Quantum State Discrimination and Quantum Cloning: Optimization and Implementation

by

Andi Shehu

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

2015

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

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

Supervisory Committee:

Prof. Mark Hillery

Prof. Christopher C. Gerry

Prof. Ed Fieldman

Prof. Neepa T. Maitra

THE CITY UNIVERSITY OF NEW YORK

Abstract

Quantum State Discrimination and Quantum Cloning: Optimization and Implementation

by

Andi Shehu

Advisor: János A. Bergou

In our work we explore the field of quantum state discrimination and quantum cloning. Recently the problem of optimal state discrimination with a Fixed Rate of Inconclusive Outcomes (FRIO strategy) has been solved for two pure quantum states and a few other highly symmetric cases. An optical implementation to FRIO for pure states is provided. The physical implementation can be carried out with the use of a six-port interferometer constructed with optical fibers beam splitters, phase shifters and mirrors. The input states are composed of qubits which are realized as photons in the dual-rail representation. The non-unitary measurements are carried out at the output for the presence or absence of a photon. The setup optimally interpolates between minimum error and unambiguous state discrimination. We also extend the FRIO strategy to two mixed states, whose eigenvectors in their spectral representation form a Jordan basis. We derive the minimum error rate P_E for a fixed inconclusive rate Q and, in particular, the optimal distribution of the total Q over the Jordan subspaces. As Q is varied between the two limits, $0 < Q < Q_c$, a structure with multiple thresholds, $Q_1^{(th)}(=0) < Q_2^{(th)} < \ldots < Q_N^{(th)} < Q_c$, emerges. We also solve the problem of state separation of two known pure states in the general case where the known

states have arbitrary prior probabilities. The solution emerges from a geometric formulation of the problem. This formulation also reveals a deeper connection between cloning and state discrimination. The results are then applied in designing a scheme for hybrid cloning which interpolates between approximate and probabilistic exact cloning. It is shown that state separation and hybrid cloning are generalized schemes to well established state discrimination and cloning strategies. The relationships between cloning, state separation and state discrimination are derived in several limints.

Acknowledgments

First and foremost I would like to thank my advisor Prof. János A. Bergou for guiding me throughout the years of the PhD program. It was due to his guidance, encouragement and positive attitude I was able to complete the program.

Special thanks go to the members of the research group: Vadim Yerokhin, colleague and very good friend of many years for the work and the good times we shared in the office; Prof. Emilio Bagan for helping us solve problems in ingenious ways which we at times believed to be impossible. In return we introduced him to the Knob Creek; Dr. Georgina Olivares for helping with the writing of the dissertation, several rehearsals of the defense over Skype and the words of encouragement; Dr. Ugur Guney for the IT support, setting up the LyX software this dissertation is being written on and some of the graphics used in here; Prof. Mark Hillery for organizing the seminars and the helpful discussions; Prof Emeritus. Ed Feldman for the work and the great stories he shared; Valeria Feherpataky-Bergou for the many gifts of chocolate and wine and for being the mother of the group.

I would also like the thank my good friends and colleagues Armando Rua, Jorge Colon, Kay Hiranaka,, Tetiana Nosuch, Zhenmao Wan, Denis Sh., Mallory Gobet, Phil Stallworth, Marc Berman and Joel DeJesus for making the PhD life memorable.

Many thanks to my parents Luan and Flutur, my siblings Elona, Evis and Kliton Shehu for the love, support and guidance throughout my life and education. I did become a doctor as my parents thought I would, just not the one to call in case of a medical emergency.

Contents

List of Figures

CHAPTER 1

Quantum State Discrimination

An integral part of quantum information and quantum processing is measurement theory [**1**]. It is the probabilistic nature of quantum mechanics that one cannot simply obtain information encoded in states [**2**], the state is not an observable in quantum mechanics [**3**]. When a quantum circuit or processor has acted on the input states to perform a task, the output needs to be read out. Thus after the processing occurs the task is to determine the state of the system. If the input states are orthogonal the process is trivial. Simply setting up detectors along the orthogonal directions and a click in those detectors will determine the state of the system. On the other hand discriminating among non orthogonal quantum states is not trivial. Since quantum mechanics does not allow for perfect discrimination of non orthogonal states the task becomes that of a measurement optimization problem. Not being able to perfectly discriminate quantum states is key to various quantum cryptographic schemes and quantum computing. The origin of the state discrimination field is attributed to the works of Helstrom [**4**] and Holevo [**5**]. The field however gained momentum in the 90's as quantum information theory became very active primarily due to the factorization work of Peter Schor [**6**] and quantum key distribution protocols such as B92 [**7**].

Various optimum state discrimination measurement strategies have been developed with respect to some figure of merit. Two of those methods which we focus on are optimum Unambiguous Discrimination (UD) and Minimum Error (ME). In UD strategy, first suggested by Ivanovic [**8**], the observer Bob, is not allowed to make an error. Whenever he is handed a state $|\psi_i\rangle$ he cannot conclude that he was given $|\psi_j\rangle$. We will show that this cannot be done with 100% success rate and that the observer must allow for inconclusive results and find an optimum measurement strategy which minimizes the average rate of inconclusive results. In the Minimum Error strategy the observer is not allowed to have inconclusive results. Thus

errors are allowed and the task is to find optimum measurements that minimize the average error rate. It has been shown that ME and UD are special cases of a more general scheme of optimum state discrimination measurement which can be approached by relaxing the conditions at either end [**9**]. In the ME scheme the optimal error rate can be further reduced by allowing for some rate of inconclusive results. Thus the optimal average error rate, P_E, becomes a function of a given rate of allowed inconclusive results Q, $P_E(Q)$. On the other hand, in UD, the optimal rate of the average inconclusive outcomes, Q, may be reduced by allowing for some error rate P_E. The failure rate becomes a function of a given error rate $Q(P_E)$.

In our work we use various quantum measurements schemes to read out information out of a quantum system. For a more thorough understanding of quantum theory of measurements we go along the lines of the review paper by J.A Bergou [**10**]. Starting with the standard quantum measurement theory due essentially to von Neumann the generalized measurements (Positive Operator Valued Measures, POVMs) are introduced as more useful measurement schemes in optimization problems. Using Neumark's theorem the POVMs can be realized experimentally.

1.0.1. Standard Quantum Measurements. We start with the postulates of standard or projective quantum measurements introduced by von Neumann [**11**] analyzing a model for the coupling of the system with the meter or ancilla and generalizing the predictions of the model.

The postulates are:

(1) Observables in quantum mechanics have a Hermintian operator χ which has a spectral representation $\chi = \sum_j^N \lambda_j |j\rangle \langle j|$, where the eigenvalues are real and assuming non-degeneracy for simplicity. The eigenvectors $\{|j\rangle\}$ form a complete orthonormal basis set.

(2) The Hilbert space is spanned by the projectors $P_j = |j\rangle \langle j|$, such that $\sum_j P_j = 1$.

(3) The eigenvalues of the projectors are 0 or 1 due to the orthogonality of the states $P_i P_j = P_i \delta_{ij}$.

(4) Any measurement of the χ will yield one of the eigenvalues λ_j.

(5) If λ_j is obtained in a measurements, the state of the system collapses onto: $|\phi_j\rangle = \frac{P_j|\psi\rangle}{\sqrt{\langle\psi|P_j|\psi\rangle}}$ if the system was initially in a pure state, $\rho_j = \frac{P_j\rho P_j}{Tr(P_j\rho)}$ if the system was initially in a mixed state.

(6) The probability of obtaining $|\phi_j\rangle$ is $p_j = ||P_j\psi||^2 = \langle\psi\,|P_j^2|\,\psi\rangle = \langle\psi\,|P_j|\,\psi\rangle$. The probability of obtaining ρ_j is $p_j = Tr(P_j\rho P_j) = Tr(P_j^2\rho) = Tr(P_j\rho)$.

(7) If a measurement is performed but the result is not recorded the post-measurement state collapses onto: $\rho = \sum_j P_j\,|\psi\rangle\,\langle\psi|\,P_j$ if the system was initially in a pure state, $\tilde{\rho} = \sum_j p_j\rho_j = \sum_j P_j\rho P_j$ if the system was initially in a mixed state.

1.0.2. POVMs. Due to the orthogonality condition of the projective measurements one cannot have more orthogonal projections than the dimensionality, hence the possible outcomes cannot exceed the number of the dimensionality. Sometimes we would like to allow for more outcomes than the dimensionality, as in the case of optimal UD measurements where we have three outcomes in a two dimensional problem.

Next we introduce a positive operator $\Pi_j \geq 0$ as a generalization of P_j^2 and the probability of obtaining state j becomes $p_j = Tr(\Pi_j\rho_j)$. To normalize the probabilities we require that the positive operators Π_j are a decomposition of the identity $\sum_j \Pi_j = I$. This is decomposition is called a Positive Operator Valued Measure (POVM) and Π_j the elements of the POVM.

The generalization of the postulates of quantum mechanics in terms of the POVM can be expressed as:

(1) The decomposition of the identity in terms of positive operators, $\Pi_j \geq 0$, $\sum_j \Pi_j = I$ is called a POVM.

(2) The elements of the POVM can be expressed in terms of the detection operators $\Pi_j = A_j^\dagger A_j$ where the operators satisfy the requirements $\sum_j A_j^\dagger A_j = I$ but they need not be Hermitian.

(3) A detection yields an element on POVM.

(4) The state of the system collapses onto: $|\phi_j\rangle = \frac{A_j|\psi\rangle}{\sqrt{\langle\psi|A_j^\dagger A_j|\psi\rangle}}$ if the system was initially in a pure state, $\rho_j = \frac{A_j^\dagger \rho A_j}{Tr(A_j \rho A_j^\dagger)} = \frac{A_j^\dagger \rho A_j}{Tr(A_j^\dagger A_j \rho)} = \frac{A_j^\dagger \rho A_j}{Tr(\Pi_j \rho)}$ if the system was initially in a mixed state.

(5) The probability of obtaining ρ_j is $p_j = Tr(A_j \rho_j A_j^\dagger) = Tr(A_j^\dagger A_j \rho_j) = Tr(\Pi_j \rho_j)$.

(6) If a measurement is performed but the result is not recorded the post-measurement state collapses onto: $\tilde{\rho} = \sum_j p_j \rho_j = \sum_j A_j \rho A_j^\dagger$.

It is these generalized measurements we will use in our optimization work. In the following sections POVM elements are used to optimize the unambiguous discrimination and minimum error schemes.

1.1. Unambiguous Discrimination

In this section we give a review of the existing schemes of Unambiguous Discrimination (UD). Particularly that of two pure states as it is directly related with our work. When performing UD the detectors are not allowed to make an error but can admit inconclusive outcomes. We first show by contradiction that it is not possible to succeed at unambiguously discriminating quantum states with 100% success rate. Then we show that in order to perform UD a third detector must be added which accounts for inconclusive results. The task is to minimize this rate of inconclusive outcomes. In Subsection (1.1.1) the problem is solved via the POVM strategy. In the following Subsection (1.1.2) we show how the solution can be implemented via the Neumark theorem.

1.1.1. Unambiguous Discrimination: Two pure states via POVM. An ensemble of quantum states is prepared with two possible pure states $|\psi_1\rangle$ or $|\psi_2\rangle$. Each state is prepared with an a priori probability η_1 or η_2, such that $\eta_1 + \eta_2 = 1$. The observer has full knowledge of the states and their priors. The preparer, Alice, picks up a state and hands it over to the observer, Bob. Bob's task it to determine which state he is given by performing a single a POVM on the individual system he is given.

As it was stated earlier, the observer is not allowed to make an error when performing a measurement. Let us assume Bob can indeed discriminate the given states with 100%

success rate. Let Π_1 and Π_2 be detectors which cover the full Hilbert space spanned by the states $|\psi_1\rangle$ and $|\psi_2\rangle$,

$$\Pi_1 + \Pi_2 = I \tag{1.1.1}$$

In the UD strategy the detector Π_i identifies only the state $|\psi_i\rangle$ and never clicks for $|\psi_j\rangle$, such that $\Pi_i|\psi_j\rangle = 0$. Multiplying Equation (1.1.1) by $|\psi_1\rangle$ from the right and $\langle\psi_1|$ from the left results in $p_1 = \langle\psi_1|\Pi_1|\psi_1\rangle = 1$, which is the probability of successfully identifying $|\psi_1\rangle$. Similarly it can be shown that the state $|\psi_2\rangle$ can be detected with a probability one, $p_2 = \langle\psi_2|\Pi_2|\psi_2\rangle = 1$. Seems as if one can indeed discriminate two non-orthogonal quantum states with a 100% success rate. However multiplying Equation. (1.1.1) with $\langle\psi_1|$ from the left and $|\psi_2\rangle$ from the right it follows that $\langle\psi_1|\psi_2\rangle = 0$, where we use $\Pi_i|\psi_j\rangle = 0$. This means that the input states are orthogonal to begin with, which is a contradiction because we started with nonorthogonal quantum states. Thus one cannot discriminate non-orthogonal quantum state with 100% success rate (orthogonal states can indeed be discriminated with no error rate, they correspond to classical states).

One can still perform Unambiguous Discrimination but with a modified scheme. Equation (1.1.1) is modified by adding a third detector Π_0 which can click for both states $|\psi_1\rangle$ and $|\psi_2\rangle$:

$$\Pi_1 + \Pi_2 + \Pi_0 = I \tag{1.1.2}$$

The clicks from Π_0 are all inconclusive, i.e we gain no information from Π_0. Defining individual failure rates $q_1 = \langle\psi_1|\Pi_0|\psi_1\rangle$ and $q_2 = \langle\psi_2|\Pi_0|\psi_2\rangle$ as the failure probabilities, the task becomes that of minimizing the overall failure rate,

$$Q = \eta_1 q_1 + \eta_2 q_2. \tag{1.1.3}$$

Equivalently optimizing the success rate

$$P_E = \eta_1 p_1 + \eta_2 p_2,$$

such that $P_E + Q = 1.$,

Let us now explicitly determine the POVM operators to be used in the optimization of Q, see Figure (1.1.1). First define the states to be in a two dimensional plane,

$$\begin{aligned} |\psi_1\rangle &= \cos\theta|0\rangle + \sin\theta|1\rangle, \\ |\psi_2\rangle &= \cos\theta|0\rangle - \sin\theta|1\rangle. \end{aligned}$$

The detectors must be orthogonal with the states for which they should not identify, i.e $\Pi_i|\psi_j\rangle = 0$,

$$\begin{aligned} \Pi_1 &= c_1|\psi_2^\perp\rangle\langle\psi_2^\perp|, \\ \Pi_2 &= c_2|\psi_1^\perp\rangle\langle\psi_1^\perp|, \end{aligned}$$

where $|\psi_2^\perp\rangle = \sin\theta|0\rangle + \cos\theta|1\rangle$ and $|\psi_1^\perp\rangle = -\sin\theta|0\rangle + \cos\theta|1\rangle$.

The coefficients $c_i \geq 0$ are yet to be determined based on the optimum strategies.

Using the definition of success probabilities $p_i = \langle\psi_i|\Pi_i|\psi_i\rangle$ the constants c_i can be replaced,

$$\begin{aligned} \Pi_1 &= \frac{p_1}{|\langle\psi_1|\psi_2^\perp\rangle|^2}|\psi_2^\perp\rangle\langle\psi_2^\perp|, \\ \Pi_2 &= \frac{p_2}{|\langle\psi_2|\psi_1^\perp\rangle|^2}|\psi_1^\perp\rangle\langle\psi_1^\perp|. \end{aligned} \tag{1.1.4}$$

To determine the failure operator, insert (1.1.4) into (1.1.2):

$$\Pi_0 = I - \Pi_1 - \Pi_2 = I - \frac{p_1}{|\langle\psi_1|\psi_2^\perp\rangle|^2}|\psi_2^\perp\rangle\langle\psi_2^\perp| - \frac{p_2}{|\langle\psi_2|\psi_1^\perp\rangle|^2}|\psi_1^\perp\rangle\langle\psi_1^\perp|. \tag{1.1.5}$$

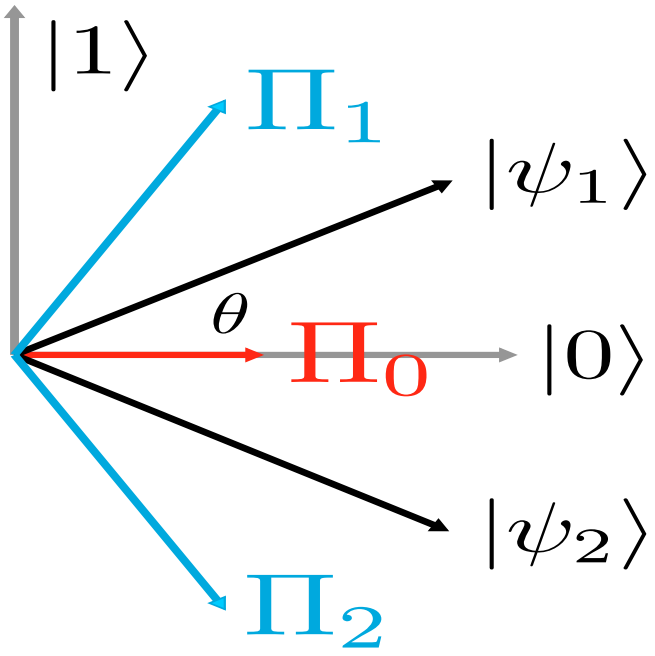


FIGURE 1.1.1. POVM setup that unambiguously discriminates between $|\psi_1\rangle$ and $|\psi_2\rangle$ optimally. The detector $D_1 = \Pi_1$ is setup along $|\psi_2^\perp\rangle$, detector $D_2 = \Pi_2$ is setup along $|\psi_1^\perp\rangle$ and the failure operator is setup symmetrically between $|\psi_1\rangle$ and $|\psi_2\rangle$ for $\eta_1 = \eta_2 = \frac{1}{2}$. When a click in the D_i detector occurs we know for certain that $|\psi_i\rangle$ was prepared $(i = 1, 2)$ as the input state since it is the only one that has a component along this direction. A click in the D_0 detector is considered inconclusive as both states have a component along this direction

direction

After writing everything explicitly, the positivity constraint of the eigenvalues of Π_0 gives the condition

$$q_1 q_2 \geq |\langle\psi_1|\psi_2\rangle|^2, \tag{1.1.6}$$

where we used $q_i = 1 - p_i$.

Using the condition (1.1.6) and taking the equality sign, the total failure rate in (1.1.3) can be expressed in terms of a single constraint. Define the overlap $s \equiv \langle\psi_1|\psi_2\rangle$ and replacing $q_1 = s^2/q_2$ into (1.1.3), $Q = \frac{\eta_1 s^2}{q_2} + \eta_2 q_2$, the optimization follows

$$0 = \frac{\partial Q}{\partial q_2} = -\frac{\eta_1 s^2}{q_2^2} + \eta_2.$$

This leads to individual failure rates $q_1 = \sqrt{\frac{\eta_2}{\eta_1}}s$ and $q_2 = \sqrt{\frac{\eta_1}{\eta_2}}s$. Inserting them back into Equation (1.1.3) gives the optimal Q which it will be defined as Q_0,

$$Q_0 = 2\sqrt{\eta_1\eta_2}s. \tag{1.1.7}$$

Let us now check the conditions where this result holds. The individual error rates must be smaller or equal to one, $q_i \leq 1$. Hence $q_1 = \sqrt{\frac{1-\eta_1}{\eta_1}}s \leq 1$ gives the lower bound on the a-prior probabilities, $\eta_1 \geq \frac{s^2}{1+s^2}$. Similarly the condition that $q_2 \leq 1$ gives the upper bound on the priors $\eta_1 \leq \frac{1}{1+s^2}$. Putting the two conditions together the POVM regime is valid in the range:

$$\frac{s^2}{1+s^2} \leq \eta_1 \leq \frac{1}{1+s^2}. \tag{1.1.8}$$

Outside of this range it is interesting to see that the measurement strategy merges into the projective measurement.

If one of the incoming states is prepared with a much higher probability, say $\eta_1 \gg \eta_2$, we design an experiment where we have only two detection operators. One of them, D_0, the failure operator, simply projects onto state $|\psi_2\rangle$, the detector D_1 projects onto $|\psi_2^{\perp}\rangle$, thus it never clicks for $|\psi_2\rangle$, so that a click on D_1 is associated with the state $|\psi_1\rangle$. A click along D_2 is failure. The setup for the detectors which produce failure rate Q_1 is shown in Figure (1.1.2). The total failure rate is:

$$Q_1 = \eta_1|\langle\psi_1|\psi_2\rangle|^2 + \eta_2. \tag{1.1.9}$$

Similarly for $\eta_2 \gg \eta_1$, the corresponding setup with detectors yielding Q_2 is shown in Figure (1.1.3)

$$Q_2 = \eta_1 + \eta_2|\langle\psi_1|\psi_2\rangle|^2. \tag{1.1.10}$$

Putting the pieces together, the minimum value of Q for the three different regimes can be written as:

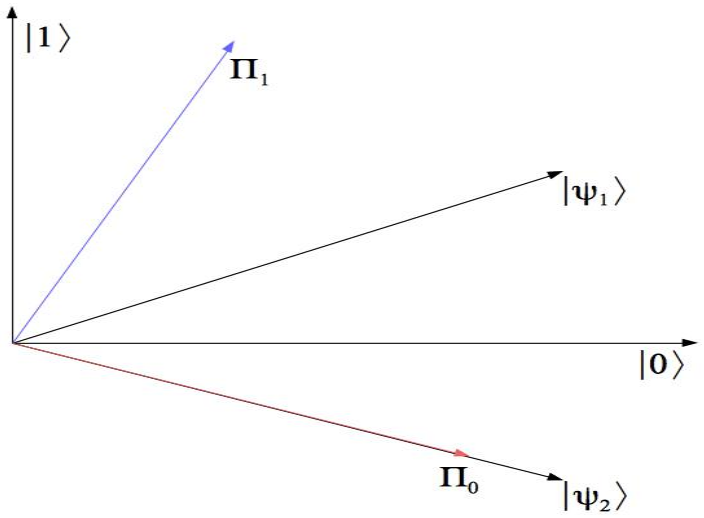


FIGURE 1.1.2. A von Neumann measurement that discriminates $|\psi_1\rangle$ unambiguously. The failure detector $\Pi_0 = P_2$ is set up along the $|\psi_2\rangle$ direction and the second detector $\Pi_1 = P_2^{\perp}$ is set up orthogonal to $|\psi_2\rangle$ therefore never clicks for $|\psi_2\rangle$. When a click in the Π_0 detector occurs we learn nothing as both states have an overlap along P_1.

$$Q = \begin{cases} 2\sqrt{\eta_1\eta_2}s & \text{if } \frac{s^2}{1+s^2} \leq \eta_1 \leq \frac{1}{1+s^2}, \\ \eta_1|\langle\psi_1|\psi_2\rangle|^2 + \eta_2 \text{ if} & \eta_1 > \frac{1}{1+s^2}, \\ \eta_1 + \eta_2|\langle\psi_1|\psi_2\rangle|^2 \text{ if} & \eta_1 < \frac{s^2}{1+s^2}. \end{cases} \tag{1.1.11}$$

It is very interesting that the POVM gives the minimum Q when it is valid. Outside the boundaries it merges with the von Neumann projective measurement.

1.1.2. Unambiguous Discrimination: Two pure states via Neumark's Theorem. Theoretically the problem of minimizing the average failure rate for two pure states has been solved in the previous section. However to be able to implement those schemes we resort to Neumark's theorem which states that any POVM operator can be realized by generalized measurements [?]. The system where the incoming states live is embedded in a

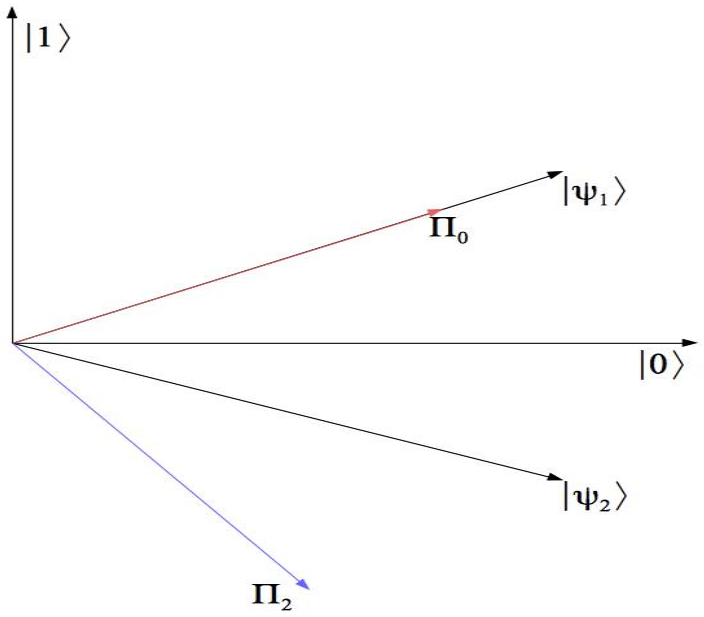


FIGURE 1.1.3. A von Neumann measurement that discriminates $|\psi_2\rangle$ unambiguously. The failure detector $\Pi_0 = P_1$ is set up along the $|\psi_1\rangle$ direction and the second detector $\Pi_2 = P_1^{\perp}$ is set up orthogonal to $|\psi_1\rangle$ therefore never clicks for $|\psi_1\rangle$. When a click in the Π_0 detector occurs, we learn nothing as both states have an overlap along P_2.

larger Hilbert space called ancilla. Then a unitary operator entangles the degrees of freedom of the system with those of the ancilla. After this interaction projective measurements are performed within this larger system in the ancilla. These measurements will also transform the system states in the original Hilbert space because of the entanglement.

To show the power of Neumark's theorem we will re-derive the optimal failure rate of two nonorthogonal states. The incoming states $\{|\psi_1\rangle_s, |\psi_2\rangle_s\}$ which live in the state Hilbert space H_S are embedded with the ancilla $|i\rangle_a$ which live in the ancilla Hilbert space H_A. Now the system and the ancilla live in the larger Hilbert space $H = H_S \otimes H_A$. The incoming states in this larger Hilbert space can be written in the product form $\{|\psi_1\rangle_s|i\rangle_a,\ |\psi_2\rangle_s|i\rangle_a\}$, where $|i\rangle_a$ is the initial state of the ancilla. The unitary operator does the following:

$$\begin{aligned} U|\psi_1\rangle_s|i\rangle_a &= \sqrt{p_1}|\psi_1'\rangle_s|1\rangle_a + \sqrt{q_1}|\phi\rangle_s|0\rangle_a, \\ U|\psi_2\rangle_s|i\rangle_a &= \sqrt{p_2}|\psi_2'\rangle_s|2\rangle_a + \sqrt{q_2}|\phi\rangle_s|0\rangle_a, \end{aligned} \tag{1.1.12}$$

where p_i is the probability of successfully identifying the state $|\psi_i\rangle_s$, q_i is the probability of failing to identify $|\psi_i\rangle_s$, and $p_i + q_i = 1$. The unitary operator aims to take the two incoming states and make them orthogonal. When there is a click on the ancilla $|1\rangle_a$ the input states have been separated and output states $|\psi_i'\rangle_s$ are orthogonal and therefore fully distinguishable. If there is a click along the ancilla $|0\rangle_a$ the incoming states have been collapsed into a single state which carries no information about the system. That is why the choice on the setup of having the failed state $|\phi\rangle_s$ be the same, there should be absolutely no information left in the failed state, otherwise it is not optimal.

Taking the inner product of the two equations in (1.1.12) gives the constraint to the optimization

$$s = \sqrt{q_1 q_2}, \tag{1.1.13}$$

where s was defined to be the overlap of the input states $s \equiv \langle\psi_1|\psi_2\rangle$. In just one line Neumark's setup has produced the constrain and the rest of the derivation, optimizing (1.1.3), is the same as in the POVM section and we do not need to repeat here. Implementation methods have been derived and we will show an example in Chapter 4.

1.2. Minimum Error Discrimination

In the Minimum Error (ME) strategy one is not allowed to abstain from identifying an incoming state, i.e for every incoming state the observer must say which state he was given. Since it was shown that perfect discrimination is not possible the detectors inevitably will make errors. A click in a detector can only identify a state with some probability of success and misidentify the state with some probability of error.

