta' - s\sqrt{1-\Delta^2}\cos\theta'\right)}, \\ \Lambda_{\text{min}} &= \frac{(1+\gamma_n)\sqrt{1-\Delta^2}s + \gamma_n\Delta\cot\theta' - Q\cos\theta'}{\sqrt{1+\gamma_n}\sqrt{1-\Delta^2}}, \end{aligned} \tag{5.3.15}$$

where $\gamma_n = s^{2n}/(1-s^{2n})$ and $\Delta = \eta_2 - \eta_1$. This completes the solution to the approximate probabilistic cloning problem. We graphically demonstrate the solution below with a plot showing the optimal fidelities versus failure rate for a particular choice of parameters, and demonstrate the convergence graphically of the fidelity to the intermediate discrimination limit.

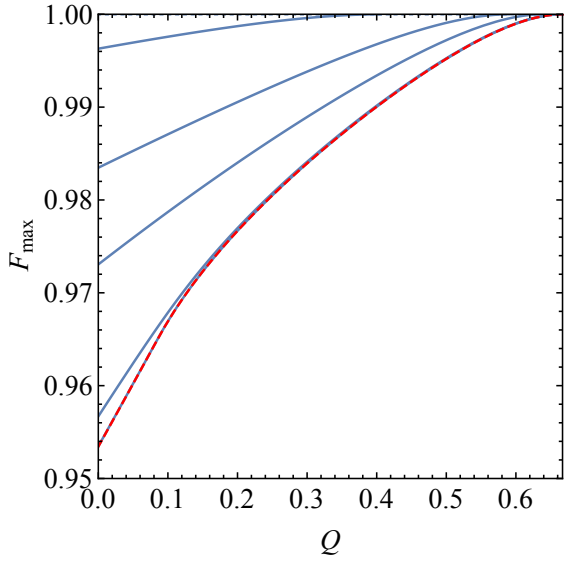


FIGURE 5.3.1. The maximum global fidelity F_{max} for a given failure rate Q for values of $n = 2, 4, 6, 14$ and 35 in blue. In red is the intermediate discrimination (IM) solution 2.4.2 nearly directly on the $n = 35$ line. Plotted for $s = .8$, $\eta_1 = .007$, $\eta_2 = 0.93$.

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. In their letter, Reck *et al.* **[54]** showed that any NxN unitary matrix can be reduced to $N(N-1)/2$ beamsplitters acting on N degrees of freedom. If we allow single photons through a set of beamsplitters to N independent paths, a single photon interferometry device with an appropriate number of optical elements can represent the action of any unitary matrix. A main benefit from seeking a solution using the Neumark setup is that it lends itself to quickly solving for the matrix elements of the unitary, thereby allowing us to express our results as experimental setups akin to those previously conducted **[14, 16]**.

For example, for interpolative discrimination we had three measurement outcomes, meaning dimensionality three, while the pure states can be represented as qubits, thereby requiring a total of three rails, one of which is ancillary and starts in the vacuum state. The three income ports 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

We use a strategy similar to that developed by J.A Bergou *et al.* **[55]**, 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.* 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.* [**56**] : 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:

$$\begin{aligned}
\cos\omega_1 &= \sqrt{\frac{r_1}{1-q_1}}, \quad \sin\omega_1 = \sqrt{\frac{p_1}{1-q_1}}, \\
\cos\omega_2 &= \sqrt{q_1}, \qquad\quad \sin\omega_2 = \sqrt{1-q_1}, \\
\cos\omega_3 &= -\frac{1}{\sqrt{1-q_1}}\left[\frac{\sqrt{q_2} - \sqrt{q_1}\cos\theta}{\sin\theta}\right], \\
\sin\omega_3 &= \frac{\sqrt{\sin^2\theta - q_1 - q_2 + 2\sqrt{q_1 q_2}\cos\theta}}{\sqrt{1-q_1}\sin\theta}.
\end{aligned}$$

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

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

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

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

An attractive simplification can be achieved by setting $\eta_1 = \eta_2$, the equal a-priori 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 IM solution this matrix minimizes the error rate and maximizes the rate of success. Hence, by setting the failure rate 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 make the states orthogonal, 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. (4.1.2) and (4.1.3), 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.

