Optimum unambiguous state discrimination and its optical realization

by

Yuqing Sun

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

2003

UMI Number: 3103176

UMI Microform 3103176

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

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.

9/8/03
Date

Chair of Examining Committee
Professor János A. Bergou

9/10/03
Date

Executive Officer
Professor Sultan Catto

Professor Mark Hillery

Professor Ying-Chih Chen

Professor Marten denBoer

Professor Daniel Greenberger

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

Optimum unambiguous state discrimination and its optical realization

by

Yuqing Sun

Adviser: Professor János A. Bergou

Discrimination of quantum states is a challenging problem in quantum physics. If the quantum states are nonorthogonal, they can not be discriminated with unit success probability. Optimum unambiguous discrimination means to have the minimum probability of getting inconclusive result. We study the general POVM formulism to derive a necessary condition of optimum unambiguous discrimination. We also provide an optical implementation of any desired unambiguous discrimination. We show that nonorthogonal quantum states, each realized as a photon split among several modes, can be conditionally distinguished by means of a linear optical network. It is discussed in detail on how to discriminate two and three nonorthogonal quantum states. We have given explicit examples together with the optical networks, which give the maximum success probabilities for several sets of states. We also consider the problem of unambiguous discrimination between subsets. We consider the simplest

instance of this problem, the situation in which we are trying to discriminate between a set containing one quantum state and another containing two. A method for finding the optimal strategy for discriminating between these two sets is presented, and analytical solutions for particular cases are given. Finally, we also study the problem of state comparison, and present the analytical solution of optimum unambiguous comparison of two quantum systems, where each of them could be in one of the two known quantum states.

Acknowledgments

I wish to express my deepest gratitude to my two mentors: Professor János A. Bergou and Professor Mark Hillery, for giving me directions, shaping my views and refining my skills in physics. I have benefited greatly from their supports.

I also thank the people who helped me during my study, and shared with me many enjoyable discussions, among them are Gabriel Drobny, Matthias Jakob, István Németh and Yonatan Abranyos.

To my family...

Table of Contents

List of Tables x

List of Figures xi

1 Introduction 1

1.1 Quantum states discrimination with minimum-error 2

1.2 Optimum unambiguous discrimination 2

2 General method of optimal unambiguous discrimination 5

2.1 Quantum measurements 5

2.2 Optimal probabilities for unambiguous discrimination among non-orthogonal quantum states 8

2.3 Realization of generalized measurement 11

2.4 Optical realization of non-unitary transformation 13

3 Application to two states and examples 21

3.1 Dual-rail representation 21

3.2 Optimal discrimination between two non-orthogonal states 22

3.3 Entanglement Enhancement 27

4 Application to three states and examples 33

5 Optimum unambiguous discrimination between $|\psi_1\rangle$ and $\{|\psi_2\rangle, |\psi_3\rangle\}$ 47

5.1 Derivation of the optimal solution 47

5.2 Derivation of the optimal solution via the method of Lagrange multipliers 57

5.3 Comparison to the case when all states are discriminated 60

5.4 Optical realization 64

6 Unambiguous quantum states comparison 74

6.1 Introduction 74

6.2 Distinguish all the states 75

6.3 States Comparison 81

7 Conclusions 89

Bibliography 92

List of Tables

Table 1 Arguments of $T_{p,q}$ for case 1. The range of arccos is the interval $[0, \pi]$. 39

Table 2 Summary of parameters and arguments of $T_{p,q}$ for case 2. 43

Table 3 Summary of parameters and arguments of $T_{p,q}$ for case 3. 45

List of Figures

Figure 1 The implementation of an optical multiport that performs the unitary transformation $M(N)$ described in Eq. (2.33). The beams are straight lines, a suitable beam splitter is at each crossing point of the beams, phase shifters are at one input of each beam splitter and at the outputs of the multiport. Each diagonal line of the multiport reduces the dimension of $M(N)$ by one. 18

Figure 2 Let $A = |O_{12}|^2$, when $\frac{A}{1+A} \leqslant \eta_1 \leqslant \frac{1}{1+A}$, the optimal measurement is a POVM. When η_1 is outside the boundaries, the optimal measurement is a standard Von Neumann measurement. The failure probability of the POVM is always less than the Von Neumann measurement, when $\eta_1 = \eta_2 = 1/2$, the difference between the two types of measurement is at the largest. 24

Figure 3 We compare Q and Q'. For $0 < s \leq \frac{\sqrt{2}}{2}$ we have that $Q' = s$ and $Q = \frac{2\sqrt{2}}{3}s$. For $\frac{\sqrt{2}}{2} < s \leq 1$, we still have that $Q' = s$, but $Q = \frac{1}{3} + \frac{2}{3}s^2$. Note that Q is always smaller than Q'. 62