1.2.1. Minimum Error: Two mixed states via POVM. Given an ensemble of two mixed states $\{\rho_1, \rho_2\}$ prepared with different a priori probabilities $\{\eta_1, \eta_2\}$ the task is to minimize the rate for which the detectors misidentify a state. The minimum error problem for two pure or mixed states was first solved by Helstrom [**4**]. We show an alternative derivation to ME of two pure states developed by Herzog [**12**] and Fuchs [**13**]. When the detector Π_i clicks for state ρ_j it is an error, $r_i = Tr(\rho_j \Pi_i)$, a clink for state ρ_i is success $p_i = Tr(\rho_i \Pi_i)$. Thus for two states we want to minimize the following expression.

$$P_E = \eta_1 Tr(\rho_1 \Pi_2) + \eta_2 Tr(\rho_2 \Pi_1). \tag{1.2.1}$$

Using the relation $\eta_1 + \eta_2 = 1$ and $\Pi_1 + \Pi_2 = I$, Equation (1.2.1) can be rewritten as:

$$\begin{aligned} P_E &= \eta_1 Tr(\rho_1 (I - \Pi_1)) + \eta_2 Tr(\rho_2 \Pi_1), \\ &= \eta_1 + Tr[(\eta_2 \rho_2 - \eta_1 \rho_1)\Pi_1, \\ &= \eta_2 - Tr[(\eta_2 \rho_2 - \eta_1 \rho_1)\Pi_2. \end{aligned}$$

Let $\Lambda = \eta_2 \rho_2 - \eta_1 \rho_1$

$$P_E = \eta_1 + Tr(\Lambda \Pi_1) = \eta_2 - Tr(\Lambda \Pi_2). \tag{1.2.2}$$

To minimize P_E, Π_1 should project onto the eigenvectors of the negative eigenvalues of Λ, on the other hand Π_2 should project onto the positive eigenvectors. Let us write Λ into its spectral decomposition.

$$\Lambda = \eta_2 \rho_2 - \eta_1 \rho_1 = \sum_{i=1}^{d} \lambda_i |\lambda_i\rangle\langle\lambda_i|. \tag{1.2.3}$$

To implement the projection of the POVM operators onto the positive (or negative) eigenvectors the eigenvalues λ_i can be split into three categories without any loss of generality: negative, positive and zero:

$$\begin{aligned}\lambda_i &< 0 \text{ for } 1 \leq i < i_o,\\ \lambda_i &> 0 \text{ for } i_o \leq i < d,\\ \lambda_i &= 0 \text{ for } d \leq i < d_s. \end{aligned} \tag{1.2.4}$$

Then from the spectral decomposition we can rewrite (2.2.6) in terms of the optimal POVM.

$$P_E = \eta_1 + \sum_{i=1}^{i_o-1} \lambda_i \langle\lambda_i|\Pi_1|\lambda_i\rangle = \eta_2 - \sum_{i=i_o}^{d_s} \lambda_i \langle\lambda_i|\Pi_2|\lambda_i\rangle, \tag{1.2.5}$$

where $\Pi_1 = \sum_{i=1}^{i_o-1} \lambda_i|\lambda_i\rangle\langle\lambda_i|$ and $\Pi_2 = \sum_{i=i_o}^{d_s} \lambda_i|\lambda_i\rangle\langle\lambda_i|$.

The POVMs need to satisfy the condition $0 \leq \langle\lambda_i|\Pi_j|\lambda_i\rangle \leq 1$ which comes from the definition of the normalized probabilities $r_i = Tr(\rho_i\Pi_j)$. These POVMs are basically von Neumann projectors onto the corresponding eigenvectors. If we now replace the detection operators by the optimal detectors the minimum error can be expressed just in terms of the eigenvalues of Λ.

$$\begin{aligned} P_E &= \eta_1 - \sum_{i=1}^{i_o-1} |\lambda_i| = \eta_2 - \sum_{i=1}^{d_s} |\lambda_i|,\\ &= \frac{1}{2}[1 - \sum_{i=1}^{d_s} |\lambda_i|] = \frac{1}{2}[1 - Tr|\Lambda|],\\ &= \frac{1}{2}[1 - Tr|\eta_2\rho_2 - \eta_1\rho_1|] \end{aligned} \tag{1.2.6}$$

When the states to be discriminated are pure, $\{|\psi_1\rangle, |\psi_2\rangle\}$, the minimum error can be reduced to

$$P_E = \frac{1}{2}[1 - \sqrt{1 - 4\eta_1\eta_2|\langle\psi_1|\psi_2\rangle|^2}]. \tag{1.2.7}$$

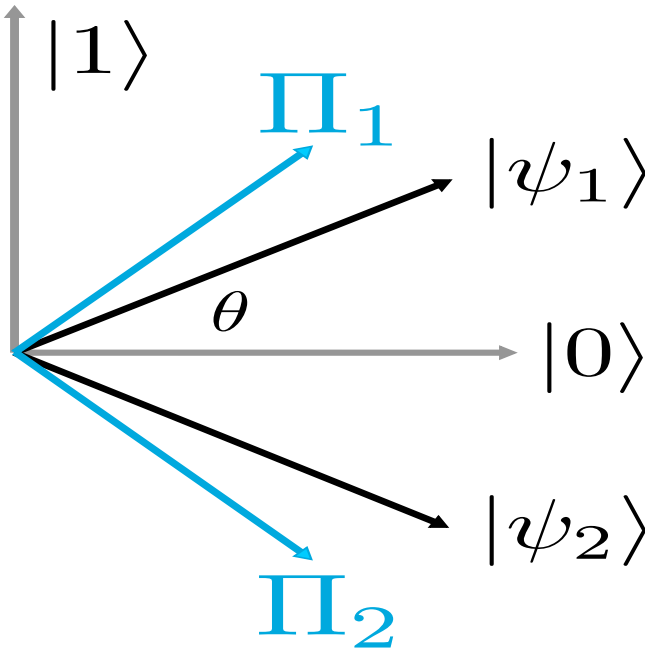


FIGURE 1.2.1. A von Neumann measurement which minimizes the error rate of two pure states prepared with equal priors. The detectors are placed symmetrically along the states $\{|\psi_1\rangle, |\psi_2\rangle\}$ for $\eta_1 = \eta_2 = \frac{1}{2}$.

1.2.2. Minimum Error: Two pure states via Neumark's theorem. Just as we did in the UD case, we will solve the ME problem of discriminating two pure states via the Neumark setup because of it lends itself into an optical implementation. Although in this case the solution is not as straightforward. The incoming states $\{|\psi_1\rangle_s, |\psi_2\rangle_s\}$ which live in the state Hilbert space H_S are embedded with the ancilla $|i\rangle_a$ which live in the ancilla Hilbert space H_A. Now the system and the ancilla live in the larger Hilbert space $H = H_S \otimes H_A$. The incoming states in this larger Hilbert space can be written in the product form $\{|\psi_1\rangle_s|i\rangle_a,\ |\psi_2\rangle_s|i\rangle_a\}$. The unitary operator does the following:

$$U|\psi_1\rangle_s|i\rangle_a = \sqrt{p_1}|\psi_1\rangle|1\rangle + \sqrt{r_1}|\psi_2\rangle|2\rangle, \tag{1.2.8}$$

$$U|\psi_2\rangle_s|i\rangle_a = \sqrt{p_2}|\psi_2\rangle|2\rangle + \sqrt{r_2}|\psi_1\rangle|1\rangle, \tag{1.2.9}$$

where p_i is the probability of having successfully identified the state $|\psi_i\rangle$ and r_i is the probability of misidentifying the state $|\psi_i\rangle$ for $|\psi_j\rangle$.

Taking the inner product of the two equations in (1.2.9) gives the constraint:

$$s = \sqrt{p_1 r_2} + \sqrt{p_2 r_1}. \tag{1.2.10}$$

The quantity we are looking to minimize in the average error rate:

$$P_E = \eta_1 r_1 + \eta_2 r_2, \tag{1.2.11}$$

subject to the constraint in (1.2.10). Adding the constraint to 1.2.11 with one Lagrange multiplier and using $p_i + r_i = 1$

$$F_E = \eta_1 r_1 + \eta_2 r_2 + \lambda \left[s - \sqrt{(1-r_1) r_2} - \sqrt{(1-r_2) r_1} \right]. \tag{1.2.12}$$

Differentiating with respect to r_i, then setting the resulting equations equal to zero yields

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

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

Rearranging the above two equations so that the left hand side is only a function of r_i and the right hand sides turn out to be equivalent,

$$\frac{2\eta_1}{\lambda} \sqrt{r_1(1-r_1)} = \sqrt{r_1 r_2} - \sqrt{(1-r_1)(1-r_2)}, \tag{1.2.13}$$

$$\frac{2\eta_2}{\lambda} \sqrt{r_2(1-r_2)} = \sqrt{r_1 r_2} - \sqrt{(1-r_1)(1-r_2)}. \tag{1.2.14}$$

The right hand sides of Eq.(1.2.13) and (1.2.14) 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 1.2.10,

$$r_i = \frac{1}{2}\left(1 \pm \sqrt{1 - \frac{\lambda^2 C^2}{\eta_i^2}}\right) = \frac{1}{2}\left(1 - \sqrt{1 - \frac{\delta^2}{\eta_i^2}}\right), \tag{1.2.15}$$

$$r_i = \frac{1}{2}\left[1 - A_i\right], \tag{1.2.16}$$

where $A_i \equiv \sqrt{1 - \frac{\delta^2}{\eta_i^2}}$ and $\delta^2 \equiv \lambda^2 C^2$. The smaller r_i is picked (lower sign in 1.2.15) as this represents error rate, which is to be minimized. Now replace r_i into the constraint (1.2.10) and solve for δ:

$$\begin{aligned}
s &= \sqrt{(1-r_1)r_2} + \sqrt{(1-r_2)r_1}, \\
2s &= \sqrt{(1+A_1)(1-A_2)} + \sqrt{(1-A_1)(1+A_2)}, \\
2s^2 &= 1 - A_1A_2 + \sqrt{(1-A_1^2)(1-A_2^2)}, \\
2s^2 &= 1 - A_1A_2 + \frac{\delta^2}{\eta_1\eta_2}, \\
(2s^2 - 1 - \tfrac{\delta^2}{\eta_1\eta_2})^2 &= 1 - \frac{\delta^2}{\eta_1^2} - \frac{\delta^2}{\eta_2^2} + \frac{\delta^4}{\eta_1^2\eta_2^2}.
\end{aligned}$$

After some tedious but trivial algebra:

$$\delta^2 = \frac{4s^2(1-s^2)\eta_1^2\eta_2^2}{1 - 4\eta_1\eta_2 s^2}. \tag{1.2.17}$$

Now substitute the value of δ from (1.2.17) into (1.2.15) to get the explicit form of the individual error rates,

$$r_i = \frac{1}{2}\left[1 - \frac{1 - 2\eta_i s^2}{\sqrt{1 - 4\eta_1\eta_2 s^2}}\right] \tag{1.2.18}$$

Inserting r_1 and r_2 into (1.2.1) Helstrom bound is retrieved **[4]**

$$
\begin{aligned}
P_E &= \frac{1}{2}\left[1 - \frac{\eta_1 - 2\eta_1\eta_2 s^2}{\sqrt{1-4\eta_1\eta_2 s^2}} - \frac{\eta_2 - 2\eta_1\eta_2 s^2}{\sqrt{1-4\eta_1\eta_2 s^2}}\right], \\
P_E &= \frac{1}{2}\left[1 - \sqrt{1-4\eta_1\eta_2 s^2}\right]. \qquad (1.2.19)
\end{aligned}
$$

As we mentioned above the advantage of solving the ME problem via Neumark is that we now have explicit expressions for the individual error rates, r_1 and r_2. In the Implementation chapter, it is shown that the unitary operator which carries out this operation can be written in terms of r_i and then be decomposed into beam splitters and phase shifters using optical interferometers.

CHAPTER 2

Optimal discrimination of a certain class of mixed states with a fixed rate of inconclusive outcome (FRIO)

In this chapter we will derive the optimal strategy with a Fixed Rate of Inconclusive Outcomes ($FRIO$) that optimally interpolates between the two well known limits, Helstrom bound for minimum error and IDP for unambiguous discrimination. In particular, as the main finding of our paper, we will show that the optimal distribution of the fixed rate of inconclusive outcomes, Q, among the 2-dimensional subspaces spanned by the pair of Jordan basis vectors is highly non-trivial and an interesting threshold-like structure emerges: As we start increasing Q from $Q = 0$, first only one subspace receives the entire inconclusive rate. Then, as we increase Q further, at a certain threshold a second subspace starts sharing the inconclusive rate. If we increase Q further, at another threshold a third subspace also starts sharing Q , and so on, until above a last threshold all subspaces share the available inconclusive rate.

2.1. Review of the FRIO solution for two pure states

We first present a brief review of the method developed in [**14**] for the two pure state optimal FRIO problem since the rest of the paper relies heavily on this method. We derive the maximum probability of success or, equivalently, the minimum probability of error in identifying the states, when a certain fixed rate of inconclusive outcomes is allowed. By varying the inconclusive rate, the scheme optimally interpolates between Unambiguous and Minimum Error discrimination (UD and ME).

In all of these scenarios (UD, ME or FRIO) one is given a system which is promised to be prepared in one of two known pure states, $|\psi_1\rangle$ or $|\psi_2\rangle$, but we don't know which. The pure states are prepared with prior probabilities η_1 and η_2, respectively, such that $\eta_1 + \eta_2 = 1$.

It is well known that two pure states can be discriminated both unambiguously and with minimum error. In Section (2.1) we showed the solution of the optimal average inconclusive rate, $Q_{\rm c}$ for UD:

$$Q_{\rm c} = \begin{cases} \eta_1 + \eta_2 \cos^2\theta, \text{if } \eta_1 < \dfrac{\cos^2\theta}{1+\cos^2\theta} \equiv \eta_1^{(l)}, \\ \eta_2 + \eta_1 \cos^2\theta, \text{if } \eta_1 > \dfrac{1}{1+\cos^2\theta} \equiv \eta_1^{(r)}, \\ 2\sqrt{\eta_1\eta_2}\cos\theta \equiv Q_0, \text{if } \eta_1^{(l)} \leq \eta_1 \leq \eta_1^{(r)} \ , \end{cases} \tag{2.1.1}$$

where $|\langle\psi_1|\psi_2\rangle| \equiv \cos\theta$ is the overlap of the input states.

In Section (2.2) we also showed the optimal average error rate for ME:

$$P_{\rm E}^{ME} = \frac{1}{2}\left(1 - \sqrt{1 - 4\eta_1\eta_2\cos^2\theta}\,\right). \tag{2.1.2}$$

It has long been suggested **[15]** that the above state discrimination points are part of a more general scheme. One which interpolates between optimal ME and UD. The FRIO strategy achieves that goal by minimizing the error rate while allowing for some rate of inconclusive results. Hence the strategy has three measurement outcomes, one that identifies with the first state, one that identifies with the second state and one that does not identify with a state at all, corresponding to the inconclusive outcome. The authors in **[14]** solve the problem via the POVM method. Three POVM elements are needed such that:

$$\Pi_1 + \Pi_2 + \Pi_0 = I \ , \tag{2.1.3}$$

A click in Π_1 is identified with state $|\psi_1\rangle$, a click in Π_2 is identified with the second state $|\psi_2\rangle$ and any clicks in the operator Π_0 corresponds to inconclusive outcomes. However the operators Π_i can also click for $|\psi_j\rangle$. We wish to minimize the average error rate,

$$\begin{aligned} P_E &= \eta_1 Tr\left[\langle\psi_1|\Pi_2|\psi_1\rangle\right] + \eta_2 Tr\left[\langle\psi_2|\Pi_1|\psi_2\rangle\right], \\ &= \eta_1 tr\left[\Pi_2|\psi_1\rangle\langle\psi_1|\right] + \eta_2 tr\left[\Pi_1|\psi_2\rangle\langle\psi_2|\right], \\ &= \eta_1 tr\left[\Pi_2\rho_1\right] + \eta_2 tr\left[\Pi_1\rho_2\right], \end{aligned}$$

where $\rho_1 = |\psi_1\rangle\langle\psi_1|$ and $\rho_2 = |\psi_2\rangle\langle\psi_2|$ are the corresponding pure state density matrices (Equivalently we wish to maximize the average success rate $P_s = \eta_1 tr[\Pi_1\rho_1] + \eta_2 tr[\Pi_2\rho_2]$), for a fixed rate of inconclusive results,

$$Q = \eta_1 tr[\Pi_0\rho_1] + \eta_2 tr[\Pi_0\rho_2] \ . \tag{2.1.4}$$

The solution to the problem involves a neat trick which transforms the three element POVM defined in Eq.(2.1.3) into a two element POVM, namely,

$$\begin{aligned} \Pi_1 + \Pi_2 &= I - \Pi_0, \\ \Omega^{-1/2}(\Pi_1 + \Pi_2)\Omega^{-1/2} &= \Omega^{-1/2}(I - \Pi_0)\Omega^{-1/2}, \\ \tilde{\Pi}_1 + \tilde{\Pi}_2 &= I, \end{aligned} \tag{2.1.5}$$

where $\Omega \equiv I - \Pi_0$ and $\tilde{\Pi}_i \equiv \Omega^{-1/2}\Pi_i\,\Omega^{-1/2}$. Should be noted that $\Omega^{-1/2} = (I - \Pi_0)^{-1/2}$ exists unless Π_0 has a unit value, in which case the problem is treated separately. It was important to notice that for optimal FRIO, Π_0 must be rank one operator. That means that Π_0 maps both incoming non-orthogonal states onto a single state which is then discarded. On the other hand the other two POVM elements map the input states onto two different states.

The FRIO problem has essentially been transformed into a new parametrized optimization scheme with two POVM elements, $\tilde{\Pi}_1$ and $\tilde{\Pi}_2$. The error probability in the new parametrized form, $P_E = (1 - Q)\tilde{P}_E$, becomes

$$\tilde{P}_e = Tr(\tilde{\eta}_1\tilde{\rho}_1\tilde{\Pi}_2) + Tr(\tilde{\eta}_2\tilde{\rho}_2\tilde{\Pi}_1), \tag{2.1.6}$$

where the normalized states $\tilde{\rho}_i$ and normalized a priori probabilities $\tilde{\eta}_i$ are

$$\tilde{\rho}_i = \frac{\Omega^{1/2}\rho_i\Omega^{1/2}}{Tr(\Omega\rho_i))}, \qquad \tilde{\eta}_i = \frac{\eta_i Tr(\Omega\rho_i)}{1 - Q}, \tag{2.1.7}$$

Equation (2.1.6) and $\tilde{\Pi}_1 + \tilde{\Pi}_2 = I$ define a ME discrimination problem for the transformed states and *priors* given in Eq. (2.1.7). The optimal solution to this ME discrimination problem immediately follows by using the tilde quantities in (2.1.2), with $|\tilde{\psi}_i\rangle = \Omega^{1/2}|\psi_i\rangle/\sqrt{\langle\psi_i|\Omega|\psi_i\rangle}$ being the properly normalized transformed states.

We can immediately write down the solution in parametrized form

$$\tilde{P}_{\rm E}^{ME} = \frac{1}{2}\left(1 - \sqrt{1 - 4\tilde{\eta}_1\tilde{\eta}_2\left|\left\langle\tilde{\psi}_1 \mid \tilde{\psi}_2\right\rangle\right|^2}\right)$$

Writing the input states as, $|\psi_i\rangle = c_i|0\rangle + s_i|1\rangle$, where $c_i \equiv \cos\theta_i$, $s_i \equiv \sin\theta_i$, we obtain the transformed states and priors from Eq. (2.1.7). Since the optimal Π_0 is a positive rank one operator it can be written as $\Pi_0 = \xi|0\rangle\langle 0|$, where ξ is its eigenvalue, $0 \le \xi \le 1$, and the eigenstate belonging to ξ is $|0\rangle$ and the orthogonal state is $|1\rangle$. In this basis $\Omega = (1-\xi)|0\rangle\langle 0| + |1\rangle\langle 1|$. The tilde error rate now becomes:

$$\tilde{P}_e^{\rm ME} = \frac{1}{2}\left\{1 - \sqrt{1 - 4\eta_1\eta_2(\cos\theta - \xi c_1 c_2)^2/(1-Q)^2}\right\}, \tag{2.1.8}$$

where $\theta_1 - \theta_2 \equiv \theta$. It follows from (2.1.4) that

$$\xi = \frac{Q}{\eta_1 c_1^2 + \eta_2 c_2^2}. \tag{2.1.9}$$

Hence Eq. (2.1.8) depends only on one parameter, say θ_1, which determines the orientation of Π_0 relative to that of the two pure states. The minimization over θ_1 simplifies considerably, using Eq. (2.1.9) and defining $c_1\,\eta_1^{1/2}(\eta_1 c_1^2 + \eta_2 c_2^2)^{-1/2} \equiv \cos\varphi$, and $c_2\,\eta_2^{1/2}(\eta_1 c_1^2 + \eta_2 c_2^2)^{-1/2} \equiv \sin\varphi$. The resulting expression is minimum for $\varphi = \pi/4$, yielding

$$P_e^{\rm min} = \frac{1}{2}\left\{1 - Q - \sqrt{(1-Q)^2 - (Q_0 - Q)^2}\right\}, \tag{2.1.10}$$

for all $Q \le Q_0 \equiv 2\sqrt{\eta_1\eta_2}s$. This is the optimal error rate for an intermediate range of the prior probabilities, a rather nice looking formula for a somewhat complicated problem. We can check that is reproduces the Helstrom bound for zero failure rate $Q = 0$,

$P_E = \frac{1}{2}\left[1-\sqrt{1-Q_0^2}\right] = \frac{1}{2}\left[1-\sqrt{1-2\eta_2\eta_2 s^2}\right]$ and the IDP bound for zero error rate $0 = \frac{1}{2}\left\{1-Q-\sqrt{(1-Q)^2-(Q_0-Q)^2}\right\} \Rightarrow Q = Q_0 \equiv 2\sqrt{\eta_1\eta_2}s$.

For the validity of (2.1.10), $\xi \leq 1$ must hold. The definitions of $\cos\varphi$ and $\sin\varphi$ after Eq. (2.1.9) give $\sqrt{\eta}_2 c_2 = \sqrt{\eta}_1 c_1$ for $\varphi = \pi/4$ which, in turn, leads to $\eta_1 c_1^2 = \eta_2 c_2^2 = \eta_1\eta_2 \sin^2\theta/(1-Q_0)$, and Eq. (2.1.9) yields $\xi = (1-Q_0)Q/(2\eta_1\eta_2\sin^2\theta)$. Setting $\xi = 1$ defines the boundary Q_b, between the projective and POVM regimes,

$$Q_{\rm b} \equiv 2\eta_1\eta_2\sin^2\theta/(1-Q_0). \tag{2.1.11}$$

Hence $\xi \leq 1$ if $Q \leq Q_{\rm b}$ and $\xi = 1$ if $Q > Q_{\rm b}$.

In Fig. 2.1.1 we plot $Q_{\rm c}$ and $Q_{\rm b}$ vs. η_1 together for a fixed overlap, $\cos\theta = 0.5$ ($\theta = \pi/3$). The two curves intersect at $\eta_1 = \eta_1^{(l)}$ and $\eta_1 = \eta_1^{(r)}$, the same points as in Eq. (2.1.1). The interval $0 \leq \eta_1 \leq 1$ is thus divided into three regions. In regions I and III, we have $Q_{\rm b} < Q_{\rm c}$ and the solution (2.1.10) is valid for $0 \leq Q < Q_{\rm b}$ only. In Region II, $\eta_1^{(l)} \leq \eta_1 \leq \eta_1^{(r)}$, we have $Q_{\rm c} = Q_0 < Q_{\rm b}$ and the solution (2.1.10) is valid for the entire $0 \leq Q \leq Q_{\rm c}$ range.

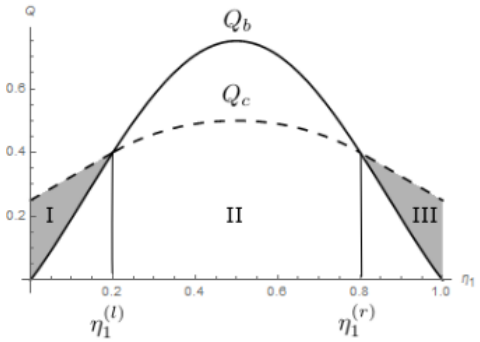


FIGURE 2.1.1. $Q_{\rm c}$ (dashed line, Eq. (2.1.1)) and $Q_{\rm b}$ (solid line, Eq. (2.1.11)) vs. η_1 for $\theta = \pi/3$. Measurements can be optimized in the area under the dashed line, Q_c. Measurements in the area above Q_c are suboptimal. In the shaded areas between Q_c and Q_b (regions I, left, and III, right) the optimal FRIO measurement is a projective measurement, in the unshaded area below Q_c (region II) the optimal measurement is a POVM.

In the shaded parts of regions I and III one has $Q_{\rm b} \leq Q \leq Q_{\rm c}$ and, necessarily, $\xi = 1$. Hence, $\Pi_0 = |0\rangle\langle 0|$ and $\Omega = |1\rangle\langle 1|$ are projectors. Therefore, $\Omega^{-1/2}$ does not exist in these areas and the case needs special consideration.

The calculation of the error probability is most easily performed by realizing that Π_1 and Π_2 become degenerate, both must be proportional to $\Pi_d = |1\rangle\langle 1|$. The three-element POVM becomes a standard two element projective measurement, $\{\Pi_d = |1\rangle\langle 1|,\ \Pi_0 = |0\rangle\langle 0|\}$. We identify a click in Π_d with ρ_1 (ρ_2) if $\eta_1 \geq \eta_2$ ($\eta_2 \geq \eta_1$), so $P_{e(s)} = \eta_2 s_2^2, \quad P_{s(e)} = \eta_1 s_1^2$, with $Q = 1 - P_e - P_s$. These equations completely determine the solution. There is nothing to optimize here, so we drop the superscript min in what follows. $\theta_1 - \theta_2 = \theta$ immediately gives $Q(P_e)$ as

$$Q = 1 - P_e - \eta_{1(2)} \left(\sqrt{\frac{P_e}{\eta_{2(1)}}} \cos\theta \pm \sqrt{1 - \frac{P_e}{\eta_{2(1)}}} \sin\theta \right)^2 . \tag{2.1.12}$$

Inverting this equation gives $P_e(Q)$ yields

$$P_e = \eta_2 \frac{2\eta_1 \cos^2\theta (1 - Q - Q_2) - (\eta_1 - \eta_2)(Q - Q_2) - 2\eta_1\eta_2 \sin\theta \cos\theta \sqrt{Q(1-Q) - \eta_1\eta_2 \sin^2\theta}}{1 - 4\eta_1\eta_2 \sin^2\theta} , \tag{2.1.13}$$

but the resulting expression is not particularly insightful. However, we note that for $P_e = 0$ (UD limit) one has $Q = Q_2$, given by the second line in Eq. (2.1.1), and for $Q = Q_{\rm b}$, P_e reduces to (2.1.10), as it should.

2.2. FRIO discrimination of two Rank 2 mixed states

In our work we extend the FRIO discrimination scheme to a particular case of mixed states. We solve the FRIO problem for two mixed states which exhibit a Jordan structure. The two states, with their respective *prior* probabilities η_1 and η_2, can be written in the spectral decomposition form as,

$$\begin{aligned} \rho_1 &= \sum_{i=1}^{N} r_i |r_i\rangle\langle r_i|, \\ \rho_2 &= \sum_{i=1}^{N} s_i |s_i\rangle\langle s_i| \ . \end{aligned} \tag{2.2.1}$$

In the Jordan structure these states satisfy the following conditions:

$$\begin{aligned} \langle r_i|r_j\rangle &= \delta_{ij} \\ \langle s_i|s_j\rangle &= \delta_{ij} \\ \langle r_i|s_j\rangle &= \delta_{ij}\cos\theta_i \end{aligned} \tag{2.2.2}$$

Due to the Jordan structure, instead of one $2N$ dimensional problem we have N mutually orthogonal 2-dimensional subspaces. The i^{th} subspace is spanned by $|r_i\rangle$ and $|s_i\rangle$ with *prior* probabilities $\eta_1 r_i$ and $\eta_2 s_i$, respectively. Thus the discrimination of the two mixed states ρ_1 and ρ_2 can be reduced into that of pure state discrimination in each subspace, by discriminating $|r_i\rangle$ and $|s_i\rangle$ in each subspace.