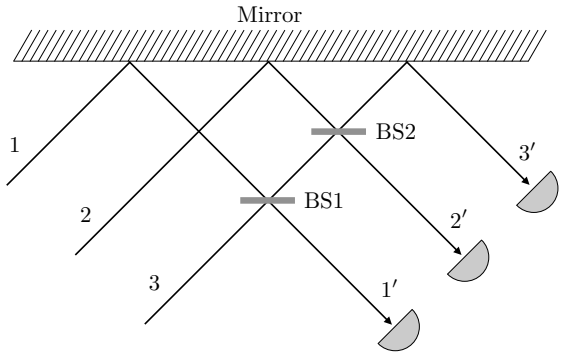


FIGURE 6.3.1. Six-port linear optics implementation of a proof-of-principle separation protocol. The transmission (reflexion) coefficients of the beam-splitters, BS1 and BS2 are given by the (off-)diagonal entries of the matrices in Eqs. (6.3.4) and (6.3.5), respectively. The input states are feed through 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. (5.2.3) can be solved explicitly. The solution is given by $Q = Q_{-1}$ in Eq. (5.2.15). 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 [**?, 55**] 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 $1 \to 2$ 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 input 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 since now we are working with 4x4 and 3x3 matrices. It was realized a shorter realization of separation exists for a different choice of basis states. We demonstrate 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 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 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

$$\langle 0|U|\psi_{1,2}\rangle = c_1 U_{01} \pm s_1 U_{02} = \sqrt{q_{1,2}} \tag{6.4.3}$$

$$\langle 1|U|\psi_{1,2}\rangle = c_1 U_{11} \pm s_1 U_{12} = c_2\sqrt{p_{1,2}} \tag{6.4.4}$$

$$\langle 2|U|\psi_{1,2}\rangle = c_1 U_{21} \pm s_1 U_{22} = \pm s_2\sqrt{p_{1,2}} \tag{6.4.5}$$

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

$$[U] = \begin{pmatrix} \bullet & \frac{\sqrt{q_1}+\sqrt{q_2}}{2c_1} & \frac{\sqrt{q_1}-\sqrt{q_2}}{2s_1} \\ \bullet & \frac{c_2(\sqrt{p_1}+\sqrt{p_2})}{2c_1} & \frac{c_2(\sqrt{p_1}-\sqrt{p_2})}{2s_1} \\ \bullet & \frac{s_2(\sqrt{p_1}-\sqrt{p_2})}{2c_1} & \frac{s_2(\sqrt{p_1}+\sqrt{p_2})}{2s_1} \end{pmatrix}. \tag{6.4.6}$$

Applying the unitarity constraint $U^\dagger U = I$ can us nine equations for the remaining three unknown 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

$$[U] = \begin{pmatrix} \bullet & \frac{\sqrt{s_2^2 - s_1^2}}{c_1 s_2} & 0 \\ \bullet & \frac{c_2 s_1}{c_1 s_2} & 0 \\ \bullet & 0 & 1 \end{pmatrix}. \tag{6.4.7}$$

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

$$[U] = \begin{pmatrix} \frac{c_2 s_1}{c_1 s_2} & \frac{\sqrt{s_2^2 - s_1^2}}{c_1 s_2} & 0 \\ -\frac{\sqrt{s_2^2 - s_1^2}}{c_1 s_2} & \frac{c_2 s_1}{c_1 s_2} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{6.4.8}$$

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

$$U|\phi_1\rangle|0\rangle = |\xi_1\rangle|\xi_1\rangle \tag{6.4.9}$$

$$U|\phi_2\rangle|0\rangle = |\xi_2\rangle|\xi_2\rangle \tag{6.4.10}$$

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

$$[U] = \begin{pmatrix} \frac{c_3^2}{c_2} & 0 & 0 & \frac{s_3^2}{c_2} \\ 0 & \frac{c_3 s_3}{s_2} & -\frac{c_3 s_3}{s_2} & 0 \\ 0 & \frac{c_3 s_3}{s_2} & \frac{c_3 s_3}{s_2} & 0 \\ \frac{s_3^2}{c_2} & 0 & 0 & \frac{c_3^2}{c_2} \end{pmatrix} \tag{6.4.11}$$