Figure 4 An optical eight-port. The beams are straight lines, a suitable beam splitter is placed at each point where two beams intersect, phase shifters are at one input of each beam splitter and at each output. 64

Figure 5 The eight-port described by Eq. (5.68) can be constructed from three beam splitters and a mirror.. 71

Chapter 1

Introduction

Quantum information theory is an emerging science, which combines two traditional disciplines: quantum mechanics and classical information theory. This subject has many fascinating potential applications for the transmission and processing of information, and yields results that cannot be achieved by classical means. The possibility of using non-orthogonal quantum states, which has no classical analogue, is especially interesting for its potential applications such as quantum cryptography [1], quantum cloning [2, 3], the purification of entanglements [4, 5], and quantum teleportation [6].

The problem of quantum states discrimination considers the following: a quantum system is prepared in a member of a known, finite set of state, and we want to know which one it is. According to the quantum theory of measurement, if the quantum states are non-orthogonal, then it is impossible to discriminate them with unit successful probability [7]. Hence different strategies exist. In minimum-error discrimination, the aim is to guess which one with the minimum probability of error. Error free discrimination is also possible, but then we must allow for the possibility of obtaining inconclusive results.

1.1. Quantum states discrimination with minimum-error

In order to devise an optimum state-discrimination measurement, strategies have been developed with respect to various criteria. The earliest strategy was first advanced by C. W. Helstrom [8], his attempt is to have the probability of getting a wrong result be as small as possible, with no inconclusive result so that all states are individually distinguished. This was known as *quantum hypothesis testing*. If the states are not orthogonal, then there will be a non-zero probability of error, P_E, optimum discrimination here means to keep P_E as minimum. Therefore, this strategy is also known as *minimum error* strategy. A minimum-error strategy of this kind has been developed by for the case when only two states are given [8], and for *specific* N state problems [9–13], including N symmetric [11] and multiply symmetric [13] states. Recently the optimum strategy has also been found for three states exhibiting a mirror-symmetry [14]. U. Herzog and J. Bergou solved the case for discrimination between two subsets that involves $N > 2$ *arbitrary* linearly dependent quantum states [15]. Using the polarization states of a single photon, minimum-error discrimination has been experimentally realized for up to four symmetric non-orthogonal states [16].

1.2. Optimum unambiguous discrimination

Another state discrimination strategy is known as *unambiguous discrimination*. This discrimination will always have a certain probability of failure, we can always tell whether or not the desired operation has failed. When the attempt fails, we obtain an inconclusive answer. When it succeeds, we discriminate the states with no error. This

allows us to achieve unambiguous discrimination. The optimum strategy problem means that finding the optimum solution which minimizes the average probability of failure.

This issue was first considered by Ivanovic [17], and then subsequently by Dieks [18] and Peres [19] on the case of distinguishing two non-orthogonal states. These authors found the optimal solution when the two states are being selected from ensemble in which they are equally likely. The optimal solution for the situation in which the states have different weights was found by Jaeger and Shimony [20]. Measurements of this kind have been performed in laboratory, first by Huttner, et. al. [21] and, more recently, by Clarke, *et al.* [22]. Both of these used the polarization states of photons to represent qubits. The case of three states was examined by Peres and Terno [24]. The general N-state problem has recently been addressed. Chefles [25] found that n non-orthogonal states can be probabilistic discriminated without error if and only if they are linearly independent. Chefles and Barnett [26] solved the case in which the probability of the procedure succeeding is the same for each of the states. Duan and Guo [27] considered general unitary transformations and measurements on a Hilbert space containing the states to be distinguished and an ancilla, which would allow one to discriminate among N states, and derived matrix inequalities which must be satisfied for the desired transformations to exist. In our paper [28], we presented the necessary conditions of optimum unambiguous discrimination, which can be used to derive optimum solution.

It is only recently the problem of generalization to mixed states has been consid-

ered. We were the first to solve the problem of optimal unambiguous discrimination between a pure state and a mixed state of rank two [29], which has been generalized for rank N [30]. Rudoph *et al.* solved other special cases, for example that of two mixed states of rank $(n-1)$ in a n dimensional subspace [31]. Necessary and sufficient conditions for optimality and some numerical methods are discussed by Fiurasek and Jezek [32] and by Eldar [33]. Finally, Raynal and Ltkenhaus derived theorems showed that the problem of optimal unambiguous discrimination of two general density matrices can be reduced to that of two density matrices that have the same rank n and are described in a Hilbert space of dimension $2n$ [34].