In this section we solve the case where $N = 2$, there is 2-dimensional subspaces for each density operator. The density operator in Eq. (2.2.1) can be expressed as

$$\begin{aligned} \rho_1 &= r_1 |r_1\rangle \langle r_1| + r_2 |r_2\rangle \langle r_2| , \\ \rho_2 &= s_1 |s_1\rangle \langle s_1| + s_2 |s_2\rangle \langle s_2| . \end{aligned} \tag{2.2.3}$$

The overall error rate is

$$\begin{aligned} P_E &= \eta_1 tr\left(\Pi_2\rho_1\right) + \eta_2 tr\left(\Pi_1\rho_2\right), \\ &= \eta_1\left(r_1 tr\left(\Pi_{s,1}|r_1\rangle\langle r_1|\right) + r_2 tr\left(\Pi_{s,2}|r_2\rangle\langle r_2|\right)\right) + \qquad (2.2.4) \\ &\quad \eta_2\left(s_1 tr\left(\Pi_{r,1}|s_1\rangle\langle s_1|\right) + s_2 tr\left(\Pi_{r,2}|s_2\rangle\langle s_2|\right)\right) \qquad (2.2.5) \end{aligned}$$

In this case the first subspace is spanned by $\{|r_1\rangle, |s_2\rangle\}$ and the second subspace by $\{|r_2\rangle, |s_2\rangle\}$. The task becomes that of performing FRIO discrimination of $\{|r_i\rangle$ and $|s_i\rangle\}$ within subspace i. The error rate in each subspace is:

$$\begin{aligned} P_{e,1} &= \eta_1 r_1 tr\left(\Pi_{s,1}\left|r_1\right\rangle\left\langle r_1\right|\right) + \eta_2 s_1 tr\left(\Pi_{r,1}\left|s_1\right\rangle\left\langle s_1\right|\right) \\ P_{e,1} &= \eta_1 r_2 tr\left(\Pi_{s,2}\left|r_2\right\rangle\left\langle r_2\right|\right) + \eta_2 s_2 tr\left(\Pi_{r,2}\left|s_2\right\rangle\left\langle s_2\right|\right) \end{aligned} \tag{2.2.6}$$

The prior probability of $|r_i\rangle$ is $\eta_1 r_i$ while the prior probability of $|s_i\rangle$ is $\eta_2 s_i$. We define their normalized probabilities as

$$\begin{aligned} \eta_{1,i} &\equiv \frac{\eta_1 r_i}{\eta_1 r_i + \eta_2 s_i}, \\ \eta_{2,i} &\equiv \frac{\eta_2 s_i}{\eta_1 r_i + \eta_2 s_i}\ , \end{aligned} \tag{2.2.7}$$

such that $\eta_{1,i} + \eta_{2,i} = 1$. The error rate in each subspace becomes

$$\begin{aligned} \tilde{P}_{e,1} &= \eta_{1,1} tr\left(\Pi_{s,1}\left|r_1\right\rangle\left\langle r_1\right|\right) + \eta_{2,1} tr\left(\Pi_{r,1}\left|s_1\right\rangle\left\langle s_1\right|\right) \\ \tilde{P}_{e,2} &= \eta_{1,2} tr\left(\Pi_{s,2}\left|r_2\right\rangle\left\langle r_2\right|\right) + \eta_{2,2} tr\left(\Pi_{r,2}\left|s_2\right\rangle\left\langle s_2\right|\right) \end{aligned} \tag{2.2.8}$$

where $\tilde{P}_{e,i} = \frac{P_{e,i}}{(\eta_1 r_i + \eta_2 s_i)}$.

We note that this reduces the problem to the FRIO discrimination of two pure states in subspace i, with prior probabilities given above. It follows immediately that the solution is given by (2.1.10) in the POVM regime of Q_i, $Q_i \leq Q_{c,i}, Q_{th,i}$, and (2.1.12) in the projective regime of Q_i, $Q_{th,i} < Q_i \leq Q_{c,i}$, again with the above substitutions.

In the following we will focus mainly on the POVM regime where the solution in each subspace is given explicitly by Eq. (2.1.10),

$$P_{e,1} = \frac{1}{2}(1 - Q_1 - \sqrt{(1-Q_1)^2 - (Q_{0,1} - Q_1)^2})\ , \tag{2.2.9}$$

$$P_{e,2} = \frac{1}{2}(1 - Q_2 - \sqrt{(1-Q_2)^2 - (Q_{0,2} - Q_2)^2}) \tag{2.2.10}$$

Where we introduced a fixed rate of inconclusive outcomes for each subspace i, Q_i, such that $0 \leq Q_i \leq Q_{c,i}$ where $Q_{c,i}$ is given by Eq. (2.1.1) with the obvious substitutions $\eta_1 \to \eta_{1,i}$,

$\eta_2 \to \eta_{2,i}$ and $\cos\theta \to \cos\theta_i = \langle r_i|s_i\rangle \Rightarrow Q_{0,i} = 2\sqrt{\eta_{1,i}\eta_{2,i}}\langle r_i|s_i\rangle$. Q_0 was introduced in Eq. (2.1.1) for the single subspace (two pure states) case, $Q_{0,i}$ is its generalization for the case of two (or many) subspaces. We introduce the weight ω_i of subspace i as

$$\eta_1 r_i + \eta_2 s_i = \omega_i \ , \tag{2.2.11}$$

where, obviously,

$$\omega_1 + \omega_2 = 1 \ . \tag{2.2.12}$$

The total error rate in Eq. (2.2.4) can be expressed as weighted sum of the error rates of the individual subspaces

$$\begin{aligned} P_e &= \omega_1 P_{e,1} + \omega_2 P_{e,2} \ , = \frac{\omega_1}{2}\left[(1-Q_1) - \sqrt{(1-Q_1)^2 - (Q_{0,1}-Q_1)^2}\right] + \\ &\quad \frac{\omega_2}{2}\left[(1-Q_2) - \sqrt{(1-Q_2)^2 - (Q_{0,2}-Q_2)^2}\right] \end{aligned} \tag{2.2.13}$$

The task it to determine the optimal distribution of Q among the two subspaces, the distribution that minimizes the error rate for a fixed amount of inconclusive results. We can write the total inconclusive rate as a weighted of the inconclusive rates of the individual subspaces,

$$Q = \omega_1 Q_1 + \omega_2 Q_2 \ . \tag{2.2.14}$$

Since the total failure rate Q is fixed, then only one of the $Q_i s$ is an independent variable. Thus the total error rate can be expressed in terms of only one variable and be optimized in that variable. Inserting $Q_2 = (Q - \omega_1 Q_1)/\omega_2$ into Eq. (2.2.13) it becomes a function of the independent variable Q_1 and the optimization with respect to this variable is straightforward:

$$
\begin{aligned}
P_e &= \frac{\omega_1}{2}\left[(1-Q_1)-\sqrt{(1-Q_1)^2-(Q_{0,1}-Q_1)^2}\right]+ \\
&\quad \frac{\omega_2}{2}\left[\left(1-\frac{Q}{\omega_2}+\frac{\omega_1 Q_1}{\omega_2}\right)-\sqrt{\left(1-\frac{Q}{\omega_2}+\frac{\omega_1 Q_1}{\omega_2}\right)^2-\left(Q_{0,2}-\frac{Q}{\omega_2}+\frac{\omega_1 Q_1}{\omega_2}\right)^2}\right].
\end{aligned}
$$

Simply setting the derivative $\frac{\partial P_e}{\partial Q_1}=0$ and solving for Q_1,

$$
\begin{aligned}
0=\frac{\partial P_e}{\partial Q_1} &= \frac{1}{2}\left[-\omega_1-\frac{\omega_1(1-Q_{0,1})}{\sqrt{(1-Q_1)^2-(Q_{0,1}-Q_1)^2}}\right]+ \\
&\quad \frac{1}{2}\left[\omega_1-\frac{\omega_1(1-Q_{0,2})}{\sqrt{\left(1-\frac{Q}{\omega_2}+\frac{\omega_1 Q_1}{\omega_2}\right)^2-\left(Q_{0,2}-\frac{Q}{\omega_2}+\frac{\omega_1 Q_1}{\omega_2}\right)^2}}\right]
\end{aligned}
$$

$$
(1-Q_{0,2})^2\left[(1-Q_1)^2-(Q_{0,1}-Q_1)^2\right] = (1-Q_{0,2})^2\left[\left(1-\frac{Q}{\omega_2}+\frac{\omega_1 Q_1}{\omega_2}\right)^2-\left(Q_{0,2}-\frac{Q}{\omega_2}+\frac{\omega_1 Q_1}{\omega_2}\right)^2\right]
$$

In order to express the optimal failure rates in compact form it will be useful to introduce at this point the following convention. Without loss of generality in what follows we assume the hierarchy

$$
Q_{0,1}\geq Q_{0,2}\ . \tag{2.2.15}
$$

We also introduce the notation

$$
Q_{th}^{(1)}\equiv 0 \qquad Q_{th}^{(2)}\equiv\frac{\omega_1(Q_{0,1}-Q_{0,2})}{1-Q_{0,2}}\ . \tag{2.2.16}
$$

Then result of the optimization can be written as

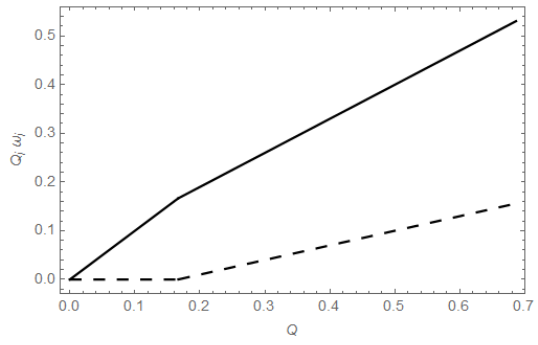
$$
Q_1^{opt}=\begin{cases}\frac{Q}{\omega_1} & \text{if } Q_{th}^{(1)}\leq Q\leq Q_{th}^{(2)},\\ \frac{1-Q_{0,1}}{1-Q_0}(Q-Q_{th}^{(2)})+\frac{Q_{th}^{(2)}}{\omega_1} & \text{if } Q_{th}^{(2)}<Q\leq Q_0,\end{cases} \tag{2.2.17}
$$

and

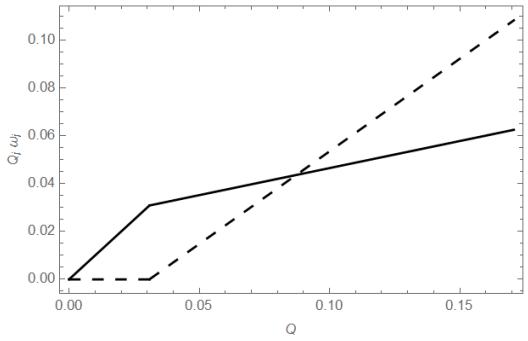
$$Q_2^{opt} = \begin{cases} 0 & \text{if } 0 \leq Q \leq Q_{th}^{(2)}, \\ \frac{1-Q_{0,2}}{1-Q_0}(Q - Q_{th}^{(2)}) & \text{if } Q_{th}^{(2)} < Q \leq Q_0. \end{cases} \tag{2.2.18}$$

The threshold structure is the interesting feature of this problem. In the region $0 \leq Q \leq Q_{th}^{(2)}$ only one of the subspaces accommodates the fixed failure rate Q, it is the subspace with the larger $Q_{0,i}$. The other subspace allows for no inconclusive rate at all, $Q = 0$, and operates in the Minimum Error regime. Above this threshold $Q_{th}^{(2)} < Q \leq Q_0$, both subspaces share the failure rate but at different values. If $\omega_1 Q_{0,1} > \omega_2 Q_{0,2}$, the first subspace always accommodates more inconclusive rate than the second, otherwise above some value of the total Q the second subspace will accommodate more inconclusive rate than the first.

The situation is depicted in Fig. 2.2.1, for some specific values of the parameters.



(a)



(b)

FIGURE 2.2.1. $\omega_1 Q_1^{\text{opt}}$ vs. Q (solid line) and $\omega_2 Q_2^{\text{opt}}$ vs. Q (dashed line). For the figure we used $\eta_1 = \eta_2 = 1/2$ and the following parameter values: (a) $\theta_1 = \pi/4$, $\theta_2 = 2\pi/7$, $r_1 = 3/4$, $r_2 = 1/4$, $s_1 = 3/4$ and $s_2 = 1/4$; (b) $\cos\theta_1 = 1/4\sqrt{3}$, $\cos\theta_2 = 1/4$, $r_1 = 3/4$, $r_2 = 1/4$, $s_1 = 3/4$ and $s_2 = 1/4$.

To get the expression of the total error rate as a function of FRIO, insert (2.2.17) and (2.2.18) into (2.2.13),

$$P_E = \begin{cases} \frac{1}{2}\left\{(1-Q) - \sqrt{(\omega_1 - Q)^2 - (\omega_1 Q_{0,1} - Q)^2} - \sqrt{(\omega_2)^2 - (\omega_2 Q_{0,2})^2}\right\} & \text{if } Q_{th}^{(1)} \leq Q \leq Q_{th}^{(2)} \ , \\ \frac{1}{2}\left\{(1-Q) - \sqrt{(1-Q)^2 - (Q_0 - Q)^2}\right\} & \text{if } Q_{th}^{(2)} < Q \leq Q_0 \ , \end{cases}$$

We will show that this is valid in all regions of the parameter Q. For $Q = 0$ it reduces to the optimal Minimum Error expression for two subspaces **[16]**, as it should. On the other hand for $P_e = 0$ it reduces to the optimal Unambiguous Discrimination of two subspaces **[17, 18]** We close this section by displaying $P_{e,i}$ vs. Q in Fig. 2.2.2, for some specific values of the parameters.

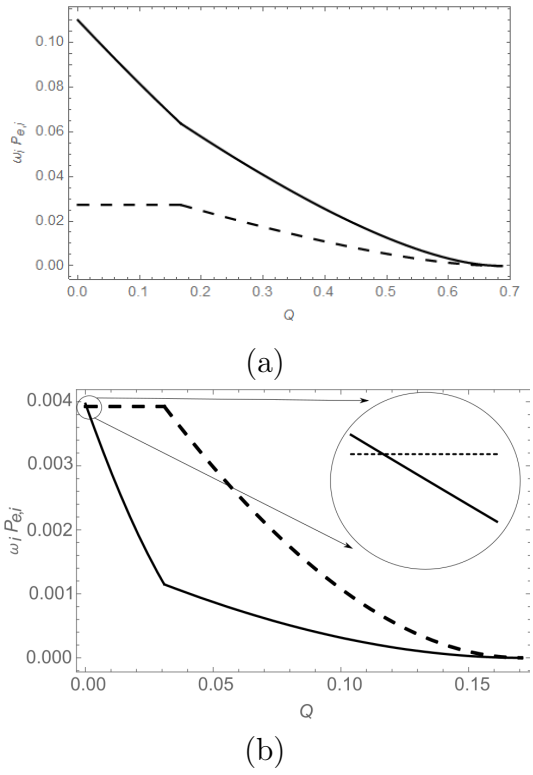


FIGURE 2.2.2. $\omega_1 P_{e,1}^{\text{opt}}$ vs. Q (solid line) and $\omega_2 P_{e,2}^{\text{opt}}$ vs. Q (dashed line). For the figure we used the same parameter values as in Fig. 2.2.1. The insert in (b) shows that the solid and dashed lines intersect for a very small value of Q.

2.3. FRIO discrimination of two Rank N mixed states, POVM regime

In this section we present the more general solution to the two rank N density matrices, where N is some arbitrary but fixed integer. The two density matrices exhibit the Jordan basis as described in (2.2.1) and (2.2.2). Solving the problem is very similar to the method for two rank 2 mixed states in the preceding section.

The definitions and properties presented in Eqs. (2.2.2)–(2.2.11) remain in effect but the part starting with Eq. (2.2.12) has to be modified accordingly.The weights of the subspaces, introduced in (2.2.11) now satisfy

$$\sum_{i=1}^{N} \omega_i = 1 \ . \tag{2.3.1}$$

The total inconclusive rate can again be written as a weighted sum of the inconclusive rates of the individual subspaces,

$$Q = \sum_{i=1}^{N} \omega_i Q_i \ . \tag{2.3.2}$$

Q is fixed, with the fixed value satisfying $0 \leq Q \leq Q_0$, where now

$$Q_0 \equiv \sum_{i=1}^{N} \omega_i Q_{0,i} \ , \tag{2.3.3}$$

is the maximal inconclusive rate that the N subspaces can accommodate.

Similarly, the total error rate is a weighted sum of the error rates of the individual subspaces,

$$P_e = \sum_{i=1}^{N} \omega_i P_{e,i} \ . \tag{2.3.4}$$

The remaining task is to determine the optimal distribution of Q between the N subspaces, the distribution that minimizes the total error rate P_e, under the constraint that Q in Eq. (2.3.2) is fixed, i.e., to determine the optimal values of Q_i as a function of the fixed Q.

In order to perform the optimization we employ the Lagrange multiplier method because it leads to symmetric and easily tractable equations. Adding the Lagrange multiplier λ times

the constraint, $Q - \sum_{i=1}^{N} Q_i = 0$, to (2.3.4) yields the function

$$F = \sum_{i=1}^{N} \omega_i P_{e,i} + \lambda(Q - \sum_{i=1}^{N} \omega_i Q_i) \ , \tag{2.3.5}$$

where $P_{e,i}$ is inserted from Eq. (2.2.9), so it is also a function of Q_i. Next we vary F treating the variables $Q_1, Q_2, ..., Q_N$ as independent. Solving the resulting equations together with the constraint (2.3.2) determines the value of the Lagrange multiplier λ which in return optimizes 2.3.4.

Before we present the results it will prove useful to introduce a hierarchy of the subspaces. So, in what follows we will assume

$$Q_{0,1} \geq Q_{0,2} \geq \ldots \geq Q_{0,N} \ , \tag{2.3.6}$$

which generalizes the ordering used in the case of two subspaces in the previous section.

Then the optimal Q_i can be written as

$$Q_i^{opt} = \begin{cases} \frac{1-Q_{0,i}}{\sum_{i=1}^{k}\omega_i - \sum_{i=1}^{k}\omega_i Q_{0,i}} Q + \frac{Q_{0,i}\sum_{i=1}^{k}\omega_i - \sum_{i=1}^{k}\omega_i Q_{0,i}}{\sum_{i=1}^{k}\omega_i - \sum_{i=1}^{k}\omega_i Q_{0,i}} \\ \text{if } Q_{th}^{(k)} \leq Q \leq Q_{th}^{(k+1)} \text{ and } i \leq k \ , \\ 0 \\ \text{if } i > k \ . \end{cases} \tag{2.3.7}$$

Here $k = 1, 2, \ldots, N$ and we introduced the notation

$$Q_{th}^{(k)} = \frac{\sum_{i=1}^{k}\omega_i Q_{0,i} - Q_{0,k}\sum_{i=1}^{k}\omega_i}{1 - Q_{0,k}} \ . \tag{2.3.8}$$

and also $Q_{th}^{(N+1)} = Q_0 = Q_{max}$, cf. (2.3.3). Obviously, for $k = 1, 2$ these results reproduce the results for two subspaces, Eqs. (2.2.16)–(2.2.18).

Inserting the optimal failure rates into the subspace-error rates $P_{e,i}$, (2.2.9), gives the optimal error rates for the subspaces, $P_{e,i}^{opt}$. Then using these optimal subspace-error rates

in (2.3.4) gives the total optimal error rate $P_E = \sum_{i=1}^{N} P_{e,i}^{opt}$,

$$P_E = \begin{cases} \frac{1}{2}\left[1 - Q - \sqrt{\left(\sum_{i=1}^{k}\omega_i - Q\right)^2 - \left(\sum_{i=1}^{k}\omega_i Q_{0,i} - Q\right)^2} - \sum_{i=k+1}^{N}\sqrt{(\omega_i)^2 - (\omega_i Q_{0,i})^2}\right] & \text{if } Q_{th}^{(k)} \leq Q \\ & \text{and } k < N \\ \frac{1}{2}\left[1 - Q - \sqrt{(1-Q)^2 - (Q_0 - Q)^2}\right] & \text{if } Q_{th}^{(N)} < Q \end{cases}$$

which is valid in all regions of the parameter Q. For $N = 2$, (2.3.9) reduces to the two-subspaces solution, (**??**). For the maximum allowable inconclusive rate, $Q = Q_0$, the second line in Eq. (2.3.9) holds and it reduces to $P_E = 0$, corresponding to optimal Unambiguous Discrimination of the subspaces , while for $Q = 0$ the first line holds and it reduces to the Minimum Error expression for N subspaces , as expected.

Again, the most interesting aspect of the optimal solution is that a structure with multiple thresholds emerges. For $Q_{th}^{(1)} = 0 \leq Q < Q_{th}^{(2)}$ the total available inconclusive rate is accommodated by the first subspace only and all others operate at the Minimum Error level. According to the hierarchy introduced in Eq. (2.3.6), the first subspace is the one with the largest $Q_{0,i}$. Then between $Q_{th}^{(2)} \leq Q < Q_{th}^{(3)}$ the second subspace, the one with the second largest $Q_{0,i}$ will also participate in sharing the available inconclusive rate, while the remaining $N-2$ subspaces continue to operate at the minimum error level. In general, in the interval $Q_{th}^{(k)} \leq Q < Q_{th}^{(k+1)}$ the first k subspaces share the available inconclusive rate and the remaining $N-k$ subspaces remain at the minimum error level. Finally, in the range $Q_{th}^{(N)} \leq Q \leq Q_0 = Q_{max}$ all N subspaces participate in sharing the available inconclusive rate.

It is easy to show that the expressions are continuous at the threshold, i.e. the expressions valid below the threshold and the ones valid above the threshold tend to the same values at the threshold, although their slopes are, in general, different below and above the threshold. Furthermore, if $\omega_i Q_{0,i} > \omega_{i+1} Q_{0,i+1}$, the i^{th} subspace always accommodates more inconclusive rate than the $i+1^{st}$ for all $i = 1, 2, \ldots, N$, otherwise above some value of the total Q the

$i+1^{st}$ subspace will accommodate more inconclusive rate than the i^{th} subspace, the $Q_i^{opt}(Q)$ curves will intersect (see part (b) of Fig. 2.2.1 for an example).

We illustrate these results on the example of $N = 3$. In Fig. 2.3.1, we plot the optimal failure rates $\omega_i Q_i^{\rm opt}$ and in Fig. 2.3.2, we plot the optimal error rates $\omega_i P_{e,i}^{\rm opt}$ for the three subspaces as a function of the total failure rate Q for some specific values of the parameters.

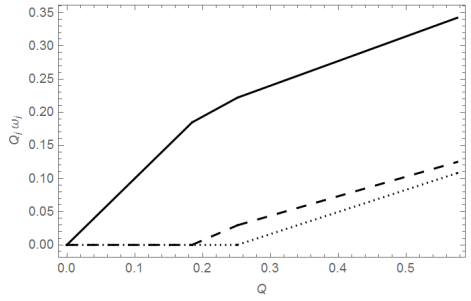


FIGURE 2.3.1. Optimal subspace failure rates $\omega_i Q_i^{opt}$ vs. Q from Eq. (2.3.7), for three subspaces ($i = 1, 2, 3$). $\omega_1 Q_1^{opt}$: solid line. $\omega_2 Q_2^{opt}$: dashed line. $\omega_3 Q_3^{opt}$: dotted line. For the figure we used $\eta_1 = \eta_2 = 1/2$ and the following parameter values: $\theta_1 = \pi/4$, $\theta_2 = \pi/3$, $\theta_3 = \pi/3$, $r_1 = 5/8$, $r_2 = 1/4$, $r_3 = 1/8$, $s_1 = 3/8$, $s_2 = 1/4$ and $s_3 = 3/8$.

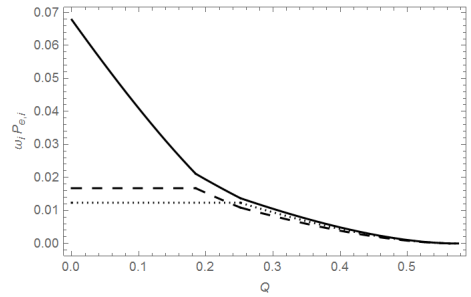


FIGURE 2.3.2. Optimal subspace error rates $\omega_i P_{e,i}^{opt}$ vs. Q from Eq. (2.2.9) with (2.3.7), for three subspaces ($i = 1, 2, 3$). $\omega_1 P_{e,1}^{opt}$: solid line. $\omega_2 P_{e,2}^{opt}$: dashed line. $\omega_3 P_{e,3}^{opt}$: dotted line. For the plots we used the same parameter values as for Fig. 2.3.1.

The results presented so far are valid if the parameters are such that in all subspaces we are in the POVM regime, i.e., Q_i is in the unshaded region (region II) of Fig. 1 for all i. When the parameters are such that in some subspaces we are in the projective regime, i.e., Q_i falls in the shaded regions of Fig. 1 (regions I and III) for some i we have to modify the

treatment to account for the fact that the error expression for the corresponding subspace, $P_{e,i}$, is no longer given by Eq. (2.2.9) but by (2.1.13). We will study this case in the next section.

2.4. Projective regime

We have seen for the single subspace case that the POVM solution is valid if the inconclusive rate is smaller than a boundary value, $Q \leq Q_b$, where Q_b is given by Eq. (2.1.11). With an obvious generalization, the POVM solution holds in subspace i if in that subspace $Q_i \leq Q_{b,i}$ holds where the subspace boundary value is given by

$$Q_{\mathrm{b,i}} \equiv 2\eta_{1,i}\eta_{2,i}\sin^2\theta_i/(1-Q_{0,i}). \tag{2.4.1}$$

In the region $Q_{b,i} \leq Q_i \leq Q_{c,i}$ the optimal measurement is a standard projective quantum measurement (SQM). The curves $Q_{b,i}$ and $Q_{c,i}$ intersect at $\eta_{1,i} = \eta_{1,i}^{(l)}$ and $\eta_{1,i} = \eta_{1,i}^{(r)}$, the same points as in Eq. (2.1.1). The interval $0 \leq \eta_{1,i} \leq 1$ is thus divided into three regions. In regions I and III, we have $Q_{\mathrm{b,i}} < Q_{\mathrm{c,i}}$ and the solution (2.1.10) is valid for $0 \leq Q_i < Q_{\mathrm{b}}$ only. In Region II, $\eta_{1,i}^{(l)} \leq \eta_{1,i} \leq \eta_{1,i}^{(r)}$, we have $Q_{\mathrm{c,i}} = Q_{0,i} < Q_{\mathrm{b,i}}$ and the solution (2.1.10) is valid for the entire $0 \leq Q_i \leq Q_{\mathrm{c,i}}$ range.

Thus, Fig. 2.1.1 is valid in every subspace, with the obvious change of axis labels to Q_i and $\eta_{1,i}$. So, $Q_i > Q_{b,i}$ occurs in regions I and III and in the shaded areas the optimal FRIO measurement is an SQM while in the unshaded area it is a POVM. We now illustrate the case when the FRIO measurement is a POVM in one and an SQM in the other subspace on an example.

The optimal distribution of the total available inconclusive rate Q between the two subspaces, Q_1^{opt} and Q_2^{opt} such that their sum satisfies $Q_1^{opt}+Q_2^{opt} = Q$, can be found by optimizing the total error rate with respect to the failure rate of the subspaces. We now have to use the error expression (2.1.13) in subspace 1 for $Q > Q_{b,1}$. This leads to a numerical optimization problem, the result of which is shown in Fig. 2.4.1.

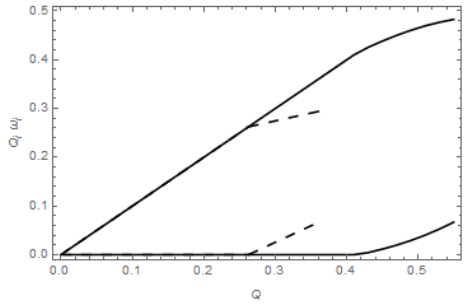


FIGURE 2.4.1. Optimal subspace failure rates $\omega_i Q_i^{opt}$ vs. Q from Eq. (2.3.7), for two subspaces ($i = 1, 2, 3$). $\omega_1 Q_1^{opt}$: solid upper line. $\omega_2 Q_2^{opt}$: solid lower line. The dashed lines indicate the behavior of the corresponding quantities if the POVM solution were valid for subspace 1. For the figure we used $\eta_1 = \eta_2 = 1/2$ and the following parameter values: $\theta_1 = \pi/16$, $\theta_2 = 3\pi/7$, $r_1 = 9/10$, $r_2 = 1/10$, $s_1 = 1/10$ and $s_2 = 9/10$.

The solution again exhibits a threshold structure. The dashed lines in the figure indicate how the optimal Q_i's would behave if the POVM solution were valid for subspace 1 instead of the SQM. It is apparent that the SQM shifts the threshold toward a higher value of Q and above the threshold the dependence on Q is nonlinear, while for the POVM regime it is always linear.

The optimal inconclusive rate for subspace i, Q_i^{opt}, displayed in Fig 2.4.1, is then inserted into Eq. (2.1.13) for $i = 1$ and (2.2.9) for $i = 2$. This yields the optimal error rate for subspace i, $P_{e,i}^{opt}$. The results are displayed in Fig. 2.4.2.

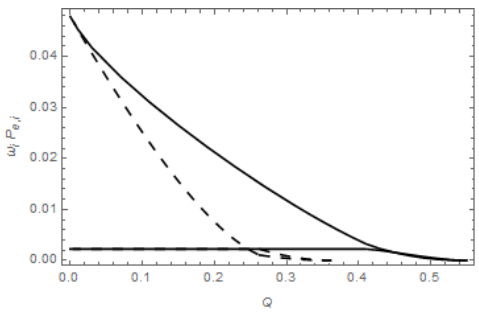


FIGURE 2.4.2. Optimal subspace error rates $P_{e,i}^{opt}$ vs. Q from Eq. (2.3.7), for two subspaces ($i = 1, 2$). $P_{e,1}^{opt}$: solid upper line. $P_{e,2}^{opt}$: solid lower line. The dashed lines indicate the corresponding quantities if the POVM solution were valid in subspace 1. For the plots we used the same parameter values as for Fig. 2.4.1.

The optimal error rates also exhibit a threshold structure. Relative to the POVM regime, the SQM shifts the threshold to a higher value of Q and the overall error rate also increases.