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

$$[M_{14}] = \begin{pmatrix} \frac{c_3^2}{c_2} & 0 & 0 & \frac{s_3^2}{c_2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{s_3^2}{c_2} & 0 & 0 & \frac{c_3^2}{c_2} \end{pmatrix}, \quad [M_{23}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{6.4.12}$$

This implementation performs the action of the necessary unitary but it does not provide us with two distinct, separable clones as we may desire for practical applications. The two qubits are in fact hidden in the statistics of the four output rails. If we perform the measurement many times and do photon statistics for the output ports we should find that they satisfy c_3^2, $c_3 s_3$, $c_3 s_3$ and s_3^2 respectively from rail 1 to rail 4. If we wish to create two distinct qubits on a single photon we may use both the polarization and location qubits as done in [**57**]. Here the polarization qubit was copied onto the location of the photon deterministically as a universal cloner, but the extension to known states is straightforward.

Bibliography

[1] M. A. Neumark, Izv. Akad. Nauk. SSSR, Ser. Mat. **4:3**, 277-318 (1940).

[2] J. A. Bergou, J. Mod. Opt. **57**, 160 (2010).

[3] A. Chefles, Contemp. Phys. **41**, 401-424 (2000).

[4] B. L. Higgins, B. M. Booth, A. C. Doherty, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde, Phys. Rev. Lett. **103**, 220503 (2009).

[5] A. Chefles, Phys. Lett. A **239**, 339-347 (1998).

[6] C. W. Helstrom,*Quantum Detection and Estimation Theory,* Academic Press, New York, 1976.

[7] M. Ban, K. Kurokawa, R. Momose, and O. Hirota, Int. J. Theor. Phys. **55**, 22 (1997).

[8] Y. C. Eldar, A. Megretski, and G. C. Verghese, IEEE Trans. Inf. Theory **IT-49**, 1007 (2003).

[9] C.-L. Chou and L. Y. Hsu, Phys. Rev. A **68**, 042305 (2003).

[10] S. M. Barnett and E. Riis, J. Mod. Opt. **44**, 1061 (1997).

[11] R. B. M. Clarke *et al.*, Phys. Rev. A **64**, 012303 (2001).

[12] I.D. Ivanovic, Phys. Lett. A **123**, 257-259 (1987).

[13] G. Jaeger and A. Shimony, Phys. Lett. A **197**, 83 (1995).

[14] B. Huttner *et al.*, Phys. Rev. A **54**, 3783 (1996).

[15] R. B. M. Clarke *et al.*, Phys. Rev. A **63**, 040305(R) (2001).

[16] Masoud Mohseni, Aephraim M. Steinberg, and János A. Bergou, Phys. Rev. Lett.**93**, 200403, (2004).

[17] S. Croke, E. Andersson, S. M. Barnett, Claire R. Gilson, and John Jeffers, Phys. Rev. Lett. **96**, 070401 (2006).

[18] U. Herzog and O. Benson, J. Mod. Opt.**57**, 188 (2010).

[19] G. A. Steudle *et al.*, Phys. Rev. A **83**, 050304(R) (2011).

[20] A. Chefles, and S. M. Barnett, Journal of Mod. Opt. Vol. **45**, 1295-1302 (1998).

[21] J. Fiurasek and M. Jezek, Phys. Rev. A **67,** 012321 (2003).

[22] H. Sugimoto, T. Hashimoto, M. Horibe, and A. Hayashi, Phys. Rev. A **80**, 052322 (2009).

[23] E. Bagan, R Munoz-Tapia, G.A. Olivares-Rentería, and J.A. Bergou, Phys. Rev. A **86**, 040303 (2012).

[24] U. Herzog, Phys Rev. A **85**, 032312 (2012).

[25] U. Herzog, Phys Rev. A **86**, 032314 (2012).