Chapter 2

General method of optimal unambiguous discrimination

2.1. Quantum measurements

The standard quantum measurement is a projective measurement, also called a von Neumann measurement [35]. If we denote the Hilbert space of the quantum system to be measured as $\mathcal{H}$ and the dimension of $\mathcal{H}$ as N, then such a measurement can be described by a set of projective operators

$$P_i = |\omega_i\rangle\langle\omega_i|, \tag{2.1}$$

where $\{|\omega_i\rangle | i = 1, \ldots, N\}$ form an orthonormal basis of $\mathcal{H}$. By this definition, P_i carry the features that

$$P_i^2 = P_i, \tag{2.2}$$

$$\sum_{i=1}^{N} P_i = \hat{1}. \tag{2.3}$$

The set of $\{P_i\}$ represents the total possible results of the measurement. If the state of the quantum system is $|\psi\rangle$, the possibility of getting result P_i is

$$p_i = \langle\psi|P_i|\psi\rangle. \tag{2.4}$$

When we take the sum over p_i, using Eq. (2.3), we have

$$\sum_{i=1}^{N} p_i = 1, \tag{2.5}$$

which confirms the theory of probability requiring that the sum of the probabilities of the possible events is 1.

Such a quantum measurement has several limitations. The first is that the eigenvalues of each P_i are either 0 or 1, no other values allowed. The second is that the number of P_i is always N, so that the number of possible results is always the same as the dimension of $\mathcal{H}$. Naturally we would ask that since a probability is just a positive number between 0 and 1, can we have a measurement represented by a set of Hermitian operators $\{\Pi_k\}$, so that each Π_k may have positive eigenvalues between 0 and 1, and the number of Π_k can be different than the dimension of $\mathcal{H}$?

The above argument can be realized by a generalized measurement [36]. A generalized measurement can be described by a set of Hermitian operators $\{\Pi_k\}$, where Π_k can be written as

$$\Pi_k = \hat{A}_k^\dagger \hat{A}_k, \tag{2.6}$$

and the number of Π_k can be denoted as M (different than N), such that each Π_k has positive eigenvalues between 0 and 1. The possibility of getting result Π_k is

$$p_k = \langle\psi|\Pi_k|\psi\rangle. \tag{2.7}$$

Similar to Eq. (2.3), the theory of probability requires that

$$\sum_{k=1}^{M} \Pi_k = \sum_{k=1}^{M} \hat{A}_k^\dagger \hat{A}_k = \hat{1}. \tag{2.8}$$

Since Π_k are positive operators, this kind of measurement is also called a *positive operator-valued measurement*, or POVM.

From the above notations, we see that generalized measurement is a much more powerful tool in the sense that both the eigenvalues and the number of the measurement operators are more flexible. The standard von Neumann measurement now becomes a special case of a generalized measurement.

As to the implementation of a generalized measurement, there is an important result, known as the Naimark theorem[37], which tells us that *any* generalized measurement can be realized with an ancillary system, a unitary operation and a von Neumann measurement. Specifically, if we wish to realism a generalized measurement with K outcomes, we need a large ancillary system. The system of interest is then made to interact unitarily with the ancilla. In general, this results in the original system and the ancilla becoming *entangled*. Following this interaction, a von Neumann measurement is performed on the ancilla. As a consequence of this entanglement, this measurement also transforms the state of our original system, and the results of this measurement give rise to the desired generalized measurement. Therefore, a generalized measurement, can be implemented by a unitary interaction with an ancilla, followed by a measurement on the latter. We will see in section 2.4 that it is the exact procedure that our optical realization scheme does.

2.2. Optimal probabilities for unambiguous discrimination among non-orthogonal quantum states

Suppose we are given a quantum system prepared in the state $|\psi\rangle$, which is guaranteed to be a member of the set of non-orthogonal states $\{|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_n\rangle\}$, but we do not know which one. We want to find a procedure which will tell us which member of the set we were given. The procedure may fail to give us any information about the state, and if it fails, it must let us know that it has, but if it succeeds, it should never give us a wrong answer. We shall refer to such a procedure as state discrimination without error. Note that this procedure has $n + 1$ outcomes; it either tells us which state we were given, or it tells us that it failed (inconclusive outcome).