2.5. Summary and conclusion

We have found analytic solutions for the optimal discrimination measurement strategy when a fixed rate of inconclusive outcomes, Q, is allowed (FRIO strategy), for a class of mixed states that exhibit a Jordan Basis structure. Thus our work extends the previously introduced FRIO strategy from pure states **[14]** to mixed states. In this strategy the probability of making an error in identifying the state, P_e, is minimized for a fixed Q and the solution optimally interpolates between the minimum error ($Q = 0$) and unambiguous discrimination ($P_e = 0$) limits. We found several surprising and unexpected conclusions. The first is that the form of the optimal error rate remains formally the same over all subspaces as in the case of two pure states, which is a consequence of the Jordan structure of the mixed states. The second is a more striking feature, the emergence of a threshold structure: as we increase the allowed inconclusive rate Q, starting from $Q = 0$, at first only one subspace accommodates all of the available Q. Above a certain threshold a second subspace starts sharing the available Q, above yet another higher threshold a third subspace starts participating, etc., until above a final threshold all of the subspaces participate. It is interesting to note that in this last regime the optimal error expression the second line in Eq. (2.3.9) is formally the same as the result for two pure states. This is a novel type of behavior and allows for experimental tests of our findings. Applications could be considered in cryptography where a key is shared over different lines to enhance security without sacrificing overall error rate.

CHAPTER 3

Quantum Cloning

One of the reasons we need to develop optimum state discrimination measurement schemes is due to the no cloning theorem of Wootters, Zurek [**19**] and Dieks [**20**]. If one could copy non-orthogonal quantum states then by making a very large number of copies, it would be possible to distinguish the states. However cloning machines which optimize some criteria with a limited degree of success have been developed. Those cloning machines fall under two categories: universal and state dependent. Universal cloning machines, which make copies of a completely unknown quantum state, were developed first by Buzek and Hillery [**21**]. This scheme makes approximate copies of an unknown quantum state while optimizing the local fidelity which is the square overlap of the approximate clones and the original state it is supposed to clone. The other category is state dependent cloning machines. In this scheme the observer has full knowledge of the prepared states but does not know which is the state he is given. There are two subcategories within this scheme: approximate and exact cloning. Approximate state-dependent cloning machines deterministically generate approximate clones from a finite set of non-orthogonal quantum states while optimizing the local or global fidelity (the average square overlap between full set of approximate clones and the states to be cloned). Hillery and Buzek [**22**] are the pioneers of this subcategory of cloning machines as well. In exact state-dependent cloning machines, the other subcategory of quantum cloning machines, the task is to probabilistically make exact copies of the incoming non-orthogonal quantum states. This comes at the expense of allowing for failure results where the scheme fails to produce a copy altogether. Duan and Guo [**23**] were first to develop probabilistic exact cloning machines for the two state input where the states are prepared with equal a priori probabilities. We recently extended this method for the more general case where the a priori probabilities of the incoming states are different []. This extension not only solves the

full problem but gives new insight into the nature of quantum cloning. The symmetry of the equal priors case completely solves the problem and no further optimization can be done. This symmetry however hides the true nature of the exact cloning machines which show up in the unequal priors case. This can be shown through a two step process: exact cloning then optimal UD on the clones. First procedure makes exact clones of the incoming two states. The exact clones are then sent to an optimal UD machine where the average failure rate is minimized. To combine the two step process the inconclusive rate from the cloning process is added to the inconclusive rate, weighed with the new a-priori probabilities, from the optimal UD. When the input states are prepared with equal a-priors the total amount of the inconclusive rate reaches the IDP limit. Hence cloning then performing optimal UD is equivalent to simply performing the optimal UD first, then prepare the clones. However this is not true for when the priors of the input states are different. After the two step process the total inconclusive rate is higher then the IDP limit. This suggests that during the cloning process some information is being leaked due to the asymmetry of the failure rate operators. When performing exact cloning, clones are produced but no measurement has been made, hence we do not know which states are being cloned. We simply know whether the procedure was successful or it failed. When it fails, the states are discarded. This is where the information leakage comes in. The state which is prepared most often shall have a higher rate of failure. This does not happen in the equal priors case because the failure rate are symmetric.

3.1. No-Cloning Theorem

Some of the schemes described in this dissertation related to discriminate quantum states would not be necessary if one could make copies of the non-orthogonal quantum states as can be done with classical states. If this were possible then the receiver, Bob, after receiving the state $|\psi_1\rangle$ or $|\psi_2\rangle$ from the preparer, Alice, makes n number of copies. After a large set of copies the states become nearly orthogonal and almost fully distinguishable. The average inconclusive rate of failing to distinguish the n copies of $|\psi_1\rangle$ or $|\psi_2\rangle$ is $Q_o = 2\sqrt{\eta_1\eta_2}s^n$. For

a large n the inconclusive rate is very small and Bob can discriminate nearly all incoming states.

Thus while in classical information it is possible to make exact copies of information, as this dissertation is printed on this paper, multiple times by a printer. The no-cloning theorem forbids the receiver doing the same with quantum states. More specifically it is non-orthogonal quantum states which cannot be copied, as classical states are a special case of quantum states, that of orthogonal states.

Let us now show a proof by contradiction of why such a quantum cloning machine cannot exist. Suppose there is such a cloning machine with an input and an output port. Inside the machine there are two slots: slot S for the system state $|\psi_i\rangle$ to be copied, and slot A for the ancilla state for where the input state is to be copied. Let the ancilla be in some blank space $|0\rangle$, then the initial state of the copying machine would be: $|\psi_i\rangle|0\rangle$. A unitary operator would copy the state $|\psi_i\rangle$ into $|0\rangle$:

$$U|\psi_i\rangle|0\rangle = |\psi_i\rangle|\psi_i\rangle \tag{3.1.1}$$

Let there be two possible input states to be copied, $\{|\psi_1\rangle, |\psi_2\rangle\}$ and we are interested in a quantum cloning machine which produces $|\psi_1\rangle|\psi_1\rangle$ when $|\psi_1\rangle$ is sent and $|\psi_2\rangle|\psi_2\rangle$ when $|\psi_2\rangle$ is sent. The unitary operator would do the following.

$$\begin{aligned} U|\psi_1\rangle|0\rangle &= |\psi_1\rangle|\psi_1\rangle \\ U|\psi_2\rangle|0\rangle &= |\psi_2\rangle|\psi_2\rangle \end{aligned} \tag{3.1.2}$$

The inner product of these equations gives $\langle\psi_2|\psi_1\rangle\langle 0|0\rangle = |\langle\psi_2|\psi_1\rangle|^2 \Rightarrow s = s^2$. This condition can be satisfied only if $s = 0$, states are orthogonal, or $s = 1$, the two states are the same. But we said that the two states are distinct and non-orthogonal. Thus one cannot design a unitary device which makes perfect clones of an unknown quantum system deterministically. Here we only proved that non-orthogonal pure states cannot be copied

through a unitary process. Other proofs exist which show that this holds for mixed states and also for other non-unitary processes.

Since such a machine cannot be designed the next logical step is to build a quantum machine which produces clones similar to the input states while allowing for some fidelity or inconclusive results. We derive some previous results considering such quantum cloning machines and also provide new results for optimal exact cloning with some inconclusive rate allowed. In a follow up section we interpolate between the deterministic cloning scheme and exact cloning by relaxing some of the conditions.

3.2. Exact cloning with failure rate

In probabilistic cloning we are concerned with making exact copies of the given quantum state. It is indeed possible to make exact copies but only probabilistically. This is similar to the Unambiguous Discrimination where one discriminates between a given set of non-orthogonal quantum states unambiguously while allowing for a failure rate.

We approach probabilistic cloning via the Neumark formulation. The system is embedded in a larger Hilbert space where the extra degrees of freedom are customarily called the ancilla. Then a unitary transformation entangles the system degrees of freedom with those of the ancilla. The input states $\{|\psi_1^M\rangle_s, |\psi_2^M\rangle_s\}$ which live in the state Hilbert space S are embedded with the ancilla $|i\rangle_a$ which live in the $N-M$ dimensional Hilbert space A. Now the system and the ancilla live in the larger Hilbert space $H = S \otimes A$. The incoming states in this larger Hilbert space can be written in the product form $\{|\psi_1^N\rangle_s|i\rangle_a, |\psi_2^N\rangle_s|i\rangle_a\}$. The unitary should do the following:

$$U|\psi_1^M\rangle|i\rangle^{N-M} = \sqrt{p_1}|\psi_1^N\rangle|\alpha_1\rangle + \sqrt{q_1}|\Phi_1\rangle, \tag{3.2.1}$$

$$U|\psi_2^M\rangle|i\rangle^{N-M} = \sqrt{p_2}|\psi_2^N\rangle|\alpha_2\rangle + \sqrt{q_2}|\Phi_2\rangle, \tag{3.2.2}$$

Where p_i is the rate of having successfully produced a perfect clone, q_i is the rate of having failed to do so. The inner product of each of the above equation with its own transpose gives

the normalized probabilities $p_i + q_i = 1$. $|\psi_i\rangle$ are the input states we wish to clone and their corresponding ancillas are $|\alpha_i\rangle$. Getting a click in $|\Phi_i\rangle$ means we have failed to clone and discard the state.

Taking the inner product of the transpose of equation (3.2.1) with (3.2.2) we get the constraint of success and failure rate in terms of the overlap of the input states.

$$s^M = \sqrt{p_1 p_2} s^N \alpha + \sqrt{q_1 q_2}\phi \tag{3.2.3}$$

where:

$$s = \langle\psi_1|\psi_2\rangle, \quad s' = \langle\Psi_1|\Psi_2\rangle, \quad \alpha = \langle\alpha_1|\alpha_2\rangle, \quad \phi = \langle\Phi_1|\Phi_2\rangle$$

In the above setup the ancilla states corresponding to successfully making the clones are distinct. The failure states are also distinct. That is the most general case. In order for the success rates to be optimal we take the ancilla states to be the same. Looking at the constraint we absorb α into s, meaning there is more copies which would come at the expense of a lower p_i . More explicitly this can be proved for the equal priors case shown below. We will now prove that the failure states should also be the same. If $|\Phi_1\rangle$ and $|\Phi_2\rangle$ are different that means that there is still information left in the failure states and we can still perform Unambiguous State discrimination and probabilistically determine whether we received $|\psi_1\rangle$ or $|\psi_2\rangle$. The optimal strategy is one which leaves no information at all in the failure states, the overlap of the failure states can be set to one $\langle\Phi_2|\Phi_1\rangle = 1$. Any click from the failure state is simply discarded as failure. The revised Neumark setup reduces to:

$$\begin{aligned} U|\psi_1^M\rangle|i\rangle &= \sqrt{p_1}|\psi_1^N\rangle\,|\alpha\rangle + \sqrt{q_1}|\Phi\rangle\,|0\rangle\,, \\ U|\psi_2^M\rangle|i\rangle &= \sqrt{p_2}|\psi_2^N\rangle|\alpha\rangle + \sqrt{q_2}|\Phi\rangle\,|0\rangle\,, \end{aligned} \tag{3.2.4}$$

The constraint simplifies into:

$$s^M = \sqrt{p_1 p_2} s^N + \sqrt{q_1 q_2}. \tag{3.2.5}$$

This is the constraint on the parameters to optimize the average failure rate $Q = \eta_1 q_1 + \eta_2 q_2$. The optimization is straightforward for the symmetric case where the input states are prepared with equal priors. The general case, states prepared with different a priori probabilities, leads to a sixth order equation the solution of which isn't very useful. A geometric picture however emerges which is more insightful.

3.2.1. Equal priors. When the input states are prepared with equal prior probabilities, $\eta_1 = \eta_2 = \frac{1}{2}$, the problem is quite trivial. It was first solved by Duan and Guo **[23]**. For equal priors the success and failure rate reduce to: $p_1 = p_2 = p$, $q_1 = q_2 = q$. The constraint in (3.2.5) reduces to:

$$s^M = ps^N + q, \tag{3.2.6}$$

using the unitarity condition $p + q = 1$, the problem can be fully solved:

$$\begin{aligned} s^M &= ps^N + (1-p), \\ p &= \frac{1-s^M}{1-s^N} \end{aligned} \tag{3.2.7}$$

and the failure rate is $q = 1 - p = \frac{s^M - s^N}{1-s^N}$.

Thus the optimization of the probabilistic exact cloning is fully solved from the constraint derived from the Neumark method and no further optimization is needed. In the limit of producing infinitely many clones, $N \to \infty$, the failure rate reduces to $q = s^M$, which is also the IDP result for unambiguous state discrimination.

3.2.2. Unequal priors I. There are a number of reasons why one might want to solve the general problem with arbitrary priori probabilities. (i) The solution to the equal priors problem is obtained using only symmetry arguments, with no need for optimization. (ii) A general solution would check the robustness of the equal priors case against variations of the prior probabilities around 1/2. This gives control over errors that are unavoidable for any physical realization. (iii) One could consider a discrimination protocol consisting of

optimal cloning followed by optimal Unambiguous Discrimination (UD) of the produced clones, which we will call discrimination by cloning. Surprisingly this is optimal for equal prior probabilities and for any number of clones (see below). This suggests that the equal priors case is very special and can provide a deceptive view of cloning. (iv) In the limit of infinitely many clones, the optimal strategy prepares the clones according to the outcomes of UD of the input states. This is a particular case of a measure and prepare protocol, which we will call cloning by discrimination. Since the UD measurement varies over the range of prior probabilities (a 3-outcome generalized measurement vs a 2-outcome projective measurement), this hints at the possibility of a similar situation for optimal cloning that can only be decided by solving the general problem.

Our solution shows that discrimination by cloning as outlined in (iii) is sub-optimal for unequal prior probabilities (unless one state is never sent). This indicates that the equal prior case is not representative of state dependent cloning. Additionally, contrary to the suggestions in (iv) above, our solution leads to a failure probability that is a smooth function of the priors. However, the strategy converges to cloning by discrimination as $N \to \infty$, implying a discontinuous second derivative of the total inconclusive rate with respect to η_i and revealing a phenomenon similar to a second order phase transition.

We can imagine 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 conditioned probability p_i. Otherwise, the output is in a refuse state. Conditioned on the input state being $|\psi_i^M\rangle$, the failure probability is $q_i = 1 - p_i$.

For cloning, optimality is usually addressed from a Bayesian viewpoint that assumes the states to be cloned are given with some prior 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

$$Q = \eta_1 q_1 + \eta_2 q_2. \tag{3.2.8}$$

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

In our formulation, similar to that in **[23]**, the Hilbert space $\mathscr{H}^{\otimes M}$ of the original M copies is supplemented by an ancillary space $\mathscr{H}^{\otimes(N-M)} \otimes \mathscr{H}_F$ that accommodates both the additional $N-M$ clones as well as the success/failure flags. Then, a unitary transformation U (time evolution) from $\mathscr{H}^{\otimes M} \otimes \mathscr{H}^{\otimes(N-M)} \otimes \mathscr{H}_F$ onto $\mathscr{H}^{\otimes N} \otimes \mathscr{H}_F$ is defined through **[23]**

$$U|\psi_i^M\rangle|0\rangle = \sqrt{p_i}|\psi_i^N\rangle|\alpha_i\rangle + \sqrt{q_i}|\Phi\rangle, \quad i = 1, 2. \tag{3.2.9}$$

Here the ancillas are initialized in a reference state $|0\rangle$. The states of the flag associated with successful cloning $|\alpha_i\rangle$ are constrained to be orthogonal to the refuse state $|\Phi\rangle$ for certainty in the outcomes of the projective measurement on the flag space $\{\mathscr{H}_F\}$. Although for optimal cloning $|\alpha_1\rangle = |\alpha_2\rangle$, we need to consider a more general setup where these two states are different to include the cloning-by-discrimination protocol where UD is used to identify the input state and then the clones are prepared accordingly [1]. For UD the success flag states must be distinguishable, so $\langle\alpha_1|\alpha_2\rangle = 0$. Taking the inner product of each equation with itself shows that our probabilities are normalized: $p_i + q_i = 1$. Similarly, by taking the product of the two equations in (3.2.9), we find the unitarity constraint

$$s^M = \sqrt{p_1 p_2}\, s^N \alpha + \sqrt{q_1 q_2}, \tag{3.2.10}$$

where the overlaps $s = \langle\psi_1 \mid \psi_2\rangle$, $0 \leq s \leq 1$ and $\alpha = \langle\alpha_1|\alpha_2\rangle$ can be chosen to be real valued without any loss of generality. . We note that for optimal cloning one has $\alpha = 1$, whereas $\alpha = 0$ for cloning by discrimination. If Eq. (3.2.10) is satisfied, it is not hard to prove that U has a unitary extension on the whole space.

Before attempting to minimize Q, we need to gain geometric insight into the meaning of the unitary constraint. The following points turn out to be important: (a) For fixed s, N

[1] Likewise, we could consider a more general setup with two refuse states $|\Phi_1\rangle$ and $|\Phi_2\rangle$ in Eqs. (3.2.9). This is necessarily sub-optimal since we could probabilistically determine whether we received $|\psi_1\rangle$ or $|\psi_2\rangle$ by applying UD to the refuse states $|\Phi_i\rangle$. Sometimes we would be certain of the input state, when we can always prepare n copies of the state, thereby increasing the overall success rate of the cloning strategy.

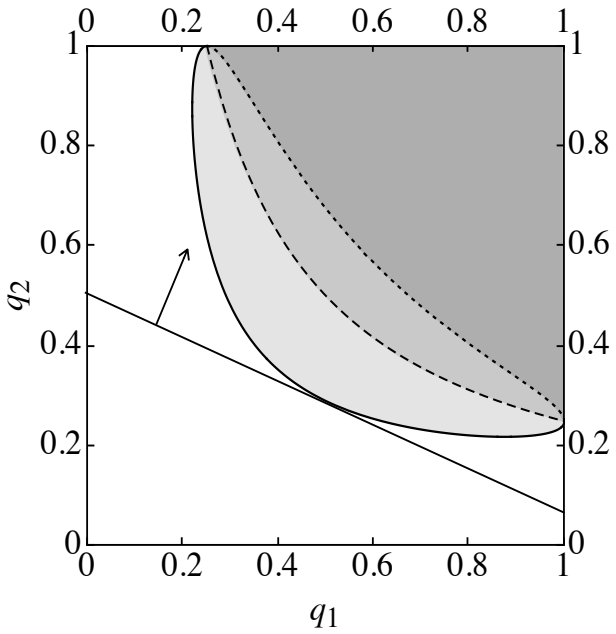


FIGURE 3.2.1. Unitarity curves in Eq. (3.2.10) and the associated sets S_α in Eq. (3.2.11) for values of α positive (solid/light gray), zero (dashed/medium gray), and negative (dotted/dark gray). The figure also shows the optimal straight segment $Q = \eta_1 q_1 + \eta_2 q_2$ and its normal vector (η_1, η_2). Plotted for $s = 0.5$, $m = 1$, $n = 2$, $\alpha = 0.8, 0, -0.8$.

and M Eq. (3.2.10) defines a class of smooth curves on the unit square $0 \leq q_i \leq 1$ (e.g., solid, dashed or dotted curves in Fig. 3.2.1). (b) All these curves meet at their endpoints, $(1, s^{2M})$ and $(s^{2M}, 1)$. (c) At the endpoints the curves become tangent to the vertical and horizontal lines $q_1 = 1$ and $q_2 = 1$ respectively, provided $\alpha \neq 0$. the curve $q_1 q_2 = s^{2M}$ (dashed line in Fig. 3.2.1). (e) Each of these curves and the segments joining their end points with the vertex $(1, 1)$ are the boundary of the sets (any of the gray regions in Fig. 3.2.1)

$$S_\alpha = \{(q_1, q_2) : \sqrt{p_1 p_2}\, s^N \alpha + \sqrt{q_1 q_2} - s^M \geq 0\}. \tag{3.2.11}$$

They satisfy $S_\alpha \subset S_{\alpha'}$ if $\alpha < \alpha'$. (f) Moreover, the sets S_α are convex if $\alpha \geq 0$. In particular S_1 is convex.

At this point a geometrical picture of the optimization problem emerges (See Fig. 3.2.1). Eq. (3.2.8) defines a straight segment on the square $0 \leq q_i \leq 1$ with a normal vector in the first quadrant parallel to (η_1, η_2). For fixed prior probabilities, the average failure

probability Q is proportional to the distance from this segment to the origin $(0,0)$. Since S_1 is convex and the stretch of its boundary given by Eq. (3.2.10) with $\alpha = 1$ is smooth, a unique point (q_1, q_2) of tangency with the segment (3.2.8) exists for any value of the priors and finite N. It gives $Q_{\min}$ and defines the optimal cloning strategy.

We note in passing that the inclusion hierarchy of the sets S_α provides a simple geometrical proof that $\alpha = 1$, i.e., $|\alpha_1\rangle = |\alpha_2\rangle$, is indeed the optimal choice. On the other hand, we recall that for cloning by discrimination we have $\langle\alpha_1|\alpha_2\rangle = \alpha = 0$. From points (b) and (c) above, it follows that for any finite N and arbitrary priors η_1 and η_2 this protocol is strictly suboptimal, i.e., $Q_{\min} < Q_{\rm UD}$, where the subscript UD is a reminder that the failure rate of cloning by discrimination is that of UD. One could say that optimal cloning is incompatible with discerning the identity of the input states for any finite number of clones. However, optimal cloning and UD become one and the same in the limit $N \to \infty$, where $s^N \to 0$ and the curve (3.2.10) collapses to the hyperbola $q_1 q_2 = s^{2M}$, as it does for $\alpha = 0$. We will come back to this point below. A more quantitative analysis requires finding a convenient parametrization of the curve (3.2.10). To this end, simpler and more manageable expressions are derived if the symmetry under $q_1 \leftrightarrow q_2$ is preserved. We write $\sqrt{q_i} = \sin\theta_i$ for $0 \le \theta_i \le \pi/2$. By further introducing the variables $x = \cos(\theta_1 + \theta_2)$ and $y = \cos(\theta_1 - \theta_2)$ we manage to linearize the constraint (3.2.10), which now reads as $2s^M = (1+s^N)y - (1-s^N)x$. A natural parametrization for this straight line is given by

$$x = \frac{1-(1+s^N)t}{s^{N-M}}, \qquad y = \frac{1-(1-s^N)t}{s^{N-M}}, \tag{3.2.12}$$

where again we have taken the most symmetrical choice. 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 (3.2.10) under this transformation are mapped into the same straight line (3.2.12). 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 $q_1 \leftrightarrow q_2$.

The allowed domain of t in Eq. (3.2.12) follows from that of x and y, readily seen from their definition to be the region $|x| \leq y \leq 1$. Hence, we have

$$\frac{1 - s^{n-m}}{1 - s^n} \leq t \leq 1. \tag{3.2.13}$$

After putting the various pieces together one can easily get rid of the trigonometric functions and express Eq. (3.2.10) in parametric form as

$$q_i = \frac{1 - xy - (-1)^i \sqrt{1 - x^2}\sqrt{1 - y^2}}{2}, \quad i = 1, 2. \tag{3.2.14}$$

Fig. 3.2.2 shows examples of the unitary curve (3.2.10) for (a) $n = 2$ and (b) $n = 5$. In both cases $m = 1$. For larger n the curves closely approximate the hyperbolae $q_1 q_2 = s^{2m}$ (dashed lines) for small and moderate values of s, while for s close to one the hyperbolas remain closer to the vertex $(1, 1)$, but still retain the same end points. As mentioned previously, in the limit $n \to \infty$ all curves become hyperbolic.

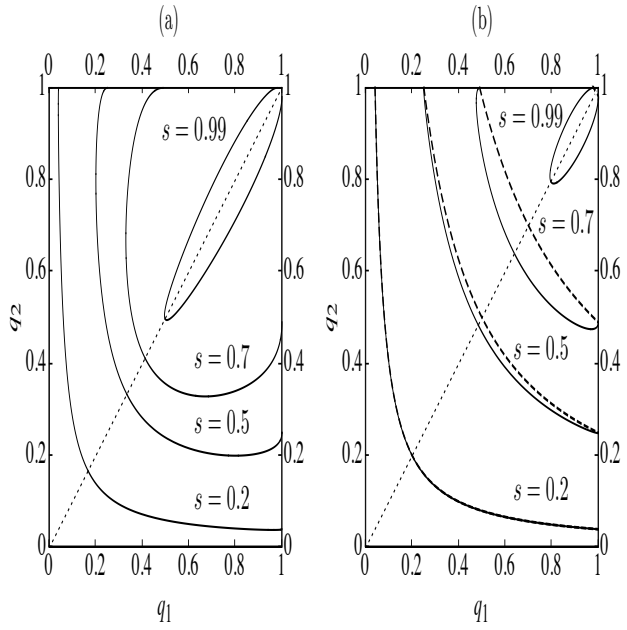


FIGURE 3.2.2. Unitarity curves for different values of s and for (a) $m = 1$, $n = 2$ and (b) $m = 1$, $n = 5$. The curves are symmetric under mirror reflexion along the (dotted) straight line $q_1 = q_2$, i.e., under the transformation $q_1 \leftrightarrow q_2$. The dashed lines in (b) are the hyperbolae $q_1 q_2 = s^{2m}$.

Now we can return to the minimization of the average failure probability Q. Despite the apparent simplicity of the problem, finding the minimum Q as an explicit function of η_1 or η_2 involves solving a quartic equation without a simple form Instead, we will derive the parametric equation of the curve $(\eta_1, Q_{\min})$. This, along with our complete description of the unitary curve , provides a full account of the solution.

Without any loss of generality we may assume that $\eta_1 \leq \eta_2$, or equivalently, that $0 \leq \eta_1 \leq 1/2$. Then the slope of the vector normal to the straight line (3.2.8) is less or equal to one and thus it can only become tangent to the lower half of the unitary curve (3.2.10) (see Fig. 3.2.2). The slope of this lower half increases monotonically as we move away from the line $q_1 = q_2$, where it has the value -1, and vanishes before we reach the line $q_1 = 1$. This follows from the properties (a)–(f) above and can be checked using Eq. (3.2.14). The values of t at which the slope is -1 and 0 are respectively

$$t_{-1} = \frac{1 - s^{n-m}}{1 - s^n}, \quad t_0 = \frac{1 - s^{2(n-m)}}{1 - s^{2n}}, \tag{3.2.15}$$

where we note that t_{-1} is the lower value of the range of t in Eq. (3.2.13). For any point $(q_1(t), q_2(t))$ with $t \in [t_{-1}, t_0]$ there is a line $Q = \eta_1 q_1 + \eta_2 q_2$ that is tangent to it, starting with $\eta_1 = \eta_2 = 1/2$ for $t = t_{-1}$ up to $\eta_1 = 0$, $\eta_2 = 1$ for $t = t_0$.

This observation enables us to derive the desired parametric expression for the optimality curve $(\eta_1, Q_{\min})$ 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. (3.2.8) with (3.2.14) we enforce contact with the unitarity curve and obtain the expression of $Q_{\min}$. The final result can be cast as:

$$\eta_1 = \frac{q_2'}{q_2' - q_1'}, \quad Q_{\min} = \frac{q_2' q_1 - q_1' q_2}{q_2' - q_1'}, \quad t_{-1} \leq t \leq t_0, \tag{3.2.16}$$

where t_{-1}, t_0 and q_i are given in Eqs. (3.2.15) and (3.2.14). The expressions for the derivatives q_i' are

$$q_i' = \frac{\sqrt{q_i(1 - q_i)}}{s^{n-m}} \left\{ \frac{1 + s^n}{\sqrt{1 - x^2}} - (-1)^i \frac{1 - s^n}{\sqrt{1 - y^2}} \right\}. \tag{3.2.17}$$

Fig. 3.2.3 shows plots of the curves $(\eta_1, Q_{\min})$ 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. (3.2.15), into Eq. (3.2.14). They are given by

$$Q_0 = q_2(t_0) = \frac{s^{2m} - s^{2n}}{1 - s^{2n}}, \quad Q_{-1} = \frac{s^m - s^n}{1 - s^n}, \tag{3.2.18}$$

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 dashed

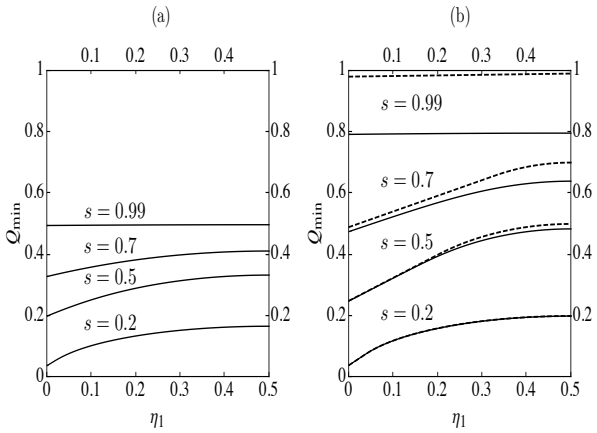


FIGURE 3.2.3. Minimum cloning failure probability $Q_{\min}$ vs. η_1 (solid lines) and UD failure probability $Q_{\rm UD}$ vs. η_1 (dashed lines) for the same values of m, n and s used in the previous figure.

lines in Fig. 3.2.3 (b) are the well known piecewise unambiguous discrimination solution **[10]**:

$$Q_{\rm UD} = \begin{cases} \eta_1 + s^{2m}\eta_2, & 0 \le \eta_1 \le \dfrac{s^{2m}}{1+s^{2m}}; \\ 2\sqrt{\eta_1\eta_2}\, s^m, & \dfrac{s^{2m}}{1+s^{2m}} \le \eta_1 \le \dfrac{1}{2}. \end{cases} \tag{3.2.19}$$

It is apparent from these plots that the optimal cloning protocol performs strictly better than cloning by discrimination, as was proved above. However, as the number of produced clones becomes larger the difference in performance reduces. In Fig. 3.2.3 (b), for only $n = 5$, a difference is hardly noticeable for $s \le 0.5$. For larger overlaps it takes larger values of n to get the same level of agreement. As discussed above, in the limit $n \to \infty$ there is perfect agreement for any $s < 1$.