[26] V. Yerokhin, A. Shehu, J. A. Bergou, *submitted for review in Phys. Rev. A.*

[27] C. A. Fuchs and A. Peres, Phys. Rev. A **53**, 20238 (1996).

[28] U. Herzog, Phys Rev A. **79** 032323 (2009).

[29] A. Chefles and S. M. Barnett, J. Phys. A: Math. Gen. **31** 10097 (1998).

[30] E. Bagan, V. Yerokhin, A. Shehu, E. Feldman, J. A. Bergou arXiv:1506.08241 *submitted for review in New Jour. of Phys.*

[31] L. M. Duan and G. C. Guo, Phys. Rev. Lett. **80**, 4999 (1998).

[32] W. K. Wooters and W. H. Zurek, Nature **299**, 802 (1982).

[33] D. Dieks, Phys. Lett. A **92**, 271 (1982).

[34] E. Pomarico, B. Sanguinetti, P. Sekatski, H. Zbinden, and N. Gisin, Optics and Spectroscopy **111**, 510 (2011).

[35] K. Bartkiewicz, A. Černoch, K. Lemr, J. Soubusta, and M. Stobinska, Phys. Rev. A **89**, 062322 (2014).

[36] V. Bužek and M. Hillery, Phys. Rev. A. **54**, 1844 (1996).

[37] N. Gisin and S. Massar, Phys. Rev. Lett. **79**, 2153 (1997).

[38] V. Bužek and M. Hillery, Phys. Rev. Lett. **81**, 5003 (1998).

[39] D. Bruß, D. P. DiVincenzo, A. Ekert, C. A. Fuchs, C. Macchiavello, and J. A. Smolin, Phys. Rev. A **57**, 2368 (1998).

[40] A. Chefles and S. M. Barnett, Phys. Rev. A. **60**, 136 (1999).

[41] J. Fiurášek, S. Iblisdir, S. Massar, and N. J. Cerf, Phys. Rev. A **65**, 040302(R) (2002).

[42] J. Fiurášek, Phys. Rev. A. **70**, 032308 (2004).

[43] C. R. Müller, C. Wittmann, P. Marek, R. Filip, C. Marquardt, G. Leuchs, and U. L. Andersen, Phys. Rev. A **86**, 010305(R) (2012).

[44] H. Barnum, J. Barrett, M. Leifer, and A. Wilce, Phys. Rev. Lett. **99**, 240501 (2007).

[45] G. Chiribella, Y. Yang, and A. Yao, Nature Comm. **4**, 2915 (2013).

[46] J. Bae and A. Acin, Phys. Rev. Lett. **97**, 030402 (2006).

[47] G. Chiribella and G. M. D'Ariano, Phys. Rev. Lett. **97**, 250503 (2006).

[48] B. Gendra, J. Calsamiglia, R. Muñoz-Tapia, E. Bagan, and G. Chiribella, Phys. Rev. Lett. **113**, 260402 (2014).

[49] V. Scarani, S. Iblisdir, N. Gisin, and A. Acin, Rev. Mod. Phys. **77**, 1225 (2005).

[50] H. Fan, Y. N. Wang, L. Jing, J. D. Yue, H. D. Shi, Y. L. Zhang, and L. Z. Mu, Phys. Rep. **544**, 241 (2014).

[51] V. Yerokhin, A. Shehu, E. Feldman, E. Bagan, J. A. Bergou, arXiv:1505.06979, *submitted for review in Phys. Rev. Lett.*

[52] J. A. Bergou, U. Futschik, and E. Feldman, Phys. Rev. Lett. **108**, 250502 (2012).

[53] E. Bagan, V. Yerokhin, A. Shehu, E. Feldman, J. A. Bergou *in preparation.*

[54] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Phys. Rev. Lett. **73** 58, (1994).

[55] J. A. Bergou, M. Hillery, and Y. Sun, Journal of Mod. Opt. **47**, 487-497 (2000).

[56] Y. Sun, M. Hillery, J. A. Bergou, Phys. Rev. A. **64**, 022311 (2001).

[57] Y. Huang, W. Li, C. Li, Y. Zhang, Y. Jiang, and G. Guo, Phys. Rev. A **64**, 012315 (2001).