In order to achieve error-free discrimination, Chefles has shown, in a very clear analysis of the problem, that the set $|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_n\rangle$ must be linearly independent [25]. If the states are not orthogonal (which we shall assume), they cannot be discriminated perfectly. That means that if we are given $|\psi_i\rangle$, we will have some probability p_i to distinguish it successfully and, correspondingly, some failure probability $q_i = 1 - p_i$ to obtain an inconclusive answer. Denote by $\mathcal{H}$ the Hilbert space spanned by the initial states $\{|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_n\rangle\}$. Since there is a chance to get an inconclusive answer, the number of outcomes of this process is larger than the dimension of $\mathcal{H}$, hence this process is a "generalized measurement" which can be represented by a

set of operators which form a resolution of the identity [25],

$$\hat{A}_I^\dagger \hat{A}_I + \sum_i \hat{A}_i^\dagger \hat{A}_i = \hat{1}, \tag{2.9}$$

where $\hat{A}_i$ is the operator that corresponds to the outcome $|\psi_i\rangle$, and $\hat{A}_I$ is the operator that corresponds to the inconclusive outcome. In more detail, if ρ is the density matrix of our given state, then the probability of obtaining the kth outcome, where k can be $1, \ldots n$ or I, is $p_k = Tr(\rho A_k^\dagger A_k)$ and if the outcome is k, then the resulting density matrix is $A_k \rho A_k^\dagger / p_k$. The requirement that the discrimination be error free implies that

$$\langle \psi_i | A_k^\dagger A_k | \psi_i \rangle = p_i \delta_{ik}. \tag{2.10}$$

From this, by an application of the Schwarz inequality, it follows that

$$\langle \psi_k | \hat{A}_i^\dagger \hat{A}_i | \psi_j \rangle = p_i \delta_{ji} \delta_{jk}, \tag{2.11}$$

If we denote by η_i the *a priori* probability that the system was prepared in the state $|\psi_i\rangle$, the average probabilities of success and of failure to distinguish the states $|\psi_i\rangle$ are, respectively,

$$\begin{aligned} P &= \sum_i \eta_i p_i, \\ Q &= \sum_i \eta_i q_i. \end{aligned} \tag{2.12}$$

Our objective is to find the set of $\{p_i\}$ that maximizes the probability of success, P, and the set of operators A_k that realize the corresponding generalized measurement.

Define the failure state, $|\phi_i\rangle$ as

$$|\phi_i\rangle = \hat{A}_I |\psi_i\rangle. \tag{2.13}$$

This is the state of the system if the input state was $|\psi_i\rangle$ and the outcome was inconclusive. Chefles [25] showed that the states $\{|\phi_i\rangle\}$ are linearly dependent when P is a maximum. The interpretation of this result is the following. Because only linearly independent states can be discriminated without error, the operator corresponding to the inconclusive outcome maps the set of linearly dependent states $\{|\psi_i\rangle\}$ onto a linearly-dependent set, which then cannot be unambiguously discriminated by any further process. As we shall see, however, this does not mean that some information cannot be extracted from an inconclusive result.

Now consider the inner product $\langle\phi_k|\phi_j\rangle$, and define the matrix C by $C_{ij} = \langle\phi_k|\phi_j\rangle$. Using equations (2.9), (2.13) and (2.11), we find

$$\langle\phi_k|\phi_j\rangle = \langle\psi_k|\psi_j\rangle - p_j\delta_{jk}. \tag{2.14}$$

The matrix C is positive-semidefinite. This can be seen by noting that for any n-dimensional vector, whose components we shall denote by b_i, where $i = 1, \ldots n$,

$$\sum_{i,j=1}^{n} b_i^* C_{ij} b_j = \| \sum_{i=1}^{n} b_i|\phi_i\rangle\|^2 \geq 0. \tag{2.15}$$

When the p_i are equal to their optimal values, i. e. the values which maximize P, the linear dependence of the $|\phi_i\rangle$ implies that

$$\det(C) = 0, \tag{2.16}$$

so that C has at least one zero eigenvalue when P is a maximum.

In conclusion, the optimum probabilities p_i can be found by maximizing P subject to the following constraints:

i) $\det(C) = 0$,

ii) C is non-negative, or, equivalently, all of the principal minors of C are non-negative.

When we consider the case of just two non-orthogonal states, the above result immediately gives the following relationship between any two failure probabilities:

$$q_1 \cdot q_2 = |\langle\psi_1|\psi_2\rangle|^2. \tag{2.17}$$

In particular, when the two states have equal *a priori* probabilities, $\eta_1 = \eta_2 = \frac{1}{2}$, we found the maximum probability of success to be,

$$P = 1 - |\langle\psi_1|\psi_2\rangle|. \tag{2.18}$$

This is the well known Ivanovic-Dieks-Peres (IDP) limit [17]-[19].