The complete UD solution in Eq. (3.2.19) emerges naturally from our geometrical approach in a straightforward manner: First, we recall that in this case the right hand side of Eq. (3.2.10) becomes $q_1 q_2 = s^{2m}$ (dashed lines in Figs. 3.2.1 and 3.2.2). The maximum slopes of these curves are at their end points and all have the value $-s^{2m}$. This implies that the boundary of S_0 has a cusp at $(1, s^{2m})$. It follows that a unique point of tangency with the line (3.2.8) exists for $s^{2m} < \eta_1/\eta_2 \leq 1$ (recall that we are assuming $\eta_1 \leq 1/2$). This condition gives the η_1 interval for that solution. The tangency condition, $(q_2, q_1) \propto (\eta_1, \eta_2)$, quickly leads us to the optimal failure rate in the second line of Eq. (3.2.19). For $s^{2m} > \eta_1/\eta_2 \geq 0$ tangency is not possible, and the optimal line (3.2.8) merely touches the cusp on the boundary of S_0, so the expression of Q becomes the first line of Eq. (3.2.19). In geometrical terms, the straight line (3.2.8) pivots on the end point as η_1 varies between 0 and $s^{2m}/(1 + s^{2m})$.

For the second case in Eq. (3.2.19), one has $q_1, p_1 \in (0, 1)$ and there are three orthogonal flag states in Eqs. (3.2.9), namely, the two success states $|\alpha_1\rangle$, $|\alpha_2\rangle$, and the failure state $|\Phi\rangle$. This 3-outcome measurement can be represented by a 3-element positive operator valued measure (POVM) on $\mathscr{H}^{\otimes m}$. For the first line in Eq. (3.2.19), $p_1 = 1 - q_1 = 0$, which leads to a 2-outcome projective measurement, as only one success flag state ($|\alpha_2\rangle$) is needed in Eqs. (3.2.9).

This two-paragraph derivation of Eq. (3.2.19) proves that the convergence of the optimal cloning failure probability $Q_{\min}$ to that of cloning by discrimination, with rate $Q_{\rm UD}$ in Eq. (3.2.19), follows from the convergence of the general unitarity curve in Eq. (3.2.10) to the hyperbola $q_1 q_2 = s^{2m}$, i.e., from $\lim_{\alpha \to 0} S_\alpha = S_0$. Interestingly enough, such convergence entails a phenomenon analogous to a second order phase transition. Our geometrical approach shows that the average failure probability $Q_{\min}(\eta_1)$ is an infinitely differentiable function of η_1 for finite n. However, as n goes to infinity (or at $\alpha = 0$, for the sake of this discussion) the limiting function $Q_{\rm UD}(\eta_1)$ has a discontinuous second derivative. Moreover, the symmetry $q_1 \leftrightarrow q_2$ breaks in the phase corresponding to the first line in Eq. (3.2.19). A similar phenomenon arises in UD of more than two pure states [**24**] .

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. (3.2.18) 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. (3.2.19) 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. (3.2.19). If $\eta_1 \to 0$ then only $|\psi_2^n\rangle$ is produced with non-vanishing probability and $Q_{\rm C} = Q_0$. Failure in the second step (UD) is given by the top line in Eq. (3.2.19) applied to n copies. The total failure rate is s^{2m}, also achieving optimality.

Using our main result in Eqs. (3.2.16), and (3.2.14) 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.

3.2.3. Unequal priors II. In this section we show a different geometric/numeric solution using the Lagrange multipliers method with the constraint given in Equation (3.2.5),

$$s = \sqrt{(1-q_1)(1-q_2)}s^2 + \sqrt{q_1 q_2}, \tag{3.2.20}$$

where we set $M = 1$, $N = 2$ for notation simplicity. This corresponds to one to two cloning but can be easily generalized for M to N cloning, $M \geq N$. The function to be maximized is:

$$F_S = \eta_1 q_1 + \eta_2 q_2 + \lambda \left[s - \sqrt{(1-q_1)\,(1-q_2)}s^2 - \sqrt{q_1 q_2} \right], \tag{3.2.21}$$

where λ is the Lagrange multiplier to be determined which optimizes the failure rate $Q = \eta_1 q_1 + \eta_2 q_2$. Set $\partial F/\partial q_i = 0$, solve for q_i,

$$\frac{2\eta_1}{\lambda} = \sqrt{\frac{q_2}{q_1}} - \sqrt{\frac{1-q_2}{1-q_1}} s^2, \tag{3.2.22}$$

$$\frac{2\eta_2}{\lambda} = \sqrt{\frac{q_1}{q_2}} - \sqrt{\frac{1-q_1}{1-q_2}} s^2. \tag{3.2.23}$$

Let $A \equiv \sqrt{\frac{q_2(1-q_1)}{q_1(1-q_2)}}$,

$$\sqrt{\frac{q_1}{q_2}} = \frac{\lambda}{2\eta_1}\left[1 - \frac{s^2}{A}\right], \tag{3.2.24}$$

$$\sqrt{\frac{q_2}{q_1}} = \frac{\lambda}{2\eta_2}\left[1 - As^2\right], \tag{3.2.25}$$

Multiplying the above two equations, setting $\delta = \frac{4\eta\eta_2}{\lambda^2}$ and $C = \frac{s^4 - \delta + 1}{s^2}$ results in a quadratic equation:

$$\begin{aligned} \delta - 1 &= s^2\left[s^2 - A - \frac{1}{A}\right], \\ 0 &= A^2 - CA + 1. \end{aligned} \tag{3.2.26}$$

It is the quadratic equation in (3.2.26) with the combination of a second emerging quadratic equation we will use to obtain the value of λ.

Another quadratic equation emerges using the two equations in (3.2.22) and (3.2.23). First let $\alpha \equiv \frac{\lambda}{2\eta_1}[1 - \frac{s^2}{A}]$ and $\frac{1}{\alpha} = \frac{\lambda}{2\eta_2}[1 - As^2]$ the relationship between the two failure rates becomes:

$$\sqrt{q_1} = \alpha\sqrt{q_2}. \tag{3.2.27}$$

Using the relationship in (3.2.27) and the definition of A, q_i can be expressed explicitly in terms of the fixed constants: $\{\eta_1, \eta_2, s\}$ and the parameter λ which is yet to be determined. The derivation starts from the definition of A,

$$\begin{aligned} A^2 &= \frac{q_2\,(1-q_1)}{q_1\,(1-q_2)}, \\ \frac{q_1}{1-q_1}A^2 &= \frac{q_2}{1-q_2}, \\ \frac{\alpha^2 q_2}{1-\alpha^2 q_2}A^2 &= \frac{q_2}{1-q_2}. \end{aligned}$$

Solving for q_2 then using the relationship $q_1 = \alpha^2 q_2$ we get:

$$q_2 = \frac{1-\alpha^2 A^2}{\alpha^2\,(1-A^2)}, \tag{3.2.28}$$

$$q_1 = \frac{1-\alpha^2 A^2}{(1-A^2)}. \tag{3.2.29}$$

This is the expression of the individual failure rates which are yet to be optimized subject to the constraint. Now q_1 and q_2 is replaced in the constraint given in (3.2.20) and from there the optimal value of λ can be obtained.

Some prior calculations before replacing q_i into the constraint will simplify the overall algebra:

$$1-q_1 = \frac{A^2\,(\alpha^2-1)}{(1-A^2)},$$

$$1-q_2 = \frac{\alpha^2-1}{\alpha^2\,(1-A^2)}.$$

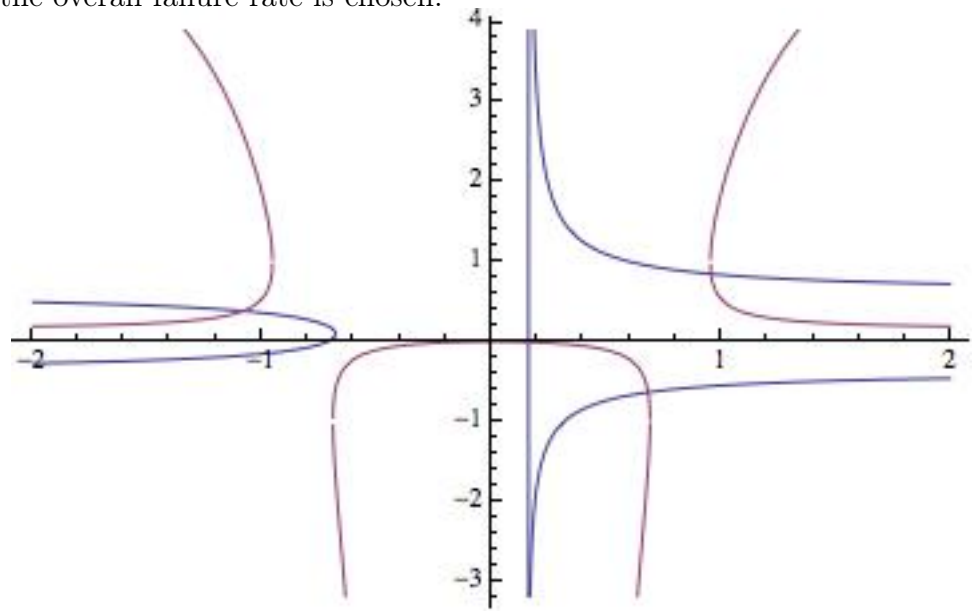
Now we are ready to replace q_i and $1-q_i$ into the constraint (3.2.20):

$$
\begin{aligned}
s &= \frac{A\left(\alpha^2-1\right)}{\alpha\left(1-A^2\right)}s^2+\frac{1-\alpha^2A^2}{\alpha\left(1-A^2\right)},\\
\left(1-A^2\right)s &= \frac{1}{\alpha}\left(1-s^2A^2\right)-\alpha A^2\left(1-\frac{s^2}{A}\right),\\
\left(1-A^2\right)s &= \frac{\lambda}{2\eta_2}\left(1-s^2A^2\right)^2-\frac{\lambda A^2}{2\eta_1}\left(1-\frac{s^2}{A}\right)^2.
\end{aligned}
$$

Here we replaced $\alpha{=}\frac{\lambda}{2\eta_1}\left[1-\frac{s^2}{A}\right]$ and $\frac{1}{\alpha}{=}\frac{\lambda}{2\eta_2}\left[1-As^2\right]$. After some trivial algebra the second quadratic equation in A emerges:

$$
A^2-\frac{2s^2(\eta_1-\eta_2)}{\eta_1 s^4-\eta_2+2\eta_1\eta_2 s/\lambda}A+\frac{\eta_1-\eta_2 s^4+2\eta_1\eta_2 s/\lambda}{\eta_1 s^4-\eta_2+2\eta_1\eta_2 s/\lambda}=0 \tag{3.2.30}
$$

The combination of Eq. (3.2.26) and Eq. (3.2.30) should give the value of λ which in turn gives explicit solution to the minimum individual failure rates, q_i, and overall optimal failure rate $Q=\eta_1 q_1+\eta_2 q_2$. Analytically such a solution is hard to achieve as one would have to solve a sixth order equation. Simply by plotting the two quadratic equations as a function of λ the solution can be obtained at the intersection of the two graphs. There will be multiple intersections, hence multiple values of λ. The one which gives the lowest value of the overall failure rate is chosen.



In our future work we would like to obtain a closed form solution of q_i. One approach could be by making some educated guess for λ.

3.3. Exact Cloning then Unambiguous Discrimination.

It is interesting to see the connection between cloning and state discrimination of non-orthogonal quantum states. To get a better understanding we consider the two step process where we first clone then make a measurement. The other way around, state discrimination first then cloning is less interesting because once we have succeeded in discriminating a state we can make as many copies as we wish by simply preparing them with the knowledge we get from the discrimination step. Making a clone first and then performing state discrimination gives some interesting results. For the case when the incoming states are prepared with equal priors we show that cloning first then performing UD on the cloned states which come with some new a priori probability p' the overall failure rate reaches the IDP limit.

The idea of the two step process is the following:

(1) Probabilistic Exact Cloning: Given an ensemble of M quantum states $\{|\psi_1\rangle, |\psi_2\rangle\}$ produce N exact clones while allowing for a rate of inconclusive outcomes.

(2) Unambiguous Discrimination: Perform optimal unambiguous discrimination on the successfully cloned states only, throwing away the failed states.

Equal priors.

(1) Probabilistic Exact Cloning

We showed in Section (3.2.1) that for equal priors and a set of M copies of two non-orthogonal states we could successfully produce exact copies with a probability $P_{clone} = \frac{1-s^M}{1-s^N}$. The average failure rate, failing to clone a state is $Q_{clone} = 1 - P_{clone} = \frac{s^M - s^N}{1-s^N}$.

(1) Optimal Unambiguous Discrimination

After the input states are sent through the deterministic exact cloning machines, N states of of the ensemble $\{|\psi_1\rangle, |\psi_2\rangle\}$ come out through the output port. The cloning machine only makes copies and does not say which state it has made a copy of. The clones are now sent

through an unambiguous discriminating machine to distinguish the incoming states. Each of the states comes with a probability $P_{clone} = \frac{1-s^M}{1-s^N}$.

The incoming states $\{|\psi_1^N\rangle_s, |\psi_2^N\rangle_s\}$ which live in the state Hilbert space S are embedded with the ancilla $|i\rangle_a$ which live in the ancilla Hilbert space A. Now the system and the ancilla live in the larger Hilbert space $H = S \otimes A$. The incoming states in this larger Hilbert space can be written in the product form $\{|\psi_1^N\rangle_s|i\rangle_a, |\psi_2^N\rangle_s|i\rangle_a\}$

$$\begin{aligned} U|\psi_1\rangle^{\oplus N}|0\rangle &= \sqrt{p}|\phi\rangle_s|1\rangle_a + \sqrt{q}|\Phi\rangle_s|0\rangle_a \\ U|\psi_1\rangle^{\oplus N}|0\rangle &= \sqrt{p}|\phi^\perp\rangle_s|2\rangle_a + \sqrt{q}|\Phi\rangle_s|0\rangle_a \end{aligned} \tag{3.3.1}$$

Here p is the probability of successfully discriminating the state $|\psi_i^N\rangle_s$, q is the probability of failing to discriminate $|\psi_i^N\rangle_s$, $p + q = 1$. The unitary operator takes the two incoming states and projects them onto a pair of orthogonal states with some success and some failure probability. When there is a click on the ancilla $|1\rangle_a$ the input states have been separated and output states $|\psi_i'\rangle_s$ are orthogonal and thus fully distinguishable. If there is a click along the ancilla $|0\rangle_a$ the incoming states have been collapsed into a single state which carries no information.

The inner product of the two equations in (3.3.1) produces the solution $s^N = q$. In addition one has to take into consideration the fact that the states come in with a priori probability of $\frac{1}{2}p = \frac{1}{2}\frac{1-s^M}{1-s^N}$. The optimal inconclusive rate for discriminating the incoming pair of states is$Q = \tilde{\eta}_1 q_1 + \tilde{\eta}_2 q_2 = \eta_1 p_1 q_1 + \eta_2 p_2 q_2 = \frac{1-s^M}{1-s^N} s^N$, where $q_1 = q_2 = s^N$ and success rate is $p_1 = p_2 = \frac{s^M - s^N}{1-s^N}$.

The total inconclusive rate for the two step process, the failure rate to clone plus failure rate to unambiguously discriminate, is

$$\begin{aligned} Q_{total} &= Q_{clone} + (1 - Q_{clone})\, Q_{UD}, \\ &= \frac{s^M - s^N}{1 - s^N} + \frac{1 - s^M}{1 - s^N} s^N = s^M. \end{aligned} \tag{3.3.2}$$

The two step process reaches the IDP limit for optimal UD of M non-orthogonal states. This is a special case and only occurs for symmetric case when states are prepared with equal prior probabilities.

3.4. State Separation

In Exact Cloning one prepares perfect clones of the input states while allowing for some failure rate in which case no clones have been produced and the states are discarded. In optimal unambiguous discrimination (UD) the input states are made orthogonal and hence fully distinguishable. It was shown by Chefles and Barnett **[25]** that these strategies are a special case of a more general scheme. Both have two outcomes: failure and success. In each strategy the overlap of the input states is decreased, in UD the overlap becomes zero, in exact cloning it is the overlap of the input states raised to the power of the desired number of the clones to be made, s^N. State separation unifies the two schemes as it produces states with an overlap s' in the range $0 \leq s' \leq s^N$ while allowing for a fixed rate of inconclusive results. The authors showed the results for the case when the states are prepared with equal a priori probabilities. Complimentary of our recent work on probabilistic exact cloning **[26]** where a geometric picture emerges we use similar tools to solve the more general state separation when the input states are prepared with different a priori probabilities.

We approach state separation via the Neumark formulation. The system is embedded in a larger Hilbert space where the extra degrees of freedom are customarily called the ancilla. Then a unitary transformation entangles the system degrees of freedom with those of the ancilla. The input states $\{|\psi_1\rangle_s, |\psi_2\rangle_s\}$ which live in the state Hilbert space H_S are embedded with the ancilla $|i\rangle_a$ which live in the Hilbert space H_A. Now the system and the ancilla live

in the larger Hilbert space $H = H_S \otimes H_A$. The incoming states in this larger Hilbert space can be written in the product form $\{|\psi_1\rangle_s|i\rangle_a, |\psi_2\rangle_s|i\rangle_a\}$. The unitary should do the following:

$$\begin{aligned} U|\psi_1\rangle|i\rangle &= \sqrt{p_1}|\phi_1\rangle|\alpha\rangle + \sqrt{q_1}|\Phi_o\rangle|f\rangle, \\ U|\psi_2\rangle|i\rangle &= \sqrt{p_2}|\phi_2\rangle|\alpha\rangle + \sqrt{q_2}|\Phi_o\rangle|f\rangle, \end{aligned} \tag{3.4.1}$$

where $|\alpha\rangle$ and $|f\rangle$ are orthogonal. A projective measurement along the ancilla $|\alpha\rangle$ means that the states have successfully become more distinguishable with a success rate of p_i, otherwise a measurement along the $|f\rangle$ space means that the process has failed to produce more distinguishable states and the states are discarded with a probability of q_i. The separation of the input states is shown in Figure. 3.4.1.

The inner product of the two equations in 3.4.1 gives the unitarity constraint:

$$s = \sqrt{p_1 p_2}s' + \sqrt{q_1 q_2}, \tag{3.4.2}$$

where $s = |\langle\psi_1 \mid \psi_2\rangle|^M$ and $s' = |\langle\phi_1 \mid \phi_2\rangle|^N$.

For given η_1, η_2 and average failure probability Q, we wish to find out the minimum value of the final overlap s' as a function of the initial overlap$s$. We could also look at the problem from a different angle: for a fixed value of the overlap of the output states what is the minimum value of failure rate Q to achieve the desired separation.

3.4.1. State Separation: equal priors. Chefles and Barnett **[25]** solve the problem of equal priors employing Kraus representation of quantum maps $\left\{\hat{A}, \hat{A}^\dagger\right\}$. However the solution is more straightforward using the Neumark setup in (3.4.1). When the input states are prepared with equal priors $\eta_1 = \eta_2$, the solution is directly derived from the constraint in (3.4.2) and no further optimization is necessary:

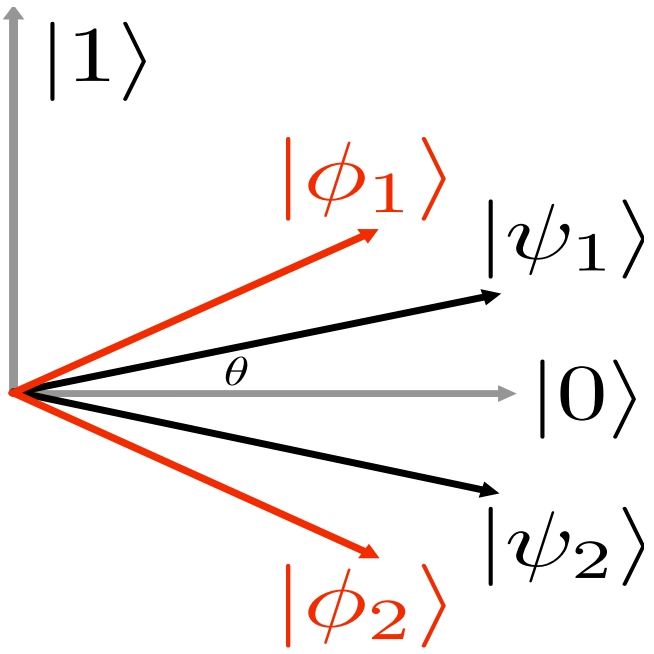


FIGURE 3.4.1. The input states $\{|\psi_1\rangle, |\psi_2\rangle\}$ are separated into a pair of more orthogonal states $\{|\phi_1\rangle, |\phi_2\rangle\}$ which are more distinguishable.

$$\begin{aligned} s &= ps' + q, \\ p &= \frac{1-s}{1-s'}. \end{aligned} \tag{3.4.3}$$

We will now show applications of state separation to exact cloning and unambiguous state discrimination. Another application to state separation for equal priors is in hybrid cloning, see Section (3.6.1). Let us begin by showing the connection to exact cloning with failure rate. Suppose we are given a quantum system in either of the pure states $\{|\psi_1\rangle, |\psi_2\rangle\}$, such that the two states are non-orthogonal. The task is to produce two perfect copies of the input state. Let us embed the quantum system with a blank state $|i\rangle$ so the system becomes $\{|\psi_1\rangle |i\rangle, |\psi_2\rangle |i\rangle\}$. A perfect copying machine should produce $\{|\psi_1\rangle |\psi_1\rangle, |\psi_2\rangle |\psi_2\rangle\}$. A modified Neumark setup from state separation in (3.4.1) for cloning is:

$$\begin{aligned} U|\psi_1\rangle|i\rangle &= \sqrt{p}\,|\psi_1\rangle\,|\psi_1\rangle\,|\alpha\rangle + \sqrt{q}|\Phi_o\rangle|f\rangle, \\ U|\psi_2\rangle|i\rangle &= \sqrt{p}\,|\psi_1\rangle\,|\psi_1\rangle\,|\alpha\rangle + \sqrt{q}|\Phi_o\rangle|f\rangle, \end{aligned} \tag{3.4.4}$$

The inner product gives the unitarity constraint which is enough to fully solve the problem:

$$\begin{aligned} s &= ps^2 + q \\ p &= \frac{1-s}{1-s^2} = \frac{1}{1+s} \end{aligned} \tag{3.4.5}$$

We notice that the success probability of producing two perfect clones with a rate of abstention (3.4.5) could be derived from state separation success rate in Eq. (3.4.3) by replacing the overlap of output states s' with the square overlap of input states s^2. This shows that the Duan-Guo limit **[23]** is a special case of state separation.

More generally, state separation can be modified to produce N copies of $\{|\psi_1\rangle, |\psi_2\rangle\}$ from M initial copies, where $N \geq M$. We derived the solution of M to N cloning in section (3.2.1). It could simply be derived from Eq. (3.4.3) replacing the overlap of input states s with s^M, the overlap of M states, and s' by s^N, the overlap of the clones.

State separation also reproduces the IDP limit in state discrimination for the case when the states are prepared with arbitrary priors. Simply setting the overlap of the output states to zero $s' = 0$, so that the states have been separated and are fully distinguishable, produces the IDP limit $p = 1 - s$.

3.4.2. State Separation: unequal priors. We now seek to generalize the results of state separation to the case when the possible input states $\{|\psi_1\rangle, |\psi_2\rangle\}$ are prepared with different a priori probabilities. All the symmetries enjoyed for equal priors break down and in order to obtain a fully analytical solution one would have to solve a sixth order equation.

Instead we resort to solving the problem geometrically as the solution turns out to be more insightful. The Neumark setup is shown in (3.4.1) and the constraint in (3.4.2).

We resort to the parametrization of the probabilities giving the curve $(s, \min s')$ in parametric form, rather than attempting to give min s' as an explicit function of s, which would require solving a high degree polynomial equation. We choose a change of variables which linearizes the unitarity constraint (3.4.2):

$$p_1 p_1 = t^2, \quad q_1 q_2 = z^2. \tag{3.4.6}$$

The condition becomes

$$z = s - s't, \qquad 0 \leq t, z \leq 1, \qquad 0 \leq s' \leq s. \tag{3.4.7}$$

It is a straight line in the first quadrant of the plane t-z with negative slope $-s'$ and crossing the z-axis at $z = s$. From the first equation in (3.4.6) we have

$$t^2 = (1 - q_1)(1 - q_2) = 1 + z^2 - q_1 - q_2.$$

Solving for q_2 and substituting back in the second equation in (3.4.6) we have

$$q_1(q_1 - 1 + t^2 - z^2) + z^2 = 0.$$

We now solve for q_1 and obtain

$$q_1 = \frac{1 + z^2 - t^2 \pm \sqrt{(1 + z^2 - t^2)^2 - 4z^2}}{2}.$$

Similarly, for q_2 we obtain

$$q_2 = \frac{1 + z^2 - t^2 \mp \sqrt{(1 + z^2 - t^2)^2 - 4z^2}}{2}.$$

Therefore, the condition $Q = \eta_1 q_1 + \eta_2 q_2$ becomes

$$2Q = 1 + z^2 - t^2 \pm (\eta_1 - \eta_2)\sqrt{(1 + z^2 - t^2)^2 - 4z^2}.$$

We now solve for z^2. After a bit of algebra we obtain

$$\begin{aligned} z^2 &= \frac{2\eta_1\eta_2(1+\tau) - 1 + Q + \sqrt{(1-4\eta_1\eta_2)\left[(1-Q)^2 - 4\eta_1\eta_2\tau\right]}}{2\eta_1\eta_2} \\ &\equiv \zeta(\tau) \end{aligned}$$

where $\tau \equiv t^2$. Since z^2 cannot be less than zero, we picked up the plus sign for the root. Let us assume that $0 \leq \eta_1 \leq 1/2$ to simplify the analysis. We need to locate the maximum of z. For that,

$$\frac{dz}{dt} = \frac{d\sqrt{\zeta}}{d\tau} 2t = \frac{d\zeta}{d\tau}\frac{t}{z}.$$

The derivative $d\zeta/d\tau$ is immediate. We find that the maximum is located at

$$t_{\min} = \begin{cases} \sqrt{\left(1 - \dfrac{Q}{2\eta_1}\right)\left(1 - \dfrac{Q}{2\eta_2}\right)}, & \text{if} \quad 0 \leq Q \leq 2\eta_1 \\ \\ 0, & \text{if} \quad 2\eta_1 < Q \leq 1. \end{cases}$$

The corresponding values of z are

$$z_{\min} = \begin{cases} \dfrac{Q}{2\sqrt{\eta_1\eta_2}}, & \text{if} \quad 0 \leq Q \leq 2\eta_1 \\ \\ \sqrt{\dfrac{Q-\eta_1}{\eta_2}}, & \text{if} \quad 2\eta_1 < Q \leq 1. \end{cases}$$

There is a second point that we need to define. We first note that for equal priors the curve is simply the hyperbola

$$z^2 = t^2 + 2Q - 1,$$

which intersects the straight line

$$z = 1 - t$$

at the point

$$(z, t) = (Q, 1 - Q)\,.$$

By trying this solution in $z^2 = \zeta(t^2)$ we note that it is actually a general solution for any η_1, η_2. Moreover, the straight line $z = 1 - t$ is tangent to $z^2 = \zeta(t^2)$ at $(Q, 1 - Q)$ for any values of η_1, η_2, as can be checked by substituting in the formula $dz/dt = (t/z)(d\zeta/d\tau)$. Note also that $z = 1 - t$ is the limiting line for the family $z = s - s't$. Hence, an obvious parametrization for the curve (s, s') is obtained as follows: *i.* define

$$s'(t) = -\frac{dz}{dt} = -\frac{t\,\zeta'(t^2)}{\sqrt{\zeta(t^2)}}, \qquad t_{\min} \leq t \leq 1 - Q,$$

and next *ii.* define

$$s(t) = z + ts'(t) = \sqrt{\zeta(t^2)} + ts'(t), \qquad t_{\min} \leq t \leq 1 - Q,$$

where

$$\zeta'(\tau) = 1 - \frac{\sqrt{1 - 4\eta_1\eta_2}}{\sqrt{(1-Q)^2 - 4\eta_1\eta_2\tau}}.$$

For $s < z_{\min}$ it is always possible to separate the initial states, i.e., $|\psi_1\rangle$ and $|\psi_2\rangle$ can be made orthogonal. We note that the condition $s = z_{\min}$ is equivalent to the unambiguous discrimination result

$$Q = 2\sqrt{\eta_1\eta_2}s, \quad Q = \eta_1 + \eta_2 s^2.$$

3.5. Deterministic State Dependent Quantum Cloning

In this section we derive the works of Chefles and Barnett **[27]** in designing an approximate quantum cloning machine for two possible input states while maximizing the global fidelity. Consider a set of K non-orthogonal quantum states with M copies each $|\psi_i\rangle^{\otimes M} = |\psi_i\rangle|\psi_i\rangle...|\psi_i\rangle$. The states are unknown and our task is to produce $N > M$ copies, as best as we can. Introducing an ancilla state $|\chi\rangle$ which is $N - M$ dimensional the goal is to transform the state $|\psi_j\rangle|\chi\rangle$ into the state which approximates the N exact copies of the input state $|\psi_j^N\rangle$.

This is deterministic cloning, although imperfect, clones are generated on demand. The authors choose the global fidelity rate to improve the quality of the clones so they resemble the given copies as closely as possible. This measure was introduced by Bruss et al **[28]**.

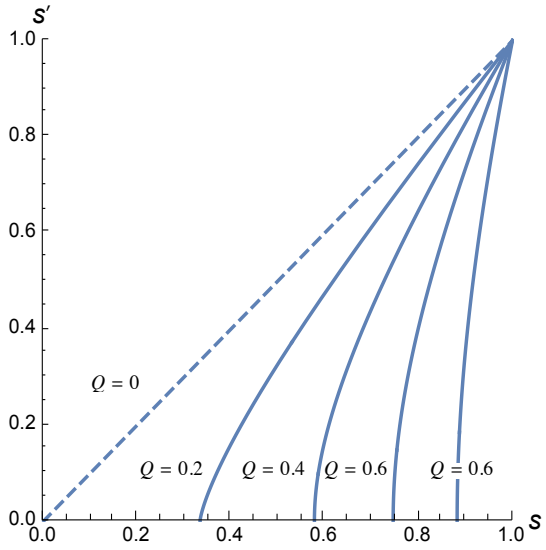


FIGURE 3.4.2. The plot is for $\eta_1 = 0.1$. As η_1 approaches $1/2$ the curves approach a straight line. The difference is more noticeable for very small values of η_1.

Thus given a set K of non orthogonal states $\left|\phi_j^N\right\rangle = |\psi_i\rangle^{\otimes M}$, we wish to produce a set K of N clones $|\phi_j\rangle^{\otimes N}$ while optimizing the global fidelity:

$$F_{MN} = \sum_{j=1}^{K} \eta_j |\langle \psi_j^N | \phi_j^N \rangle|^2 \tag{3.5.1}$$

Other figures of merit can also be used to improve the quality of the clones, such as local fidelity. The local fidelity is the average fidelity of each of the individual clones of each of the N subsystems measured against the input states $|\psi_j\rangle$. The authors choose the global fidelity due to its close connection to state discrimination. We are also very interested in the connection between cloning and state discrimination, particularly in the two step process where clones are first produced then the clones are unambiguously discriminated.

The global fidelity can be expressed differently if unitary operator acts on the input states $U|\psi_j^M\rangle|\chi\rangle = |\phi_j^N\rangle$.

$$F_{MN} = \sum_{j=1}^{K} \eta_j |\langle \psi_j^N | U | \psi_j^M \rangle | \chi \rangle|^2 \tag{3.5.2}$$

The problem of maximizing the fidelity can be explicitly solved for a set of two possible input states $K = 2$, $\{|\psi_1\rangle, |\psi_2\rangle\}$. It was originally solved by Bruss *et al* **[28]** for the case when the the incoming states are prepared with equal a priori probabilities. They noticed that the optimum clones $\{|\Phi_1\rangle, |\Phi_2\rangle\}$ lie in the subspace spanned by the input states to be cloned $\{|\psi_1\rangle, |\psi_2\rangle\}$.

A unitary produces N copies $|\phi_1^N\rangle$ or $|\phi_2^N\rangle$, to resemble the original states as best as possible.

$$U|\psi_1^M\rangle|i\rangle = |\phi_1^N\rangle \tag{3.5.3}$$

$$U|\psi_2^M\rangle|i\rangle = |\phi_2^N\rangle \tag{3.5.4}$$

The inner product of the above two equations gives a relationship between the input and the output states

$$|\langle\psi_1|\psi_2\rangle|^M = |\langle\phi_1|\phi_2\rangle|^N, \tag{3.5.5}$$

$$s^M = s'^N, \tag{3.5.6}$$

where s^M=$|\langle\psi_1|\psi_2\rangle|^M$ and $s'^N = |\langle\phi_1|\phi_2\rangle|$.

The input states can be expressed as:

$$\begin{aligned} |\psi_1\rangle &= \cos\theta|1\rangle + \sin\theta|0\rangle, \\ |\psi_2\rangle &= \cos\theta|1\rangle - \sin\theta|0\rangle, \end{aligned} \tag{3.5.7}$$

similarly the clones yet to be optimized can be expressed as:

$$
\begin{aligned}
|\phi_1\rangle &= \cos\phi_1|1\rangle + \sin\phi_1|0\rangle, \\
|\phi_2\rangle &= \cos\phi_2|1\rangle - \sin\phi_2|0\rangle. \qquad (3.5.8)
\end{aligned}
$$

Using this general representation of input and output states and using the overlap relation in 3.5.6 we see that the sum of the output angles is fixed as $|\langle\psi_1|\psi_2\rangle|^M = \cos^M 2\theta$ and $|\langle\phi_1|\phi_2\rangle|^N = \cos^N(\phi_1+\phi_2) \Rightarrow \cos^M 2\theta = \cos^N(\phi_1+\phi_2)$.

The global fidelity in terms of the angles becomes:

$$
\begin{aligned}
F_{MN} &= \eta_1 |\,|\langle\psi_1|\phi_1\rangle| + \eta_2\,|\langle\psi_2|\phi_2\rangle|^2\,, \qquad (3.5.9) \\
&= \eta_1\,(\cos\theta\cos\phi_1 + \sin\theta\sin\phi_1) + \eta_2\,(\cos\theta\cos\phi_2 + \sin\theta\sin\phi_2)\,, \\
&= \eta_1\cos^2(\theta-\phi_1) + \eta_2\cos^2(\theta-\phi_2)\,. \qquad (3.5.10)
\end{aligned}
$$

Rewriting the fidelity in terms of the sum and difference of the output angles,

$$
\begin{aligned}
F_{MN} &= \eta_1\cos^2\left(\theta - \frac{\phi_1+\phi_2}{2} - \frac{\phi_1-\phi_2}{2}\right) + \eta_2\cos^2\left(\theta - \frac{\phi_1+\phi_2}{2} + \frac{\phi_1-\phi_2}{2}\right), \\
&= \eta_1\cos^2(\alpha - x) + \eta_2\cos^2(\alpha + x)\,, \\
&= \frac{1}{2}\left[\eta_1\,(\cos 2(\alpha-x)+1)\right] + \frac{1}{2}[\eta_2\,(\cos 2(\alpha+x)+1)\Big]\,, \\
&= \frac{1}{2}\,[1 + \eta_1(\cos 2(\alpha-x) + \eta_2(\cos 2(\alpha+x)]\,, \qquad (3.5.11)
\end{aligned}
$$

where $\alpha = \theta - \frac{\phi_1+\phi_2}{2}$ is fixed and $x = \frac{\phi_1-\phi_2}{2}$ is the only variable subject to optimization. Differentiating with respect to x we get

$$\begin{aligned}
\eta_1 \sin(\alpha - x) &= \eta_2 \sin(\alpha - x) \\
\eta_1 \left[\sin 2\alpha \cos 2x - \cos 2\alpha \cos 2x\right] &= \eta_2 \left[\sin 2\alpha \cos 2x + \cos 2\alpha \cos 2x\right] \\
(\eta_1 - \eta_2) \sin 2\alpha \cos 2x &= \cos 2\alpha \cos 2x \\
(\eta_1 - \eta_2) \tan 2\alpha &= \tan 2x \qquad (3.5.12)
\end{aligned}$$

This is the relationship that gives the optimal clones and x should be replaced in F_{MN}. To be able to use this relationship we re-express F_{MN} in a different way:

$$\begin{aligned}
F_{MN} &= \frac{1}{2}\left[1 + \eta_1 \cos 2(\alpha - x) + \eta_2 \cos 2(\alpha + x)\right], \\
&= \frac{1}{2}\left[\cos 2\alpha \cos 2x \left\{1 + (\eta_1 - \eta_2) \tan 2\alpha \tan 2x\right\}\right] + \frac{1}{2},
\end{aligned}$$

we can rewrite (3.5.12) as $(\eta_1 - \eta_2)\tan 2\alpha = \frac{\sin 2x}{\cos 2x}\cos 2x = \frac{\sqrt{1-\cos^2 2x}}{(\eta_1-\eta_2)\tan 2\alpha} = \frac{1}{\sqrt{1+(\eta_1-\eta_2)^2 \tan^2(2\alpha)}}$. Finally the optimal fidelity can be expressed in terms of the a-priori probabilities and the overlap of the input states only .

$$\begin{aligned}
F_{MN} &= \frac{1}{2}\left[1 + \cos 2\alpha \cos 2x \left\{1 + (\eta_1 - \eta_2) \tan 2\alpha \tan 2x\right\}\right], \\
&= \frac{1}{2}\left[1 + \frac{\cos 2\alpha}{\sqrt{1 + (\eta_1 - \eta_2)^2 \tan^2(2\alpha)}}\{1 + (\eta_1 - \eta_2)^2 \tan^2 2\alpha\right], \\
&= \frac{1}{2}\left[1 + \cos 2\alpha \sqrt{1 + (\eta_1 - \eta_2)^2 \tan^2 2\alpha}\right], \\
&= \frac{1}{2}\left[1 + \sqrt{\cos^2 2\alpha + (1 - 4\eta_1\eta_2)\sin^2 2\alpha}\right], \\
&= \frac{1}{2}\left[1 + \sqrt{1 - 4\eta_1\eta_2 \sin^2 2\alpha}\right], \\
&= \frac{1}{2}\left[1 + \sqrt{1 - 4\eta_1\eta_2 \sin^2 (2\theta - (\phi_1 + \phi_2))}\right]. \qquad (3.5.13)
\end{aligned}$$

In the asymptotic limit, producing infinitely many copies $N \to \infty$, fidelity merges into the Helstrom bound in the discrimination of non-orthogonal pure states. Using $\cos 2\theta = s^M$, $\cos(\phi_1 + \phi_2) = s^N$ and expanding the sin term under the square root,

$$\begin{aligned}
\sin\left(2\theta - (\phi_1 + \phi_2)\right) &= \sin 2\theta \cos\left(\phi_1 + \phi_2\right) - \cos 2\theta \sin\left(\phi_1 + \phi_2\right) \\
&= \sqrt{1 - \cos^2(2\theta)} \cos\left(\phi_1 + \phi_2\right) - \cos 2\theta \sin\left(\phi_1 + \phi_2\right) \\
&= \sqrt{1 - s^{2M}} s^N - s^M \sqrt{1 - s^{2N}} \\
&= -s^M
\end{aligned}$$

Substituting back into Eq. (3.5.13) Helstrom bound emerges:

$$F_{M\infty} = \frac{1}{2}[1 + \sqrt{1 - 4\eta_1\eta_2 \cos^2(2\theta)}] \tag{3.5.14}$$

While the mathematics shows the convergence of Fidelity into optimal minimum error state discrimination it is not clear as to why this connection should exist at all. On the one hand fidelity optimizes the average overlap between the clones and input states while ME state discrimination minimizes the error rate of failing to distinguish the incoming state.

We will show the connection in the following two step process: first measure and prepare then optimize the fidelity rate.

- Step 1: Measure and prepare

Discriminate the incoming states $\left\{\left|\psi_1^M\right\rangle, \left|\psi_2^M\right\rangle\right\}$ with optimal ME then prepare states $\left\{\left|\psi_1^N\right\rangle, \left|\psi_2^N\right\rangle\right\}$ with the corresponding probabilities.

$$U|\psi_1^M\rangle|0\rangle = \sqrt{p_1}|\psi_1^N\rangle|1\rangle + \sqrt{r_1}|\psi_2^N\rangle|2\rangle, \tag{3.5.15}$$

$$U|\psi_2^M\rangle|0\rangle = \sqrt{r_2}|\psi_1^N\rangle|1\rangle + \sqrt{p_2}|\psi_2^N\rangle|2\rangle, \tag{3.5.16}$$

When state $\left|\psi_i^M\right\rangle$ is received we successfully prepare the state $\left|\psi_i^N\right\rangle$ with probability p_i and mistakenly prepare the state $\left|\psi_j^N\right\rangle$ with probability r_i, $i = 1, 2$.

- Step 2: Optimize the fidelity

The global fidelity for state $\left|\psi_1^N\right\rangle$ is

$$\begin{aligned} F_1 &= p_1 |\langle\psi_1 \mid \psi_1\rangle|^N + r_1 |\langle\psi_2 \mid \psi_1\rangle|^{2N}, \\ &= p_1 + r_1 |\langle\psi_2 \mid \psi_1\rangle|^{2N}. \end{aligned}$$

Similarly for state $\left|\psi_2^N\right\rangle$

$$F_2 = p_2 + r_2 |\langle\psi_1 \mid \psi_2\rangle|^{2N}.$$

The average fidelity for both states is:

$$\begin{aligned} F_{MN} &= \eta_1 F_1 + \eta_2 F_2 \\ &= \eta_1 \left(p_1 + r_1 |\langle\psi_2 \mid \psi_1\rangle|^{2N}\right) + \eta_2 \left(p_2 + r_2 |\langle\psi_1 \mid \psi_2\rangle|^{2N}\right) \\ &= \eta_1 p_1 + \eta_2 p_2 + \eta_1 r_1 s^{2N} + \eta_2 r_2 s^{2N} \end{aligned} \tag{3.5.17}$$

It is now clear that in the asymptotic limit $N \to \infty$ fidelity reproduces the Helstrom bound $F_{MN} = \eta_1 p_1 + \eta_2 p_2$. However the relationship between optimal fidelity and optimal state discrimination for a finite number of copies N remains an open question.

3.6. Hybrid Cloning: Interpolation between exact and approximate cloning

In this section we seek to interpolate between probabilistic exact cloning and approximate cloning machines using our results from state separation. Exact cloning machines produce perfect clones while allowing for some inconclusive outcomes. Approximate cloning machines produce copies on demand which resemble the input states by maximizing the fidelity. One can imagine a scheme where fidelity can be higher then maximum fidelity in the approximate

cloning machine while it allows for a fixed rate of inconclusive outcomes, FRIO. This scheme should reproduce exact cloning and approximate cloning machines by setting FRIO to Q_o and zero respectively. Chefles and Barnett [**27**] solved the problem for the case when the input states are prepared with equal a priori probabilities. We extend the solution to the more general case when the states are prepared with different priors. Such a solution is possible due to our recent work on making N perfect clones from M copies of one of two known pure states with minimum failure probability in the general case where the known states have arbitrary priori probabilities.

3.6.1. Equal priors. The solution to the interpolation of cloning for equal a priori probabilities has been derived by Chefles et al [**27**]. The authors develop a scheme which, depending on the fidelity of the clones, can interpolate between exact cloning with inconclusive results in one extreme and optimal approximate cloning on the other extreme. In our work this scheme has been generalized for the case when the input states are prepared with different a priori probabilities. First we show the derivation of the equal priors as it will help to better understand the general case.

For $\eta_1 = \eta_2 = 1/2$, the output states are symmetric, $\phi_1 = \phi_2 = \phi$, and the optimal global fidelity, F_{MN}, in Eq.(3.5.13) reduces to:

$$\begin{aligned} F_{MN} &= \frac{1}{2}\left[1 + \sqrt{1 - \sin^2(2\theta - 2\phi)}\right], \\ &= \frac{1}{2}\left[1 + \cos^2(2\theta - 2\phi)\right]. \end{aligned} \tag{3.6.1}$$

Duan and Guo [**23**] 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:

$$P_{MN} = \frac{1 - s^M}{1 - s^N}, \tag{3.6.2}$$

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:

$$P_{12} = \frac{1}{1+s}. \tag{3.6.3}$$

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:

$$U|\psi_1\rangle^{\otimes M}|i\rangle = \sqrt{p}|\phi_1\rangle^{\oplus N}|1\rangle + \sqrt{q}|f\rangle|0\rangle \tag{3.6.4}$$

$$U|\psi_2\rangle^{\oplus M}|i\rangle = \sqrt{p}|\phi_2\rangle^{\oplus N}|1\rangle + \sqrt{q}|f\rangle|0\rangle \tag{3.6.5}$$

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 (3.6.4) and (3.6.5) gives the constraint:

$$s^M = ps'^N + q \tag{3.6.6}$$

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

$$p = \frac{1 - s^M}{1 - s'^N} \tag{3.6.7}$$

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

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

$$|\langle\phi_1|\phi_2\rangle|^N = 1 - \frac{1 - |\langle\psi_1|\psi_2\rangle|^M}{P_S} \tag{3.6.8}$$

$$\cos^N(\phi_1 + \phi_2) = 1 - \frac{P_{IDP}}{P_S} \tag{3.6.9}$$

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 $|\phi_{1,2}\rangle = \cos\phi_1|1\rangle \pm \sin\phi_1|0\rangle$.

The fidelity rate for equal priors is:

$$F_{MN} = \frac{1}{2}\left[1 + \cos(2\theta - (\phi_1 + \phi_2)\right], \tag{3.6.10}$$

and we want to use the relationship in (3.6.9). Let us expand the cosine term

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

The fidelity becomes:

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

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

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

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

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

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

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

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

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

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

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

3.6.2. 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.

$$\begin{aligned} U|\psi_1^M\rangle|0\rangle &= \sqrt{p_1}|\phi_1\rangle|1\rangle + \sqrt{q_1}|\Phi\rangle|2\rangle, \\ U|\psi_2^M\rangle|0\rangle &= \sqrt{p_2}|\phi_1\rangle|1\rangle + \sqrt{q_2}|\Phi\rangle|2\rangle, \end{aligned} \tag{3.6.11}$$

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

$$s = \sqrt{p_1 p_2}s' + \sqrt{q_1 q_2} \tag{3.6.12}$$

- Step 2: Optimize Fidelity

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

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

where the normalized a priori probabilities are $\tilde{\eta}_i = \frac{\eta_i p_i}{1-Q}$. The average fidelity is the same as calculated in (3.5.13) with the new normalized probabilities:

$$\begin{aligned} F_{MN} &= \frac{1}{2}\left[1 + \sqrt{1 - 4\tilde{\eta}_1\tilde{\eta}_2 \sin^2\left(2\theta - (\phi_1 + \phi_2)\right)}\right], \\ &= \frac{1}{2(1-Q)}\left[(1-Q) + \sqrt{(1-Q)^2 - 4\eta_1\eta_2 p_1 p_2 \sin^2\left(2\theta - (\phi_1 + \phi_2)\right)}\right] \end{aligned} \tag{3.6.13}$$

It can be seen that in the limit $N \to \infty$, expanding the sin term as we did in the previous section, the FRIO **[14]** results are recovered. It again shows a close relationship between fidelity and state discrimination.

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

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

Here $u \equiv \sqrt{q_1 q_2}$, $\quad v \equiv \frac{1}{2}\left(q_1 + q_2\right)$ and we used the constraint from the unitarity in (3.4.2) to replace $\sqrt{p_1 p_2}s' = s - \sqrt{q_1 q_2}$.

CHAPTER 4

Experimental realization to FRIO

Choosing a physical system to realize quantum information processes, which have otherwise been solved theoretically, is central challenge to building a quantum computer. Some of the systems in use today are: energy levels of ions, the orientation of nuclear spin, the presence or absence of a photon in a cavity [**29, 30, 31, 32**] and dual rail representation of a qubit proposed by Milburn [**33**] and later by Chuang and Yamamoto [**34**]. We will realize the implementation of our works using the dual rail representation of photons combined with a six-port, which is a linear device with three input and three output ports. The six-port can be realized with beamsplitters and phase shifters. First we will demonstrate the power and simplicity of this system by working out the implementation of UD following the work by Bergou *et al*. [**35**].

In our work [**36**] we generalized the optical implementation schemes to state discrimination which optimally interpolates between UD and ME with a fixed rate of inconclusive results FRIO. Ever since the interpolation scheme of a general measurement with FRIO was derived by Bagan *et al* [**14**] there has been a quest for a physical realization. The authors solved the problem using an operator transformation technique that reformulated the intermediate problem into a ME problem with an extra optimization parameter. Essentially they reduced the problem from a three element POVM to two element POVM similar to ME. Inspired by the work of Reck and Zeilinger [**37**] in which they prove that any discreet finite dimensional unitary operator can be constructed in the lab as a multi-port interferometer using beamsplitters and phase shifters we set out to solve the FRIO using the Neumark setup as it lends itself into an optical implementation. This gives us a closed form solution as in Bagan *et al*. In addition it gives a three dimensional unitary operator where all the coefficients are explicitly calculated in terms of a priori probabilities, overlap of the input

states and FRIO. Using the Reck *et al*. algorithm the unitary is decomposed in terms of beam splitters with corresponding coefficients of transmittance and reflectivity. The setup reduces into the UD by setting the error rate to zero, in turn it reproduces the work of Bergou *et al* **[35]**. At the other extreme it produces the setup to ME for FRIO equal to zero.

4.1. Analytical Solution of Interpolation

In this section we solve the interpolation with FRIO in two different ways. In the first method, through a parametrization, the problem is converted into a minimum error whose solution is well known. Then there is one last optimization with respect to a FRIO. This method generates the same solution as in Bagan *et al* in a few lines. However to obtain an experimental realization we give an alternative solution using Neumark's theorem. It is a generalized measurement procedure in which the system is embedded in a larger Hilbert space with extra degrees of freedom. A unitary transformation entangles the system with the extra degree of freedom known as ancilla. After this interaction has taken place, projective von Neumann measurements are carried out on the ancilla. Our input states are qubits, which can be expressed in general as unit length vectors in the two dimensional basis spanned by $|1\rangle$ and $|2\rangle$. In the output of the transformation, we associate the basis state $|1\rangle$ with $|\psi_1\rangle$, $|2\rangle$ with $|\psi_2\rangle$, and $|0\rangle$ gives no information about the system. The unitary transformation should do the following:

$$
\begin{aligned}
U|\psi_1\rangle_s|0\rangle_a &= \sqrt{p_1}|1\rangle_s\,|1\rangle_a + \sqrt{r_1}|2\rangle_s\,|2\rangle_a + \sqrt{q_1}|0\rangle_s\,|0\rangle_a\,,\\
U|\psi_2\rangle_s|0\rangle_a &= \sqrt{r_2}|1\rangle_s\,|1\rangle_a + \sqrt{p_2}|2\rangle_s\,|2\rangle_a + \sqrt{q_2}|0\rangle_s\,|0\rangle_a\,,
\end{aligned}
\tag{4.1.1}
$$

where: p_i is the probability that state $i\,(i=1,2)$ is correctly identified, r_i is the probability that the detector mistakenly identifies state i for j, and q_i is the failure probability, the detector fails to identify the state at all. A click in the ancilla $|0\rangle_a$ means the results are

inconclusive and we learn nothing from the measurement. From the unitarity conditions we obtain the normalized probabilities $p_i + r_i + q_i = 1$.

We wish to maximize the probability of success, $P_s = \eta_1 p_1 + \eta_2 p_2$, and minimize the error rate, $P_e = \eta_1 r_1 + \eta_2 r_2$, for a fixed failure rate $Q = \eta_1 q_1 + \eta_2 q_2$. Clearly $P_s + P_e + Q = 1$.

The inner product of the two equations in (4.1.1) gives the overlap of the input states in terms of r_i, p_i and q_i,

$$s = \sqrt{p_1 r_2} + \sqrt{p_2 r_1} + \sqrt{q_1 q_2}, \tag{4.1.2}$$

where $s \equiv \langle \psi_1 | \psi_2 \rangle$. This is a constraint on the optimization of ME with a FRIO. First we show the solution for the case when the input states are prepared with equal priors for which the problem is fully solved from the constraint. It was initially solved by Chefles and Barnett using a different approach [**15**]. Then we solve the general case for different priors and reproduce the results of Bagan *et al*. with two different methods.

4.1.1. Equal priors. Let us first present the solution where the incoming states are given with equal a-priori probabilities, $\eta_1 = \eta_2 = \frac{1}{2}$. This implies equal error, success and failure rates: $r_1 = r_2$, $p_1 = p_2$ and $q_1 = q_2$. Thus the total error and failure rates reduce to: $P_E = \eta_1 r_1 + \eta_2 r_2 = r$ and $Q = \eta_1 q_1 + \eta_2 q_2 = q$.

We can immediately solve the constraint in Equation (4.1.2) by replacing $p = 1 - r - Q$, solving the quadratic equation for the error rate in terms of the failure rate and overlap s, which is also the overall failure rate in the IDP limit for the equal priors: $Q_o \equiv 2\sqrt{\eta_1 \eta_2} s = s$,

$$\begin{aligned} s &= \sqrt{pr} + \sqrt{pr} + Q, \\ Q_o &= 2\sqrt{r(1-r-Q)} + Q, \\ (Q_o - Q)^2 &= 4r(1-r-Q), \\ \frac{1}{4}(Q_o - Q)^2 &= r(1-Q) - r^2, \\ 0 &= r^2 - (1-Q)r + \frac{1}{4}(Q_o - Q)^2. \end{aligned} \tag{4.1.3}$$

Solving the quadratic equation and taking the smaller root, *i.e* the smaller error rate:

$$r = P_E = \frac{1}{2}[(1-Q) - \sqrt{(1-Q)^2 - (Q_o - Q)^2}], \tag{4.1.4}$$

$$p = P_S = \frac{1}{2}[(1-Q) + \sqrt{(1-Q)^2 - (Q_o - Q)^2}]. \tag{4.1.5}$$

The equal priors case requires no further optimization. Simply solving the quadratic equation in (4.1.3) derived from the constraint the optimal solution is carried out. By varying the failure rate Q from zero to Q_o we recover the Helstrom and IDP bounds. In the Helstrom bound [**4**] one is not allowed to have inconclusive results, hence setting failure rate to zero, $Q = 0$ results in

$$P_E = \frac{1}{2}[1 - \sqrt{1 - Q_o^2}] \tag{4.1.6}$$

In the IDP limit [**8, 38, 39**] where one is not allowed to make an error $r = 0$, while allowing for inconclusive outcomes:

$$\begin{aligned} 0 &= \frac{1}{2}[(1-Q) - \sqrt{(1-Q)^2 - (Q_o - Q)^2}], \\ Q &= Q_o, \end{aligned}$$

where Q_o in the overlap of the input states for equal priors $Q_o = s$.

4.1.2. Arbitrary priors. Because of the recent interest in this problem, we feel it is beneficial to show two new and different approaches to its solution. The first is more conceptual: a renormalization inspired by E. Bagan et al. [**14**] allows us to rewrite the problem as a ME problem with an implicit dependence on the last free parameter, the failure rate of one state with relation to the other. This greatly simplifies the problem as the solution to the first part is well known and the second is a straight-forward derivative. The second solution employs a Lagrange multiplier method that is algebraically difficult but useful in its explicit results of individual error rates. This in turn is useful in designing an

implementation scheme where the reflective and transmittance coefficients are expressed in terms of individual error and success rate.

4.1.3. Transformation of the problem into the Helstrom form. Through a renormalization our problem is converted into the well known Helstrom form the solution of which is well known. First we define the useful quantity $\omega \equiv s - \sqrt{q_1 q_2}$ which will serve as normalized overlap and the unitarity constraint from Eq.(4.1.2) reduces to:

$$\omega = \sqrt{p_1 r_2} + \sqrt{p_2 r_1}. \tag{4.1.7}$$

Next we renormalize all probabilities in the problem:

$$\begin{aligned} \tilde{p_i} &= \frac{p_i}{\alpha_i}, \\ \tilde{r_i} &= \frac{r_i}{\alpha_i}, \\ \tilde{\omega} &= \frac{\omega}{\alpha_1 \alpha_2}, \end{aligned} \tag{4.1.8}$$

where $\alpha_i = 1 - q_i$. Now the probabilities are normalized $\widetilde{r}_i + \tilde{p}_i = 1$, and the new normalized overlap is in terms of $\tilde{r_i}$ and $\tilde{p_i}$,

$$\tilde{\omega} = \sqrt{\tilde{p_1}\tilde{r_2}} + \sqrt{\tilde{p_2}\tilde{r_1}} \tag{4.1.9}$$

Using the above transformation of r_i the error rate can be expressed as:

$$\tilde{P}_E = \tilde{\eta}_1 \tilde{r}_1 + \tilde{\eta}_2 \tilde{r}_2, \tag{4.1.10}$$

where $\tilde{P}_E = \frac{P_E}{\eta_1\alpha_1 + \eta_2\alpha_2} = \frac{P_E}{1-Q}, \tilde{\eta}_i = \frac{\eta_i \alpha_i}{\eta_1\alpha_1 + \eta_2\alpha_2} = \frac{\eta_i \alpha_i}{1-Q}$ and $\tilde{\eta}_1 + \tilde{\eta}_2 = 1$.