2.3. Realization of generalized measurement

Once we know the set of optimum discrimination probabilities $\{p_i\}$, we would like to find a realizable experimental procedure to achieve it. We shall do this first abstractly, and then show how it can be realized by linear optical elements. Let us first summarize the procedure, and subsequently fill in the details. We begin with a total Hilbert space $\mathcal{K}$, which is the direct sum of two subspaces, $\mathcal{K} = \mathcal{H} \oplus \mathcal{A}$. The space $\mathcal{H}$ is an n-dimensional space that contains the vectors $|\psi_i\rangle$, and $\mathcal{A}$ is the space that will contain the failure vectors $|\phi_i\rangle$. We shall denote the dimension of $\mathcal{A}$ by m. The input state of the system is one of the vectors $|\psi_i\rangle$, which is now a vector in the subspace $\mathcal{H}$ of the total space $\mathcal{K}$. A unitary transformation, U, which acts in the entire space

$\mathcal{K}$ is now applied to the input vector, resulting in the state $|\psi_i^{\mathcal{K}}\rangle_{out}$. A measurement is performed on the part of $|\psi_i^{\mathcal{K}}\rangle_{out}$ in $\mathcal{A}$, and, if the proper result is obtained, the vector $|\psi_i^{\mathcal{K}}\rangle_{out}$ is projected onto the vector $|e_i^{\mathcal{H}}\rangle$, which lies in the subspace $\mathcal{H}$. The probability of this occuring is p_i. The vectors $\{|e_i^{\mathcal{H}}\rangle, i = 1, \ldots n\}$ are orthonormal and can be distinguished perfectly. The effect of the unitary transformation on an extended space and the measurement is to map a set of non-orthogonal vectors onto a set of orthogonal ones.

We now need to specify U and the measurement, and let us discuss the latter first. The measurement has two outcomes, one of them corresponding to the operator, $P_{\mathcal{H}}$, which projects onto the subspace $\mathcal{H}$, and the other to the operator $P_{\mathcal{A}} = I - P_{\mathcal{H}}$, which projects onto the subspace $\mathcal{A}$. The first outcome corresponds to the successful transformation of $|\psi_i^{\mathcal{K}}\rangle_{out}$ into $|e_i^{\mathcal{H}}\rangle$, and its probability of occurrence is p_i. This implies that

$$|\psi_i^{\mathcal{K}}\rangle_{out} = \sqrt{p_i}|e_i^{\mathcal{H}}\rangle + |\phi_i^{\mathcal{A}}\rangle, \tag{2.19}$$

where $|\phi_i^{\mathcal{A}}\rangle$ is a failure state, and we have added a superscript $\mathcal{A}$ to denote the fact that it is in the subspace $\mathcal{A}$. The other outcome corresponds to obtaining an inconclusive answer and transforms $|\psi_i^{\mathcal{K}}\rangle_{out}$ into $|\phi_i^{\mathcal{A}}\rangle$.

Eq. (2.19) and the fact that U is unitary implies that

$$\langle\phi_k^{\mathcal{A}}|\phi_j^{\mathcal{A}}\rangle = \langle\psi_k^{\mathcal{H}}|\psi_j^{\mathcal{H}}\rangle - p_j\delta_{jk}, \tag{2.20}$$

which is just Eq. (2.14). If (and only if) the matrix C is non-negative, we can always find vectors $|\phi_i^{\mathcal{A}}\rangle$ that satisfy this equation. This follows from the fact that

a non-negative matrix can be written as the product of a matrix and its adjoint, in particular, we can express C as

$$C = A^{\dagger}A, \tag{2.21}$$

for some matrix A. If we define $|\phi_j\rangle = A|j\rangle$, where $|j\rangle$ is the vector whose jth component is one and all of whose other components are zero, then we have that $C_{jk} = \langle\phi_j|\phi_k\rangle$. Once we have found these vectors and specified the vectors $|e_i^{\mathcal{H}}\rangle$, then the operator U can be found by means of Eq. (2.19). These conditions may not completely determine U; if they do not, then there is freedom in choosing it. This will be the case if the dimension of $\mathcal{A}$ is greater than one. U maps vectors in $\mathcal{A}$ to vectors in the subspace, $\mathcal{S}$, of $\mathcal{K}$ that consists of the vectors that are orthogonal to all of the vectors $|\psi_i^{\mathcal{K}}\rangle_{out}$. The dimension of $\mathcal{S}$ is m. The freedom in choosing U comes from the fact that Eq. (2.19) does not specify how $\mathcal{A}$ is mapped into $\mathcal{S}$. If both are one-dimensional, then the mapping is determined (up to an overall phase), but if their dimension is greater than two it is not.

Once the measurement and the operator U have been specified, our realization of the generalized measurement is completely determined. The next task is to find a physical system with which to implement it.

2.4. Optical realization of non-unitary transformation

We now want to propose an experimental procedure to achieve our non-unitary transformation by using optical devices. We shall show how this can be accomplished by using a single-photon representation of the states $|\psi_i\rangle$ and an optical multiport

together with photodetectors at the output ports to carry out the desired non-unitary transformation.