We have transformed the problem into a discrimination between two pure states with overlap $\tilde{\omega}$ and no explicit failure rate. Hence we can simply write down the expression for the minimum error solution of two pure states (the Helstrom bound), and then replace the

normalized quantities with the original expressions:

$$\tilde{P}_E = \frac{1}{2}[1 - \sqrt{1 - 4\tilde{\eta}_1\tilde{\eta}_2\tilde{\omega}^2}],$$
$$P_E = \frac{1}{2}[(1-Q) - \sqrt{(1-Q)^2 - 4\eta_1\eta_2(s - \sqrt{q_1q_2})^2}]. \tag{4.1.11}$$

There is now one last optimization. Given a fixed rate of the average of inconclusive outcomes, Q, what are the individual failure rates q_i. To minimize Eq. (4.1.11) we maximize the square root term $\sqrt{(1-Q)^2 - 4\eta_1\eta_2(s - \sqrt{q_1q_2})^2}$, which in turn means to minimize $(s - \sqrt{q_1q_2})^2 = (s^2 + q_1q_2 - 2s\sqrt{q_1q_2})$. The overlap term s^2 is fixed and we are left with $q_1q_2 - 2s\sqrt{q_1q_2}$. Only one of the $q_i's$ is an independent variable as we fix the overall failure rate $Q = \eta_1q_1 + \eta_2q_2$,

$$\begin{aligned}
\Theta &= q_1q_2 - 2s\sqrt{q_1q_2}. \\
\Theta &= \frac{q_1(Q_0 - \eta_1q_1^2)}{\eta_2} - 2s\sqrt{\frac{q_1(Q_0 - \eta_1q_1^2)}{\eta_2}} \\
\Theta &= \frac{\eta_1q_1Q_0 - \eta_1^2q_1^2}{\eta_1\eta_2} - 2s\sqrt{\frac{\eta_1q_1Q_0 - \eta_1^2q_1^2}{\eta_1\eta_2}} \\
\Theta &= \frac{\tilde{q}_1Q_0 - \tilde{q}_1^2}{\eta_1\eta_2} - 2s\sqrt{\frac{\tilde{q}_1Q_0 - \tilde{q}_1^2)}{\eta_1\eta_2}}
\end{aligned}$$

where $\eta_1q_1 = \tilde{q}$. Let's optimize with respect to $\tilde{q}_1$.

$$0 = \frac{\partial\Theta}{\partial\tilde{q}_1} = \frac{Q_0 - 2\tilde{q}_1}{\eta_1\eta_2} - 2s\sqrt{\frac{Q_0 - 2\tilde{q}_1}{\eta_1\eta_2}} \tag{4.1.12}$$

this leads to the optimality condition $Q_0 = 2\tilde{q}_1 = 2\eta_1q_1 = 2\eta_2q_2$, giving the minimal error rate in discriminating two pure states with a fixed rate of failure as

$$P_E = \frac{1}{2}[(1-Q) - \sqrt{(1-Q)^2 - (Q_o - Q)^2}]. \tag{4.1.13}$$

Here $Q_o = 2\sqrt{\eta_1\eta_2}s$ is the failure rate in the optimal unambiguous state discrimination, which our expression reaches when we set $P_E = 0$. On the other hand when the failure rate is zero we can recover the Helstrom bound for two pure states $P_E = \frac{1}{2}[1 - \sqrt{1 - 4\eta_1\eta_2 s^2}]$.

4.2. Lagrange Multipliers Method

While the above method gives a closed form solution of the average error rate in terms of a FRIO it does not produce individual error or success rates, i.e the error rates of mistaking state $|\psi_i\rangle$ for state $|\psi_j\rangle$, which are needed for the implementation in calculating the transmittance and reflection coefficients of the beam splitters. To obtain these expressions we show another solution to the interpolation using the Lagrange multipliers method with the constraint in Eq. (4.1.2).

We want to minimize the average error rate $P_E = \eta_1 r_1 + \eta_2 r_2$ subject to the constraint $s = \sqrt{(1 - r_1 - q_1)r_2} + \sqrt{(1 - r_2 - q_2)r_1} + \sqrt{q_1 q_2}$, setting up the function with one Lagrange multiplier λ :

$$F = \eta_1 r_1 + \eta_2 r_2 + \lambda(s - \sqrt{(1 - r_1 - q_1)r_2} - \sqrt{(1 - r_2 - q_2)r_1} - \sqrt{q_1 q_2}). \tag{4.2.1}$$

Setting the derivative $dF_{(r_1,r_2,\lambda)}/dr_i$ to zero then solving for $r_i(\lambda)$, we exploit the symmetry in the resulting equations to solve for the individual error rates r_i as a function of the failure rates q_i. Subsequent substitution into the constraint gives us the optimal value of λ. Then we can obtain the total minimum error by replacing the expressions of r_i into P_e and minimizing P_e under the additional constraint that $\eta_1 q_1 + \eta_2 q_2 = Q$. This gives us the optimal relationship between failure rates as $\eta_1 q_1 = \eta_2 q_2$ and the total optimal error rate as $Q = 2\eta_1 q_1 = 2\eta_2 q_2$.

Setting $dF/dr_1 = 0$ and re-arranging we get:

$$(2\eta_1/\lambda)\sqrt{(1 - q_1 - r_1)r_1} = \sqrt{(1 - q_2 - r_2)(1 - q_1 - r_1)} - \sqrt{r_1 r_2}.$$

Similarly $dF/dr_2 = 0$ gives

$$(2\eta_2/\lambda)\sqrt{(1-q_2-r_2)r_2} = \sqrt{(1-q_2-r_2)(1-q_1-r_1)} - \sqrt{r_1 r_2}. \tag{4.2.2}$$

This step is algebraically challenging and requires the insight that the resulting equations can each be separated into two expressions, left hand side depending on only r_1 or r_2 and the right hand sides are equivalent. Because both equations have the same multivariable expression, we can set the left hand sides equal to a constant, C, which is yet to be determined.

$$(2\eta_1/\lambda)\sqrt{(1-q_1-r_1)r_1} = (2\eta_2/\lambda)\sqrt{(1-q_2-r_2)r_2} \equiv C. \tag{4.2.3}$$

This greatly simplifies the problem, turning it into a quadratic equation.

$$(2\eta_i/\lambda)\sqrt{(1-q_i-r_i)r_i} = C. \tag{4.2.4}$$

Let $\alpha_i \equiv 1 - q_i$

$$\begin{aligned}
\sqrt{(\alpha_1 - r_i)r_i} &= (\lambda C)/(2\eta_i), \\
r_i^2 - \alpha_i r_i + \lambda^2 C^2/4\eta_i^2 &= 0, \\
r_i(\pm) &= 1/2\left(\alpha_i \pm \sqrt{\alpha_i^2 - \frac{\lambda^2 C^2}{\eta_i^2}}\right).
\end{aligned}$$

r_i is the error rate which we want to be minimized, thus we take the smaller root $r_i(-)$.

$$r_1 = 1/2\left(\alpha_1 - \sqrt{\alpha_1^2 - \frac{\delta}{\eta_1^2}}\right) = 1/2\left(\alpha_1 - A_1\right), \tag{4.2.5}$$

$$r_2 = 1/2\left(\alpha_2 - \sqrt{\alpha_2^2 - \frac{\delta}{\eta_2^2}}\right) = 1/2\left(\alpha_2 - A_2\right), \tag{4.2.6}$$

where $\delta \equiv \lambda^2 C^2$ and $A_i \equiv \sqrt{\alpha_i^2 - (\frac{\lambda^2 C^2}{\eta_i^2})}$.

Insert the expression of r_i from (4.2.5) and (4.2.6) into the constraint from (4.1.2). First let us rewrite the constraint so it simplifies the algebra later:

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

where $\omega \equiv (s - \sqrt{q_1 q_2})^2$. We use the definition in (4.2.4) to reduce the square root term $\sqrt{(1 - r_1 - q_1)(1 - r_2 - q_2)r_1 r_2} = \frac{\lambda^2 C^2}{4\eta_1 \eta_2}$, giving

$$(s - \sqrt{q_1 q_2})^2 = (1 - r_1 - q_1)r_2 + (1 - r_2 - q_2)r_1 + \frac{\lambda^2 C^2}{4\eta_1 \eta_2}.$$

Using $\alpha_i = 1 - q_i$, the above expression becomes:

$$\begin{aligned}
\omega &= (\alpha_1 - q_1)r_2 + (\alpha_2 - q_2)r_1 + \frac{\delta}{4\eta_1 \eta_2} \\
&= \alpha_1 r_2 - \alpha_2 r_1 - 2r_1 r_2 + \frac{\delta}{4\eta_1 \eta_2} \\
&= \frac{\alpha_1}{2}(\alpha_2 - A_2) + \frac{\alpha_2}{2}(\alpha_1 - A_1) - \frac{1}{2}(\alpha_1 - A_1)(\alpha_2 - A_2) + \frac{\delta}{4\eta_1 \eta_2} \\
&= \alpha_1 \alpha_2 - A_1 A_2 + \frac{\delta}{4\eta_1 \eta_2}.
\end{aligned}$$

Replacing A_i by their respective value

$$\begin{aligned}
\omega &= \frac{1}{2}\left[\alpha_1 \alpha_2 - \sqrt{\alpha_1^2 - \frac{\delta}{\eta_1^2}}\sqrt{\alpha_2^2 - \frac{\delta}{\eta_2^2}} + \delta/\eta_1 \eta_2\right], \\
\sqrt{\alpha_1^2 - \frac{\delta}{\eta_1^2}}\sqrt{\alpha_2^2 - \frac{\delta}{\eta_2^2}} &= \alpha_1 \alpha_2 + \delta/\eta_1 \eta_2 - 2\omega.
\end{aligned}$$

Squaring both sides, using $\frac{\alpha_1}{\eta_2} + \frac{\alpha_2}{\eta_1} = \frac{\alpha_1 \eta_1 + \alpha_2 \eta_2}{\eta_1 \eta_2} = \frac{\eta_1(1-q_1) + \eta_2(1-q_2)}{\eta_1 \eta_2} = \frac{1-Q}{\eta_1 \eta_2}$ and after some trivial algebra we get:

$$\begin{aligned}
\delta\left[\frac{4\omega}{\eta_1\eta_2}-\left(\frac{\alpha_1}{\eta_2}+\frac{\alpha_2}{\eta_1}\right)\right] &= 4\omega^2-4\omega\alpha_1\alpha_2,\\
\delta\left[\frac{4\omega}{\eta_1\eta_2}-\frac{(1-Q)^2}{\eta_1^2\eta_2^2}\right] &= 4\omega^2-4\omega\alpha_1\alpha_2,\\
\delta &= \frac{(\omega^2-\omega\alpha_1\alpha_2)\,4\eta_1^2\eta_2^2}{4\omega\eta_1\eta_2-(1-Q)^2}, \qquad (4.2.7)
\end{aligned}$$

Now substitute δ into r_1 in (4.2.5)

$$\begin{aligned}
r_1 &= \frac{1}{2}\left(\alpha_1-\sqrt{\alpha_1^2-\frac{\lambda^2C^2}{\eta_1^2}}\right)\\
&= \frac{1}{2}\left(\alpha_1-\sqrt{\alpha_1^2-\frac{1}{\eta_1^2}\frac{(\omega^2-\omega\alpha_1\alpha_2)4\eta_1^2\eta_2^2}{4\omega\eta_1\eta_2-(1-Q)^2}}\right)\\
&= \frac{1}{2}\left(\alpha_1-\sqrt{\frac{\alpha_1^2[4\omega\eta_1\eta_2-(1-Q)^2]-4\eta[\omega^2-\omega\alpha_1\alpha_2]}{4\omega\eta_1\eta_2-(1-Q)^2}}\right). \qquad (4.2.8)
\end{aligned}$$

The numerator can be greatly simplified:

$\alpha_1^2[4\omega\eta_1\eta_2-(1-Q)^2]-4\eta_2^2[\omega^2-\omega\alpha_1\alpha_2] = -\alpha_1^2(1-Q)^2-4\eta_2^2\omega^2+4\omega\eta_2\alpha_1[\eta_1\alpha_1+\eta_2\alpha_2]$

$= -[\alpha_1(1-Q)-2\eta_2\omega]^2 = -[(1-q_1)(1-Q)-2\eta_2(s-\sqrt{q_1q_2})^2]^2.$

The calculation for r_2 goes along the same line. Expression for r_1 and r_2 become:

$$\begin{aligned}
r_1 &= \frac{1}{2}\left[(1-q_1)-\frac{(1-q_1)\,(1-Q)-2\eta_2\left(s-\sqrt{q_1q_2}\right)^2}{\sqrt{((1-Q))^2-4\eta_1\eta_2\left(s-\sqrt{q_1q_2}\right)^2}}\right], \qquad (4.2.9)\\
r_2 &= \frac{1}{2}\left[(1-q_2)-\frac{(1-q_2)\,(1-Q)-2\eta_1\left(s-\sqrt{q_1q_2}\right)^2}{\sqrt{(1-Q)^2-4\eta_1\eta_2\left(s-\sqrt{q_1q_2}\right)^2}}\right]. \qquad (4.2.10)
\end{aligned}$$

Finally r_1 and r_2 can be substituted into the overall average error rate $P_E=\eta_1r_1+\eta_2r_2$:

$$
\begin{aligned}
P_E &= \frac{1}{2}\left[1-(\eta_1 q_1+\eta_2 q_2)-\frac{(1-Q)(\eta_1+\eta_2)-(\eta_1 q_1+\eta_2 q_2)-4\eta_1\eta_2\left(s-\sqrt{q_1 q_2}\right)^2}{\sqrt{(1-Q)^2-4\eta_1\eta_2\left(s-\sqrt{q_1 q_2}\right)^2}}\right] \\
&= \frac{1}{2}\left[(1-Q)-\frac{(1-Q)^2-4\eta_1\eta_2\left(s-\sqrt{q_1 q_2}\right)^2}{\sqrt{(1-Q)^2-4\eta_1\eta_2\left(s-\sqrt{q_1 q_2}\right)^2}}\right] \\
&= \frac{1}{2}\left[(1-Q)-\sqrt{(1-Q)^2-4\eta_1\eta_2\left(s-\sqrt{q_1 q_2}\right)^2}\right] \qquad (4.2.11)
\end{aligned}
$$

It has been showed in previous sections that (4.2.11) is optimized for a fixed value of failure rate when $\eta_1 q_1 = \eta_2 q_2$. Eq (4.2.11) then reduces to the now well know FRIO form:

$$P_E = \frac{1}{2}\left[(1-Q)-\sqrt{(1-Q)^2-(Q-Q_0)^2}\right]. \qquad (4.2.12)$$

The individual error and success rates can now be expressed explicitly in terms of η_i, Q_o and most importantly the fixed failure rate Q as:

$$r_i = \frac{1}{2}\left[\left(1-\frac{Q}{2\eta_i}\right)-\frac{\left(1-\frac{Q}{2\eta_i}\right)(1-Q)-\frac{1}{2\eta_i}(Q_o-Q)^2}{\sqrt{(1-Q)^2-(Q-Q_o)^2}}\right], \qquad (4.2.13)$$

$$p_i = \frac{1}{2}\left[\left(1-\frac{Q}{2\eta_i}\right)+\frac{\left(1-\frac{Q}{2\eta_i}\right)(1-Q)-\frac{1}{2\eta_i}(Q_o-Q)^2}{\sqrt{(1-Q)^2-(Q-Q_o)^2}}\right]. \qquad (4.2.14)$$

This is the first time that the individual error and failure rates have been calculated. An immediate application is that it can now be used in the calculations of the coefficients beam splitters.

4.3. Choosing the physical implementation

The main reason to seek a solution using the Neumark setup is because it lends itself to an optical implementation. This implementation, as we will see, can be carried out using only linear optical elements (beam splitters and a mirror). The possible states are represented by single photons and a photodetector will carry out the measurement process at the output. To choose the basis we start with the two mode vacuum state $|00\rangle$. The total number of

photons in both modes is one. This two dimensional Hilbert space where our photons live can be spanned by the states $\{|10\rangle, |01\rangle\}$, where $|10\rangle = a_1^\dagger|00\rangle$ and $|01\rangle = a_2^\dagger|00\rangle$ and $a_i^\dagger$ are creation operators corresponding to two different modes of the electromagnetic field. The basis $\{|10\rangle, |01\rangle\}$ corresponds to the basis of qubit in $\{|0\rangle, |1\rangle\}$. The most general input state can be written as $|\psi_i\rangle = \alpha_i|0\rangle + \beta_i|1\rangle = \alpha_i|10\rangle + \beta_i|01\rangle$ which can be produced by sending a photon into a beam splitter with some transmission and reflection coefficient, where the modes 1 and 2 correspond to the output modes of the beam splitter.

A general $2N$ port is a linear device with N input and N output ports. It can be constructed from beam splitters and mirrors **[35]**. For our work we will need a six-port, three input ports and three output ports. Let us denote the annihilation operators corresponding to the input modes by $a_{j\,\mathrm{in}}$, $j = 1, 2, 3$ then the output operators are given by

$$a_{j\,\mathrm{out}} = U^{-1} a_{i\,\mathrm{in}}\, U = \sum M_{jk} a_{k\,\mathrm{in}} \,,$$

where M_{jk} are the elements of the $N \times N$ unitary matrix M. In the Schrodinger picture, the *in* and *out* states are related by

$$|\psi\rangle_{out} = U\,|\psi\rangle_{in}\,.$$

In general an *in* state that contains a single photon can be described by

$$|\psi\rangle_{in} = \sum_{j=1}^{3} c_j a_j^\dagger\,|000\rangle\,,$$

where $\sum_{j=1}^{3} |c_j| = 1$. The output state is given by

$$\begin{aligned} |\psi\rangle_{in} &= U \sum_{j=1}^{3} c_j a_j^\dagger U^{-1}\,|000\rangle \\ &= \sum_{j,k=1}^{3} c_j M_{jk}^T a_j^\dagger\,|000\rangle\,, \end{aligned}$$

where we have made use of the fact that the vacuum is invariant under the transformation, U. To simplify the notation in the implementation section let:$|100\rangle \equiv |1\rangle$, $|010\rangle \equiv |2\rangle$, $|001\rangle \equiv |3\rangle$.

4.4. Implementation: equal priors

We will first show the implementation of FRIO for equal priors. The Neumark in direct sum notation setup will be used to calculate the unitary matrix. The two input states to be discriminated can be written as $|\psi_1\rangle_{in} = |1\rangle$ and $|\psi_2\rangle_{in} = \cos\theta|1\rangle + \sin\theta|2\rangle$.

$$U|1\rangle = \sqrt{p}|1\rangle + \sqrt{r}|2\rangle + \sqrt{q}|3\rangle, \tag{4.4.1}$$

$$U(\cos\theta|1\rangle + \sin\theta|2\rangle) = \sqrt{r}|1\rangle + \sqrt{p}|2\rangle + \sqrt{q}|3\rangle, \tag{4.4.2}$$

where the error and success rate was calculated in section 4.1.1:

$$r = \frac{1}{2}\left[(1-Q) - \sqrt{(1-Q)^2 - (Q_o - Q)^2}\right], \tag{4.4.3}$$

$$p = \frac{1}{2}\left[(1-Q) + \sqrt{(1-Q)^2 - (Q_o - Q)^2}\right], \tag{4.4.4}$$

From the Neumark setup we can read out six out of nine elements of the unitary matrix. Eq. (4.4.1) gives the first column, Eq. (4.4.2) gives the second column and the last column can be constructed from the conditions of unitarity.

To get first column, multiply Eq. (4.4.1) from the left with $\langle 1|$, $\langle 2|$ and $\langle 3|$:

$$\langle 1|U|1\rangle = U_{11} = \sqrt{p},$$

$$\langle 2|U|1\rangle = U_{21} = \sqrt{q},$$

$$\langle 3|U|1\rangle = U_{31} = \sqrt{r}.$$

To get the second column, multiply 4.4.2 from the left with $\langle 1|$, $\langle 2|$ and $\langle 3|$:

$$\cos\theta\langle 1|U|1\rangle + \sin\theta\langle 1|U|2\rangle = \cos\theta U_{11} + \sin\theta U_{12} = \sqrt{r},$$

$$\cos\theta\langle 2|U|1\rangle + \sin\theta\langle 2|U|2\rangle = \cos\theta U_{21} + \sin\theta U_{22} = \sqrt{p},$$

$$\cos\theta\langle 3|U|1\rangle + \sin\theta\langle 3|U|2\rangle = \cos\theta U_{31} + \sin\theta U_{32} = \sqrt{q}.$$

Using the solution of the first column U_{i1}, gives the entries to second column U_{i2} $(i = 1, 2, 3)$,

$\cos\theta\sqrt{p}_1 + \sin\theta U_{12} = \sqrt{r_2} \Rightarrow U_{12} = \frac{\sqrt{r}-\sqrt{p}Q_o}{\sqrt{1-Q_o^2}}$,
$\cos\theta\sqrt{q_1} + \sin\theta U_{22} = \sqrt{p_2} \Rightarrow U_{22} = \frac{[\sqrt{p}-\sqrt{r}Q_o]}{\sqrt{1-Q_o^2}}$,
$\cos\theta\sqrt{r_1} + \sin\theta U_{32} = \sqrt{q_2} \Rightarrow U_{32} = \sqrt{\frac{Q(1-Q_o)}{1+Q_o}}$,
where $\cos\theta = Q_o$ and $\sin\theta = \sqrt{1-Q_o^2}$

The remaining elements be calculated from the conditions of the unitarity, $U^TU = I$, $U_{i1}^2 + U_{i2}^2 + U_{i3}^2 = 1$ where $i = 1, 2, 3$

$U_{13} = \pm\sqrt{1-U_{11}^2-U_{12}^2} = \pm\sqrt{1-p-\frac{r+pQ_o^2-2\sqrt{pr}Q_0}{1-Q_0^2}} = \pm\sqrt{\frac{Q-Q_o^2+2\sqrt{pr}Q_0}{1-Q_0^2}} = \pm\sqrt{\frac{Q}{1-Q_0}}$,

the relation $\sqrt{pr} = \frac{1}{2}(Q_o - Q)$ was used, which is derived from the multiplication of (4.4.3) and (4.4.4).

$U_{23} = \pm\sqrt{1-U_{21}^2-U_{22}^2} = \pm\sqrt{1-r-\frac{p+rQ_o^2-2\sqrt{pr}Q_0}{1-Q_0^2}} = \pm\sqrt{\frac{Q}{1-Q_0}}$,
$U_{33} = \pm\sqrt{1-U_{31}^2-U_{32}^2} = \pm\sqrt{1-Q-\frac{Q(1-Q_o)}{1+Q_o}} = \pm\sqrt{\frac{1+Q_o-2Q}{1+Q_o}} = \pm\frac{\sqrt{p}+\sqrt{r}}{\sqrt{1+Q_o}}$.

It is important to notice the relation $\sqrt{(1+Q_o-2Q)} = \sqrt{p}+\sqrt{r}$ which is somewhat unexpected.

The full unitary, with the signs of the last column elements chosen so that $U_{13}^2+U_{23}^2+U_{33}^2 = 1$, is

$$U = \begin{pmatrix} \sqrt{p} & \frac{\sqrt{r}-\sqrt{p}Q_o}{\sqrt{1-Q_o^2}} & -\sqrt{\frac{Q}{1+Q_o}} \\ \sqrt{r} & \frac{\sqrt{p}-\sqrt{r}Q_o}{\sqrt{1-Q_o^2}} & -\sqrt{\frac{Q}{1+Q_o}} \\ \sqrt{Q} & \sqrt{\frac{Q(1-Q_o)}{1+Q_o}} & \frac{\sqrt{p}+\sqrt{r}}{\sqrt{1+Q_o}} \end{pmatrix}. \tag{4.4.5}$$

In this representation the above matrix can be shown that it satisfies all the unitary conditions.

The unitary in (4.4.5) interpolates between the optimal ME and UD schemes varying the fixed rate of inconclusive results Q. Setting the failure rate to zero, $Q = 0$, collapses it into the unitary of optimal ME:

$$U_{ME} = \begin{pmatrix} \sqrt{p} & \frac{\sqrt{r}-\sqrt{p}Q_o}{\sqrt{1-Q_o^2}} & 0 \\ \sqrt{r} & \frac{\sqrt{p}-\sqrt{r}Q_o}{\sqrt{1-Q_o^2}} & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which can be simplified further by noticing $(U_{12})^2 = \left(\frac{\sqrt{r}-\sqrt{p}Q_o}{\sqrt{1-Q_o^2}}\right)^2 = r$, similarly $(U_{22})^2 = \left(\frac{\sqrt{p}-\sqrt{r}Q_o}{\sqrt{1-Q_o^2}}\right)^2 = p$, simplifying the unitary into:

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

only one beam splitter in needed for optimal ME measurements.

On the other hand setting the error rate of the unitary 4.4.5 to zero gives the optimal UD unitary,

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

All three beamsplitters are still necessary for a general UD measurement. This is because the measurement is essentially two-step: in the first step the states are made orthogonal, then upon succeeding a projective measurement is performed.

Let us now express the interpolating unitary in terms of three beamsplitters, as $U = M_1 M_2 M_3$. This ordering was derived using the Reck-Zeilinger algorithm which says that

any discreet finite unitary matrix can be expressed in terms of beamsplitters and phase shifters. The beamsplitters can be written in terms of $\sin\omega_i$ and $\cos\omega_i$ and it is easy to check the unitarity condition, $\sin^2\omega_i + \cos^2\omega_i = 1$. Then the task is that of calculating ω_i.

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

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

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

$$U = M_1M_2M_3 = \begin{pmatrix} \sqrt{p} & \frac{\sqrt{r}-\sqrt{p}Q_o}{\sqrt{1-Q_o^2}} & \sqrt{\frac{Q}{1+Q_o}} \\ \sqrt{r} & \frac{[\sqrt{p}-\sqrt{r}Q_o]}{\sqrt{1-Q_o^2}} & \sqrt{\frac{Q}{1+Q_o}} \\ \sqrt{Q} & \sqrt{\frac{Q(1-Q_o)}{1+Q_o}} & \sqrt{\frac{1+Q_o-2Q}{1+Q_o}} \end{pmatrix}$$

$$= \begin{pmatrix} \sin\omega_1\sin\omega_2 & \cos\omega_1\sin\omega_3 + \sin\omega_1\cos\omega_2\cos\omega_3 & \cos\omega_1\cos\omega_3 - \sin\omega_1\cos\omega_2\sin\omega_3 \\ \cos\omega_1\sin\omega_2 & -\sin\omega_1\sin\omega_3 + \cos\omega_1\cos\omega_2\cos\omega_3 & -\sin\omega_1\cos\omega_3 - \cos\omega_1\cos\omega_2\sin\omega_3 \\ \cos\omega_2 & -\sin\omega_2\cos\omega_3 & \sin\omega_2\sin\omega_3 \end{pmatrix}.$$

This gives nine equations and only three independent variables to be calculated. All the elements can be obtained by using just U_{31}, U_{32}, U_{21} :

$\cos\omega_2 = U_{31} = \sqrt{Q}, \sin\omega_2 = \sqrt{1-r^2} = \sqrt{1-Q}$,

$\cos\omega_3 = -\sqrt{\frac{Q(1-Q_o)}{(1+Q_o)(1-Q)}}, \sin\omega_3 = \sqrt{\frac{1+Q_o-2Q}{(1-Q)(1+Q_o)}} = \frac{\sqrt{p}+\sqrt{r}}{\sqrt{(1-Q)(1+Q_o)}}$,

$\cos\omega_1 = \sqrt{\frac{r}{1-Q}}, \sin\omega_1 = \sqrt{\frac{p}{1-Q}}$.

The three beamsplitters with the proper coefficients of reflectivity and transmittance are:

$$M_1 = \begin{pmatrix} \sqrt{\frac{p}{1-Q}} & \sqrt{\frac{r}{1-Q}} & 0 \\ \sqrt{\frac{r}{1-Q}} & -\sqrt{\frac{p}{1-Q}} & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$M_2 = \begin{pmatrix} \sqrt{1-Q} & 0 & \sqrt{Q} \\ 0 & 1 & 0 \\ \sqrt{Q} & 0 & -\sqrt{1-Q} \end{pmatrix},$$

$$M_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{\frac{1+Q_o-2Q}{(1-Q)(1+Q_o)}} & -\sqrt{\frac{Q(1-Q_o)}{(1+Q_o)(1-Q)}} \\ 0 & -\sqrt{\frac{Q(1-Q_o)}{(1+Q_o)(1-Q)}} & -\sqrt{\frac{1+Q_o-2Q}{(1-Q)(1+Q_o)}} \end{pmatrix},$$

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

4.5. Implementation: Unequal priors

In this section we derive the beamsplitter coefficients necessary to interpolate minimum error measurements with a $FRIO$ when the input states are prepared with different a priori probabilities. The two input states to be discriminated are $|\psi_1\rangle_{in} = |1\rangle$ and $|\psi_2\rangle_{in} = \cos\theta|1\rangle + \sin\theta|2\rangle$. A unitary operator carries out the operation:

$$U|1\rangle = \sqrt{p_1}|1\rangle + \sqrt{r_1}|2\rangle + \sqrt{q_1}|3\rangle, \tag{4.5.1}$$

$$U(\cos\theta|1\rangle + \sin\theta|2\rangle) = \sqrt{r_2}|1\rangle + \sqrt{p_2}|2\rangle + \sqrt{q_2}|3\rangle. \tag{4.5.2}$$

The error and success rates where calculated in Section 4.1.4

$$r_i = \frac{1}{2}\left[(1-\frac{Q}{2\eta_i}) - \frac{(1-\frac{Q}{2\eta_i})(1-Q) - \frac{1}{2\eta_i}(Q_o-Q)^2}{\sqrt{(1-Q)^2-(Q-Q_o)^2}}\right], \tag{4.5.3}$$

$$p_i = \frac{1}{2}\left[(1-\frac{Q}{2\eta_i}) + \frac{(1-\frac{Q}{2\eta_i})(1-Q) - \frac{1}{2\eta_i}(Q_o-Q)^2}{\sqrt{(1-Q)^2-(Q-Q_o)^2}}\right]. \tag{4.5.4}$$

From the two equation in (4.5.1) and (4.5.2) we can read out six of nine elements of the three by three unitary matrix. Multiplying (4.5.1) on the left hand side by $\{\langle 1|, \langle 2|, \langle 3|\}$ will give the elements U_{i1}, $i = 1, 2, 3$. Similarly we can obtain three more elements for the second column.

The first column is:

$\langle 1|U|1\rangle = U_{11} = \sqrt{p_1}$,

$\langle 2|U|1\rangle = U_{21} = \sqrt{q_1}$,

$\langle 3|U|1\rangle = U_{31} = \sqrt{r_1}$.

Second Column:

$\cos\theta\langle 1|U|1\rangle + \sin\theta\langle 1|U|2\rangle = \cos\theta U_{11} + \sin\theta U_{12} = \sqrt{r_2}$,

$\cos\theta\langle 2|U|1\rangle + \sin\theta\langle 2|U|2\rangle = \cos\theta U_{21} + \sin\theta U_{22} = \sqrt{p_2}$,

$\cos\theta\langle 3|U|1\rangle + \sin\theta\langle 3|U|2\rangle = \cos\theta U_{31} + \sin\theta U_{32} = \sqrt{q_2}$.

Using the solution of the first column U_{i1}, gives the explicit entries to second column U_{i2} :

$\cos\theta\sqrt{p}_1 + \sin\theta U_{12} = \sqrt{r_2} \Rightarrow U_{12} = \frac{\sqrt{r_2}-\sqrt{p_1}\cos\theta}{\sin\theta}$,

$\cos\theta\sqrt{q_1} + \sin\theta U_{22} = \sqrt{p_2} \Rightarrow U_{22} = \frac{\sqrt{p_2}-\sqrt{r_1}\cos\theta}{\sin\theta}$,

$\cos\theta\sqrt{r_1} + \sin\theta U_{32} = \sqrt{q_2} \Rightarrow U_{32} = \frac{\sqrt{q_2}-\sqrt{q_1}\cos\theta}{\sin\theta}$,

The last column be calculated from the conditions of the unitarity, $U^TU = I$, $U_{i1}^2 + U_{i2}^2 + U_{i3}^2 = 1$ where $i = 1, 2, 3$

$U_{13} = \pm\sqrt{1-U_{11}^2-U_{12}^2} = \pm\sqrt{1-p_1-\frac{r_2+p_1\cos\theta-2\sqrt{p_1r_2}\cos\theta}{\sin^2\theta}} = \pm\frac{\sqrt{\sin^2\theta-p_1-r_2+2\sqrt{p_1r_2}\cos\theta}}{\sin\theta}$.

Similarly:

$U_{23} = \pm\frac{\sqrt{\sin^2\theta-r_1-p_2+2\sqrt{p_2r_1}\cos\theta}}{\sin\theta}$,

$U_{33} = \pm\frac{\sqrt{sin^2\theta-q_1-q_2+2\sqrt{q_1q_2}\cos\theta}}{\sin\theta}$.

Now that all the elements have been calculated the unitary is:

$$U = \begin{pmatrix} \sqrt{p_1} & \frac{\sqrt{r_2}-\sqrt{p_1}\cos\theta}{\sin\theta} & -\frac{\sqrt{\sin^2\theta-p_1-r_2+2\sqrt{p_1 r_2}\cos\theta}}{\sin\theta} \\ \sqrt{r_1} & \frac{\sqrt{p_2}-\sqrt{r_1}\cos\theta}{\sin\theta} & -\frac{\sqrt{\sin^2\theta-r_1-p_2+2\sqrt{p_2 r_1}\cos\theta}}{\sin\theta} \\ \sqrt{q_1} & \frac{\sqrt{q_2}-\sqrt{q_1}\cos\theta}{\sin\theta} & +\frac{\sqrt{sin^2\theta-q_1-q_2+2\sqrt{q_1 q_2}\cos\theta}}{\sin\theta} \end{pmatrix}. \tag{4.5.5}$$

It is worth mentioning that all equations in this section referencing r_i and p_i are using the optimal values (4.2.13) and (4.2.14) derived in the previous section.

Now that we have a full unitary matrix we want to express it in terms of linear optical devices. Again the Reck-Zeilinger algorithm is used to decompose the unitary in terms of beamsplitters and their ordering 4.5.1. 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 $\sqrt{T_i} = \cos\omega_i$.

$$U = M_1 M_2 M_3 = \begin{pmatrix} \sqrt{p_1} & \frac{\sqrt{r_2}-\sqrt{p_1}cos\theta}{sin\theta} & \pm\frac{\sqrt{sin^2\theta-p_1-r_2+2\sqrt{p_1 r_2}cos\theta}}{sin^2\theta} \\ \sqrt{r_1} & \frac{\sqrt{p_2}-\sqrt{r_1}cos\theta}{sin\theta} & \pm\frac{\sqrt{sin^2\theta-r_1-p_2+2\sqrt{p_2 r_1}cos\theta}}{sin^2\theta} \\ \sqrt{q_1} & \frac{\sqrt{q_2}-\sqrt{q_1}cos\theta}{sin\theta} & \pm\frac{\sqrt{sin^2\theta-q_1-q_2+2\sqrt{q_1 q_2}cos\theta}}{sin^2\theta} \end{pmatrix}$$

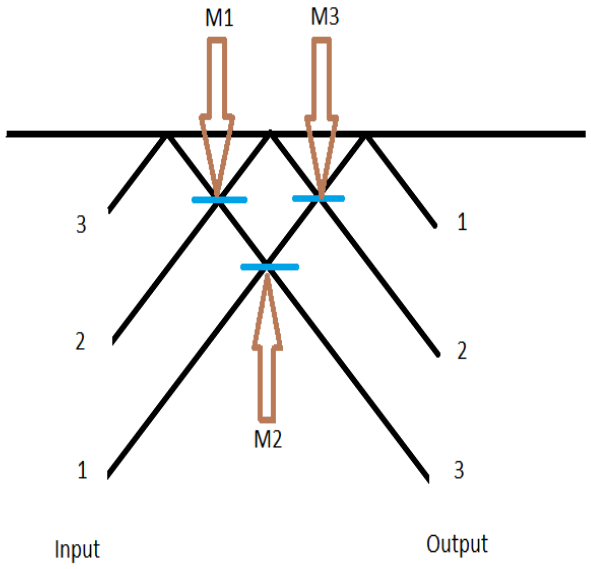


FIGURE 4.5.1

$$= \begin{pmatrix} sin\omega_1 sin\omega_2 & cos\omega_1 sin\omega_3 + sin\omega_1 cos\omega_2 cos\omega_3 & cos\omega_1 cos\omega_3 - sin\omega_1 cos\omega_2 sin\omega_3 \\ cos\omega_1 sin\omega_2 & -sin\omega_1 sin\omega_3 + cos\omega_1 cos\omega_2 cos\omega_3 & -sin\omega_1 cos\omega_3 - cos\omega_1 cos\omega_2 sin\omega_3 \\ cos\omega_2 & -sin\omega_2 cos\omega_3 & sin\omega_2 sin\omega_3 \end{pmatrix}$$

The coefficients of reflectivity and transmittance can be calculated by matching the corresponding elements of the unitary and it's decomposition. We can get all the elements by using just U_{31}, U_{32} and U_{33}.

$U_{31} = \sqrt{q_1} = \cos\omega_2, \sin\omega_2 = \sqrt{1-\cos\omega_2^2} = \sqrt{1-q_1}$,

$U_{32} = \frac{\sqrt{q_2}-\sqrt{q_1}cos\theta}{sin\theta} = -\sin\omega_2\cos\omega_3 = -\sqrt{1-q_1}\cos\omega_3 \Rightarrow \cos\omega_3 = -\frac{1}{\sqrt{1-q_1}}[\frac{\sqrt{q_2}-\sqrt{q_1}cos\theta}{sin\theta}]$,

$\sin\omega_3 = \sqrt{1-\cos\omega_3} = -\frac{\sqrt{sin^2\theta - q_1 - q_2 + 2\sqrt{q_1 q_2}cos\theta}}{\sqrt{1-q_1}sin\theta}$,

$U_{21} = \sqrt{r_1} = \cos\omega_1 \sin\omega_2 = \cos\omega_1\sqrt{1-q_1} \Rightarrow \cos\omega_1 = \sqrt{\frac{r_1}{1-q_1}}$,

$\sin\omega_1 = \sqrt{1-\cos\omega_1} = \sqrt{\frac{p_1}{1-q_1}}$.

Substituting the coefficients of reflectivity and transmittance the beamsplitters are:

$$M_1 = \begin{pmatrix} \sqrt{\frac{p_1}{1-q_1}} & \sqrt{\frac{r_1}{1-q_1}} & 0 \\ \sqrt{\frac{r_1}{1-q_1}} & -\sqrt{\frac{p_1}{1-q_1}} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; M_2 = \begin{pmatrix} \sqrt{1-q_1} & 0 & \sqrt{q_1} \\ 0 & 1 & 0 \\ \sqrt{q_1} & 0 & -\sqrt{1-q_1} \end{pmatrix},$$

$$M_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{sin^2\theta - q_1 - q_2 + 2\sqrt{q_1 q_2}cos\theta}}{\sqrt{1-q_1}sin\theta} & -\frac{1}{\sqrt{1-q_1}}\left[\frac{\sqrt{q_2}-\sqrt{q_1}cos\theta}{sin\theta}\right] \\ 0 & -\frac{1}{\sqrt{1-q_1}}\left[\frac{\sqrt{q_2}-\sqrt{q_1}cos\theta}{sin\theta}\right] & -\frac{\sqrt{sin^2\theta - q_1 - q_2 + 2\sqrt{q_1 q_2}cos\theta}}{\sqrt{1-q_1}sin\theta} \end{pmatrix}.$$

All the coefficients can be expressed in terms of the $FRIO$. 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 above expressions of success and error rates.

$\cos\omega_1 = \sqrt{\frac{r_1}{1-Q/2\eta_1}}$, $\sin\omega_1 = \sqrt{\frac{p_1}{1-Q/2\eta_1}}$,
$\cos\omega_2 = \sqrt{Q/2\eta_1}$, $\sin\omega_2 = \sqrt{1-Q/2\eta_1}$,
$\cos\omega_3 = -\frac{\sqrt{Q/2\eta_2} - Q_o/2\eta_1\sqrt{Q/2\eta_2}}{\sqrt{(1-Q/2\eta_1)(1-Q_o^2/4\eta_1\eta_2)}}$,
$\sin\omega_3 = \frac{\sqrt{1-Q_o^2/4\eta_1\eta_2 - Q/(2\eta_1\eta_2) + QQ_o/(2\eta_1\eta_2)}}{\sqrt{(1-Q/2\eta_1)(1-Q_o^2/4\eta_1\eta_2)}}$,

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

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

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

Let us now check the bounds of the general unitary matrix for equal priors to see if it reproduces the unitary in (4.4.5). Indeed, everything checks out and the equal priors unitary

matrix is reproduced:

$$U = \begin{pmatrix} \sqrt{p} & \frac{\sqrt{r}-\sqrt{p}Q_o}{\sqrt{1-Q_o^2}} & \sqrt{\frac{Q}{1+Q_o}} \\ \sqrt{r} & \frac{[\sqrt{p}-\sqrt{r}Q_o]}{\sqrt{1-Q_o^2}} & \sqrt{\frac{Q}{1+Q_o}} \\ \sqrt{Q} & \sqrt{\frac{Q(1-Q_o)}{1+Q_o}} & -\frac{\sqrt{p}+\sqrt{r}}{\sqrt{1+Q_o}} \end{pmatrix}, \tag{4.5.6}$$

In summary in this chapter we have shown that two nonorthogonal quantum states, each realized as a photon split between two modes, in combination with a six-port interferometer can be used to implement state discrimination with FRIO. The implementation requires only optical elements: beam splitters and mirrors. All the proper coefficients of transmittance and reflectivity were calculated and we believe it should be possible to construct the setup in a laboratory. The setup should give the experimentalist more freedom when designing a quantum computation network as now only one setup is needed to perform UD, ME and interpolate with a fixed rate of inconclusive results.

Appendix 1: Reck-Zeilinger Algorithm

Optimizing the function $f(x_1, x_2, ..., x_n)$ we differentiate with respect to all the independent variables $(x_1, x_2, ..., x_n)$ and follow the procedure defined above.

In their letter **[37]** prove that any discrete finite-dimensional unitary operator can be constructed using optical devices only. Then they provide a general algorithm which decomposes any N X N unitary matrix into a product of two-dimensional $U(2)$ transformations which can be expressed as beam splitters, phase shifters and mirrors. This optical multi-port can act upon various fields such as electrons, neutrons, atoms, photons etc. The authors decide to work with photons purely for convenience and widespread availability of high power lasers. It is this very proof which allows us to implement our various works in state discrimination and cloning. In addition the proof has greatly simplified the experimental realizations of many quantum computation, quantum information and quantum cryptography schemes. Besides these very practical applications it has also answered a long standing question: Does an experiment measuring the variables corresponding to any arbitrary Hermitian operator exists? They show that indeed an experimental realization does exist for an arbitrary operator in a finite dimensional Hilbert space.

It has long been known that a lossless beam splitter and a phase shifter can implement any $U(2)$ transformation: a beam splitter and a phase shifter at one output port transforms the input operators into output operators as

$$\begin{pmatrix} a_1' \\ a_2' \end{pmatrix} = \begin{pmatrix} e^{i\phi}\sin\omega & e^{i\phi}\cos\omega \\ \cos\omega & -\sin\omega \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \tag{4.5.7}$$

where, ϕ is the phase shifter which can be realized as an external phase shifter after the beam splitter, ω represents the transmittance and reflectivity coefficient, $\sqrt{T} = \cos\omega$, $\sqrt{R} = \sin\omega$.

In their Letter, Reck *et al*. considered the use of a Mach-Zehner interferometer to simulate the effect of a beam splitter which splits the incoming beam according to the given parameters of transmittance and reflectivity. For an actual two by two beam splitter the coefficients of transmittance and reflectivity should be $\sqrt{T} = \sin\omega$ and $\sqrt{R} = \sin\theta$.

The authors show that starting with an $N \times N$ unitary matrix $U(N)$, it can be expressed into a succession of two-dimensional matrices which correspond to beam splitters and phase shifters. Hence the $U(N)$ unitary matrix can be realized in the full N dimensional Hilbert space through a succession of two-dimensional $U(2)$ matrices.

The order in which the matrices are multiplied correspond to the sequence in which the beamsplitters are set up. The task of realizing the experimental setup of an arbitrary unitary matrix becomes that of factorizing the matrix in terms of two dimensional beam splitter matrices with phase shifters which can be placed after the beam splitters.

Define an $N-$dimensional identity matrix T_{pq} which multiplies the N dimensional unitary matrix from the right to reduce the dimensionality to $N-1$. In the identity matrix T_{pq} the elements $I_{pq},\ I_{pp},\ I_{qp},\ I_{qq}$ are replaced by the corresponding beam splitter matrix elements $(\cos\omega, \sin\omega)$. Thus:

$$U(N) \times T_{N,N-1} \times T_{N,N-2} \times ...T_{N,1} = \begin{pmatrix} U(N-1) & 0 \\ 0 & e^{i\phi} \end{pmatrix}. \tag{4.5.8}$$

This reduces the dimensionality of $U(N)$ to $U(N-1)$. The process is repeated again until all the off diagonal elements of the original unitary matrix are zero.

$$U(N)\cdot T_{N,N-1}\cdot T_{N,N-2}\cdot ...T_{N,1}\cdot T_{N-1,N-2}\cdot T_{N-2,N-2}\cdot ...T_{2,1}\cdot ...T_{2,1} = \begin{pmatrix} e^{i\alpha_1} & 0 & .. & 0 \\ \vdots & e^{i\alpha_2} & & \\ & & \ddots & \\ 0 & \cdots & & e^{i\alpha N} \end{pmatrix}. \tag{4.5.9}$$

Let:

$$D = \begin{pmatrix} e^{-i\alpha_1} & 0 & .. & 0 \\ \vdots & e^{-i\alpha_2} & & \\ & & \ddots & \\ 0 & \cdots & & e^{-i\alpha N} \end{pmatrix}. \tag{4.5.10}$$

Then we have

$$U(N) \cdot T_{N,N-1} \cdot T_{N,N-2} \cdot ...T_{N,1} \cdot T_{N-1,N-2} \cdot T_{N-2,N-2} \cdot ...T_{2,1} \cdot D = I \tag{4.5.11}$$

The unitary matrix can be expressed in terms of Tp, q and D :

$$U(N) = D^{-1} \cdot T_{2,1}.... \cdot T^{-1}_{N-2,N-2} \cdot T^{-1}_{N-1,N-2} \cdot T^{-1}_{N,1}.... \cdot T^{-1}_{N,N-2} \cdot T^{-1}_{N,N-1}. \tag{4.5.12}$$

Since the product of matrices represents the order in which the beam splitters are set up, then (4.5.12) is all one needs to implement a finite dimensional unitary matrix. Since this algorithm is recursive, it can factorize any finite dimensional unitary operator. For example a 3×3 unitary matrix, three beam splitters are needed T_{21}, T_{31}, T_{32}, a 4×4 unitary matrix requires six beamsplitters $T_{4,3}, T_{4,2}, T_{4,1}, T_{32}, T_{31}, T_{21}$ in reversed order. In general the maximum number of beam splitters required for any N dimensional unitary operator is $\binom{N}{2} = \frac{N(N-1)}{2}$. In practice this method involves a triangular array of beamsplitters (4.5.2), with each diagonal row effectively reducing the dimension of the Hilbert space by one.

Let us now give an example to see explicitly how this algorithm works. For a three dimensional unitary operator $U(3)$, the algorithm in (4.5.12) gives:

$$U(3) = D^{-1} \times T^{-1}_{2,1} \times T^{-1}_{3,1} \times T^{-1}_{3,2} \tag{4.5.13}$$

where:

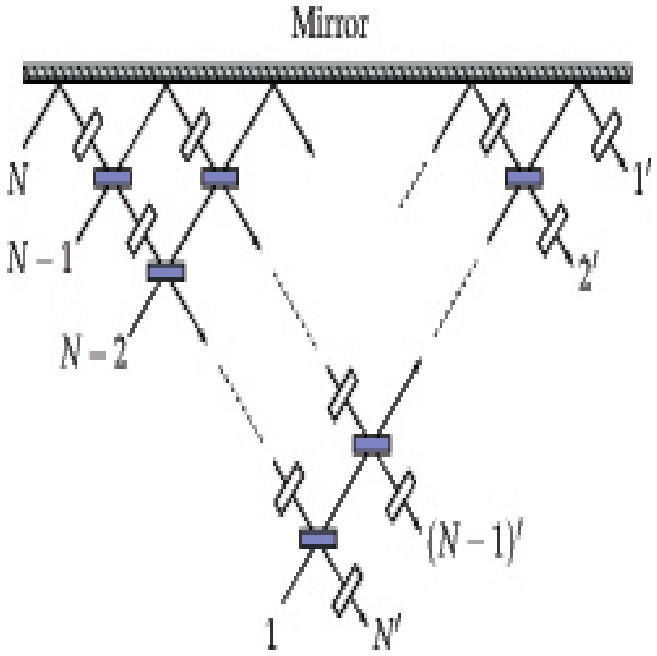


FIGURE 4.5.2

$$D^{-1} = \begin{pmatrix} e^{i\alpha_1} & 0 & 0 \\ 0 & e^{i\alpha_2} & 0 \\ 0 & 0 & e^{i\alpha_3} \end{pmatrix}, \; T_{21} = \begin{pmatrix} \sin\omega_1 & \cos\omega_1 & 0 \\ \cos\omega_1 & -\sin\omega_1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$T_{31} = \begin{pmatrix} \sin\omega_2 & 0 & \cos\omega_2 \\ 0 & 1 & 0 \\ \cos\omega_2 & 0 & -\sin\omega_2 \end{pmatrix}, \; T_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sin\omega_3 & \cos\omega_3 \\ 0 & \cos\omega_3 & -\sin\omega_3 \end{pmatrix}.$$

To obtain the transmittance and reflective coefficients match the corresponding entries of U_{ij} with the elements on right hand side multiplying the three beam splitter matrix with the phase shifter matrix. In the process nine equations and six independent unknowns are produced.

Appendix 2: Lagrange Multipliers

In our works we have relied quite heavily on the method of Lagrange multipliers when optimizing a function which was under the restriction of a constraint. We now show how it works [**40**]. The method can be applied to a function of any number of variables but it can be more clearly explained in two variables. Suppose that we need to find the stationary points of a function $f(x, y)$, where x and y are the two variables, subject to the constraint $g(x, y) = 0$. If the constraint is simple then we can solve for x in terms of y, plug it into the function and solve $\partial f/\partial y = 0$. However for a more complicated constraint this can easily lead to a very high order equation with cannot be solved analytically. In the case of exact cloning, doing just so leads to a sixth order equation which is of little use.

To find the stationary points of a function of two variables such as $f(x, y)$, we could just take the total differential df and set it to zero

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy = 0, \tag{4.5.14}$$

which leads to two conditions:

$$\frac{\partial f}{\partial x} = 0, \quad \frac{\partial f}{\partial y} = 0. \tag{4.5.15}$$

However there is a constraint which means that the differentials dx and dy are not independent, they are related to the total differential of g by:

$$dg = \frac{\partial g}{\partial x}dx + \frac{\partial g}{\partial y}dy = 0. \tag{4.5.16}$$

Multiplying (4.5.16) by the Lagrange parameter λ and adding it to (4.5.14) we get

$$d(f+\lambda g) = (\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x})dx + (\frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y})dy \tag{4.5.17}$$

This equation can be satisfied by choosing the Lagrange multiplier λ such that the following two conditions are satisfied:

$$\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} = 0 \tag{4.5.18}$$

and

$$\frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y} = 0 \tag{4.5.19}$$

To get the stationary points of $f(x, y)$ follow this procedure:

- Solve the two equations: (4.5.18) and (4.5.19) in terms of λ, $x(\lambda)$ and $y(\lambda)$;
- Plug $x(\lambda)$ and $y(\lambda)$ into the constraint $g(x, y)$;
- Solve for λ;
- Plug the value of λ into $x(\lambda)$ and $y(\lambda)$;
- Plug $x(\lambda)$ and $y(\lambda)$ into the function which was to be optimized $f(x, y)$.

Now that we have seen how the Lagrange multipliers method works, we can simplify the procedure by optimizing the function following function:

$$F(x, y) = f(x, y) + \lambda g(x, y) \tag{4.5.20}$$

with respect to the the independent variable x and y. Differentiating (4.5.20) with respect to x and y, we obtain equations (4.5.18) and (4.5.19). The rest of the procedure is the same. This is the exact procedure we used for our works, for example in optimizing the error rate $P_E(r_1, r_2)$ with one constraint $s(r_1, r_2)$.

Bibliography

[1] János A. Bergou and Mark Hillery. *Introduction to the Theory of Quantum Information Processing*. Springer NY„ 2013.

[2] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[3] A. Peres. *Quantum Theory: Concepts and Methods*. Springer NY„ 1995.

[4] Carl W. Helstrom. Quantum detection and estimation theory. *Journal of Statistical Physics*, 1(2):231–252, 1969.

[5] Alexander S Holevo. *Probabilistic and Statistical Aspects of Quantum Theory (Statistics & Probability)*. Springer Science & Business Media, 1982.

[6] P.W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pages 124–134. IEEE Comput. Soc. Press, 1994.

[7] Charles Bennett. Quantum cryptography using any two nonorthogonal states. *Physical Review Letters*, 68(21):3121–3124, May 1992.

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

[9] Anthony Chefles and Stephen M. Barnett. Optimum unambiguous discrimination between linearly independent symmetric states. *Physics Letters A*, 250(4-6):223–229, December 1998.

[10] János A. Bergou. Discrimination of quantum states. *Journal of Modern Optics*, 57(3):160–180, February 2010.

[11] John Von Neumann. *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, 1955.

[12] Ulrike Herzog. Minimum-error discrimination between a pure and a mixed two-qubit state. *Journal of Optics B: Quantum and Semiclassical Optics*, 6(3):S24–S28, March 2004.

[13] Christopher A. Fuchs. Distinguishability and Accessible Information in Quantum Theory. *quant-ph/9601020*, January 1996.

[14] E. Bagan, R. Muñoz Tapia, G. A. Olivares-Rentería, and J. A. Bergou. Optimal discrimination of quantum states with a fixed rate of inconclusive outcomes. *Physical Review A - Atomic, Molecular, and Optical Physics*, 86(4):1–5, 2012.

[15] Anthony Chefles and Stephen M. Barnett. Quantum State Separation, Unambiguous Discrimination and Exact Cloning. page 5, 1998.

[16] János A. Bergou, Vladimír Bužek, Edgar Feldman, Ulrike Herzog, and Mark Hillery. Programmable quantum-state discriminators with simple programs. *Physical Review A*, 73(6):062334, June 2006.

[17] Ulrike Herzog and János A. Bergou. Optimum unambiguous discrimination of two mixed quantum states. *Physical Review A*, 71(5):050301, May 2005.

[18] János Bergou, Edgar Feldman, and Mark Hillery. Optimal unambiguous discrimination of two subspaces as a case in mixed-state discrimination. *Physical Review A*, 73(3):032107, March 2006.

[19] W. K. Wootters and W. H. Zurek. A single quantum cannot be cloned naturevol. 299(5886)p. *Nature 299, 802 - 803*, 299(5886):802–803, 1982.

[20] D. Dieks. Communication by EPR devices. *Physics Letters A*, 92(6):271–272, November 1982.

[21] V Buzek and M Hillery. Quantum copying:beyond the no-cloning theorem. *Physical Review A*, 54(3):1844–1852, 1996.

[22] M. Hillery and V. Bužek. Quantum copying: Fundamental inequalities. *Physical Review A*, 56(2):1212–1216, August 1997.

[23] Lu-Ming Duan and Guang-Can Guo. Probabilistic Cloning and Identification of Linearly Independent Quantum States. *Physical Review Letters*, 80(22):4999–5002, June 1998.

[24] János A. Bergou, Ulrike Futschik, and Edgar Feldman. Optimal Unambiguous Discrimination of Pure Quantum States. *Physical Review Letters*, 108(25):250502, June 2012.

[25] Anthony Chefles and Stephen M Barnett. Quantum state separation, unambiguous discrimination and exact cloning. *Journal of Physics A: Mathematical and General*, 31(50):10097–10103, December 1998.

[26] A Feldman E Bagan E V. Yerokhin A.Shehu E. Feldman E. Bagan J A Bergou Yerokhin, V Shehu. Probabilistically perfect cloning of two pure states: A geometric approach. *To be submitted in PRL*, 2015.

[27] Anthony Chefles and Stephen Barnett. Strategies and networks for state-dependent quantum cloning. *Physical Review A*, 60(1):136–144, July 1999.

[28] Dagmar Bruss, Artur Ekert, and Chiara Macchiavello. Optimal Universal Quantum Cloning and State Estimation. *Physical Review Letters*, 81(12):2598–2601, September 1998.

[29] J. Cirac and P. Zoller. Quantum Computations with Cold Trapped Ions. *Physical Review Letters*, 74(20):4091–4094, May 1995.

[30] N. A. Gershenfeld. Bulk Spin-Resonance Quantum Computation. *Science*, 275(5298):350–356, January 1997.

[31] Q. A. Turchette, C. J. Hood, W. Lange, H. Mabuchi, and H. J. Kimble. Measurement of Conditional Phase Shifts for Quantum Logic. *Physical Review Letters*, 75(25):4710–4713, December 1995.

[32] P. Domokos, J. M. Raimond, M. Brune, and S. Haroche. Simple cavity-QED two-bit universal quantum logic gate: The principle and expected performances. *Physical Review A*, 52(5):3554–3559, November 1995.

[33] G. J. Milburn. Quantum optical Fredkin gate. *Physical Review Letters*, 62(18):2124–2127, May 1989.

[34] Isaac L. Chuang and Yoshihisa Yamamoto. Quantum Bit Regeneration. *Physical Review Letters*, 76(22):4281–4284, May 1996.

[35] János A. Bergou, Mark Hillery, and Yuqing Sun. Non-unitary transformations in quantum mechanics: An optical realization. *Journal of Modern Optics Volume 47, Issue 2-3, 2000*, July 2009.

[36] Vadim Yerokhin Andi Shehu and János A. Bergou. Optical realization of frio for pure states. *To be submitted in PRA*, 2015.

[37] Michael Reck and Anton Zeilinger. Experimental realization of any discrete unitary operator. *Phys. Rev. Lett*, 73(1):58–61, 1994.

[38] D. Dieks. Overlap and distinguishability of quantum states. *Physics Letters A*, 126(5-6):303–306, January 1988.

[39] Asher Peres. How to differentiate between non-orthogonal states. *Physics Letters A*, 128(1-2):19, March 1988.

[40] Stephen Barnett. *Quantum Information*. Oxford University Press, 2009.