Our Hilbert space will consist of a single photon, which is divided among $n+m$ modes. The modes themselves could be distinguished by having different wave vectors or they might be modes of different optical fibers. A basis for this space consists of the single photon states $\{a_j^\dagger|0\rangle|j=1,\ldots n+m\}$, where $|0\rangle$ is the vacuum state and $a_j^\dagger$ is the creation operator for the jth mode. The states $\{a_j^\dagger|0\rangle|j=1,\ldots n\}$ form a basis for the space $\mathcal{H}$, and the states $\{a_j^\dagger|0\rangle|j=n+1,\ldots n+m\}$ form a basis for the space $\mathcal{A}$. The initial states $|\psi_i\rangle$ can be represented as single photon states in $\mathcal{H}$, which can be written as

$$|\psi_i\rangle = \sum_{j=1}^{n} d_{ij}|e_j^{\mathcal{H}}\rangle = \sum_{j=1}^{n} d_{ij}\hat{a}_j^\dagger|0\rangle, \tag{2.22}$$

where we have chosen the states $|e_j^{\mathcal{H}}\rangle$ to be $|e_j^{\mathcal{H}}\rangle = \hat{a}_j^\dagger|0^{\mathcal{H}}\rangle$.

An optical $2N$-port is a lossless linear device with N input ports and N output ports. Its action on the input states can be described by a unitary operator, U_{2N}, and physically it consists of an arrangement of beam splitters, phase shifters, and mirrors. Choosing $N=n+m$, we send the single photon state $|\psi_i\rangle$ into the first n input ports, which correspond to $\mathcal{H}$, and the vacuum into the remaining m input ports, which correspond to $\mathcal{A}$. Photodetectors are placed at the last m output ports (the ones corresponding to $\mathcal{A}$), and if there is no photon detected, the desired non-unitary transformation will have been carried out. In particular, with $|\psi_i^{\mathcal{K}}\rangle_{out} = U_{2N}|\psi_i\rangle$, where $|\psi_i^{\mathcal{K}}\rangle_{out}$ is given by Eq. (2.19), the action of the measurement, if successful, is

to project the output state onto $|e_i^{\mathcal{K}}\rangle$, and the probability to achieve this is p_i.

If we denote the annihilation operators corresponding to the input modes of the $2N$-port by $a_{j\mathrm{in}}$, $j = 1, 2, \ldots, N$, then the output operators are given by

$$a_{j\mathrm{out}} = U_{2N}^{-1} a_{j\mathrm{in}} U_{2N} = \sum_{k=1}^{N} M_{jk} a_{k\mathrm{in}}, \tag{2.23}$$

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

$$|\psi^{\mathcal{K}}\rangle_{out} = U_{2N} |\psi^{\mathcal{K}}\rangle_{in}. \tag{2.24}$$

In general, for an *in* state that contains a single photon

$$|\psi^{\mathcal{K}}\rangle_{in} = \sum_{j=1}^{N} c_j a_j^{\dagger} |0\rangle, \tag{2.25}$$

where $\sum_{j=1}^{N} |c_j|^2 = 1$, the *out* state is given by

$$\begin{aligned} |\psi^{\mathcal{K}}\rangle_{out} &= U_{2N} |\psi^{\mathcal{K}}\rangle_{in} \\ &= U_{2N} \sum_{j=1}^{N} c_j a_{j\mathrm{in}}^{\dagger} U_{2N}^{-1} |0\rangle \\ &= \sum_{j,k=1}^{N} c_j M_{jk}^{T} a_{k\mathrm{in}}^{\dagger} |0\rangle. \end{aligned} \tag{2.26}$$

Note that we have made use of the fact that the vacuum is invariant under the transformation, U_{2N}. This implies that the matrix elements M_{il} is the same as the matrix element of U_{2N} between the single-particle states $|i\rangle = a_{i\mathrm{in}}^{\dagger}|0\rangle$ and $|l\rangle = a_{l\mathrm{in}}^{\dagger}|0\rangle$. Choosing $c_j = \delta_{jl}$ in the above equation and then taking the inner product of the result with $|i\rangle$, we find that

$$\langle i | U_{2N} | l \rangle = M_{il}. \tag{2.27}$$

The desired matrix M can be found from Eq. (2.19), and our next task is to decompose it in such a way that it corresponds to a linear optical network.

This problem has been solved by M. Reck *et al.* [38], and we shall summarize their method. They gave an algorithmic procedure to factorize any $N \times N$ unitary matrix into a product of two-dimensional $U(2)$ transformations, and it is this procedure that we shall adopt here to construct our $2N$-port, which is characterized by the matrix M of equation (2.23).

It is well known that a lossless beam splitter and a phase shifter with appropriate parameters can implement any $U(2)$ transformation; a beam splitter with a phase shifter at one output port transforms the input operators into output operators as

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}_{out} = \begin{pmatrix} e^{i\phi}\sin\omega & e^{i\phi}\cos\omega \\ \cos\omega & -\sin\omega \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}_{in} \tag{2.28}$$

where a_1, a_2 are the annihilation operators of modes 1 and 2 respectively. In their paper, Reck, *et al.* considered the use of a Mach-Zehnder interferometer to simulate the effect of a beam splitter that does not split the incoming beam equally, in which case ω describes the reflectivity and transmittance of the effective beam splitter with $\sqrt{R} = \sin\omega$, $\sqrt{T} = \cos\omega$, and ϕ describes the effect of the phase shifter. If the matrix describes an actual beam splitter, then $\sqrt{R} = \cos\omega$, and $\sqrt{T} = \sin\omega$. Any $N \times N$ unitary matrix $U(N)$ can be reduced to an $(N-1) \times (N-1)$ unitary matrix, $U(N-1)$, by multiplying from the right by a succession of two-dimensional unitary

matrices

$$U(N)\cdot R(1)=\begin{pmatrix} e^{i\alpha_1} & 0 \\ 0 & U(N-1) \end{pmatrix}. \tag{2.29}$$

Here $R(1)=T_{1,2}\cdot T_{1,3}\cdots T_{1,N}$, and $T_{p,q}$ is defined as an N-dimensional identity matrix with elements I_{pp}, I_{pq}, I_{qp}, I_{qq} replaced by the corresponding elements of a $U(2)$ matrix. It performs a unitary transformation on a two-dimensional subspace of the full N dimensional space, and can be implemented by attaching a beam splitter and a phase shifter to ports p and q.

We can repeat the above transformation, decreasing the dimension of the remaining unitary matrix by one at each step. Applying this procedure to the matrix M of equation (2.23), we have that

$$\begin{aligned} M \cdot R(1)\cdot R(2)\cdots R(n+m) \\ =\begin{pmatrix} e^{i\alpha_1} & 0 & & 0 \\ 0 & e^{i\alpha_2} & & \\ 0 & 0 & \ddots & \\ 0 & 0 & & e^{i\alpha_{n+m}} \end{pmatrix}. \end{aligned} \tag{2.30}$$

Denoting by $D(\alpha_1,\alpha_2\ldots\alpha_n)$ the diagonal matrix

$$D=\begin{pmatrix} e^{-i\alpha_1} & & 0 & \\ & e^{-i\alpha_2} & & \\ 0 & & \ddots & \\ & & & e^{-i\alpha_{n+m}} \end{pmatrix}, \tag{2.31}$$

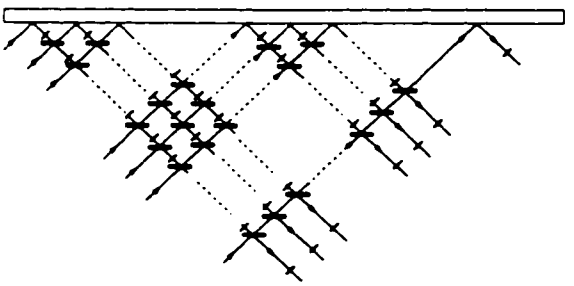

FIG. 1: The implementation of an optical multiport that performs the unitary transformation $M(N)$ described in Eq. (2.33). The beams are straight lines, a suitable beam splitter is at each crossing point of the beams, phase shifters are at one input of each beam splitter and at the outputs of the multiport. Each diagonal line of the multiport reduces the dimension of $M(N)$ by one.

we have

$$M \cdot R(1) \cdot R(2) \cdots R(n+m-1) \cdot D = \mathbf{1}, \tag{2.32}$$

i.e.,

$$M = D^{-1} \cdot R(n+m-1)^{-1} \cdots R(1)^{-1}. \tag{2.33}$$

Since the product of matrices is equivalent to setting up experimental devices in sequence, Eq. (2.33) implies that to get M, the actual experimental setup is made of a series of $U(2)$ blocks to achieve $R(n+m)^{-1} \cdots R(1)^{-1}$, and $n+m$ appropriate phase shifters attached to the output ports to produce D^{-1}. Figure 1 gives a picture of the practical implementation of M.

It is possible to save some steps by modifying this procedure. As was mentioned earlier, the matrix M is not always completely determined. In particular, there is freedom in choosing the matrix elements M_{jk} for $k > n$. Let us now see what happens if we apply the procedure of Reck, et al. to the transpose of M, M^T, instead of M

itself. It is now the matrix elements $(M^T)_{jk}$, for $j > n$ that are not completely determined, and we shall leave them that way for now. The matrixes making up $R(1)$ are chosen to make all of the elements, except the first, of the first row of $M^T R(1)$ zero. In finding each of the matrixes $T_{1,q}$, we only need to use the matrix elements that are completely determined (this is not true if we start with M instead of M^T). Now if the first row of $M^T R(1)$ is zero except for the first element, then by unitarity, the first column is also zero, except for its first element. Continuing in this way we have that

$$\begin{aligned} M^T \cdot R(1) \cdot R(2) \cdots R(n) \\ = \begin{pmatrix} e^{i\alpha_1} & 0 & 0 \\ 0 & e^{i\alpha_2} & \\ 0 & 0 & \ddots \\ 0 & 0 & M_m \end{pmatrix}, \end{aligned} \tag{2.34}$$

where M_m is an $m \times m$ unitary matrix that contains the information about the matrix elements in M that are not completely specified. At this point, we can choose M_m to be any unitary matrix, and the simplest choice is the $m \times m$ identity matrix, I_m. Defining

$$D' = \begin{pmatrix} e^{-i\alpha_1} & 0 & 0 \\ 0 & e^{-i\alpha_2} & \\ 0 & 0 & \ddots \\ 0 & 0 & I_m \end{pmatrix}, \tag{2.35}$$

we have that

$$M = [R(1)R(2)\cdots R(n)D']^{*}. \tag{2.36}$$

Chapter 3

Application to two states and examples

3.1. Dual-rail representation

Let $a_1^\dagger$ and $a_2^\dagger$ be creation operators corresponding to two different modes of the electromagnetic field. We shall denote the two-mode vacuum state by $|00\rangle$. If we consider only states in which the total number of photons in both modes is one, then our Hilbert space is spanned by the states

$$|10\rangle = a_1^\dagger|00\rangle \qquad |01\rangle = a_2^\dagger|00\rangle. \tag{3.1}$$

We shall identify the state $|10\rangle$ with the qubit state $|0\rangle$, and the state $|01\rangle$ with the qubit state $|1\rangle$. The general qubit state is then

$$\begin{aligned} \alpha|0\rangle + \beta|1\rangle &= (\alpha a_1^\dagger + \beta a_2^\dagger)|00\rangle \\ &= \alpha|10\rangle + \beta|01\rangle. \end{aligned} \tag{3.2}$$

Such a qubit state could be produced by sending a single photon into a beam splitter with the proper transmission and reflection coefficients, where the modes 1 and 2 would correspond to the output modes of the beam splitter. This representation of a qubit is the so-called dual-rail representation proposed by Milburn [39], and later

by Chuang and Yamamoto [40, 41]. It is a special case of the representation that we proposed in Chapter 2.4. It is this representation we shall employ here.

3.2. Optimal discrimination between two non-orthogonal states

In this section we apply the above considerations to the problem of realizing optimal discrimination between two non-orthogonal but linearly-independent quantum states [5, 42].

From Eq. (2.12), the probability of failure is

$$Q = \eta_1 q_1 + \eta_2 q_2. \tag{3.3}$$

The requirement of the linear dependence of the $|\phi_i\rangle$ vectors ($i = 1, 2$) leads to the constraint given by Eq. (2.16) as

$$\Delta = q_1 q_2 - |O_{12}|^2 = 0,$$

where $O_{12} = \langle\psi_1|\psi_2\rangle$.

Employing the Lagrange multiplier method, we wish to minimize the quantity

$$Q' = q_1 q_2 - |O_{12}|^2 + \lambda\Delta, \tag{3.5}$$

which immediately leads to the conditions

$$\begin{aligned} \frac{\partial Q'}{\partial q_1} &= \eta_1 + \lambda q_2 = 0, \\ \frac{\partial Q'}{\partial q_2} &= \eta_2 + \lambda q_1 = 0, \end{aligned} \tag{3.6}$$

where λ is a Lagrange multiplier.

By using Eq. (3.4), (3.6) we can solve λ, q_1 and q_2, by keeping in mind that q_1 and q_2 must be between 0 and 1, we finally can get the following conclusion:

(i) if $\frac{|O_{12}|^2}{1+|O_{12}|^2} \leqslant \eta_1 \leqslant \frac{1}{1+|O_{12}|^2}$, then

$$\begin{aligned} q_1 &= \sqrt{\frac{\eta_2}{\eta_1}}|O_{12}|, \\ q_2 &= \sqrt{\frac{\eta_1}{\eta_2}}|O_{12}|, \\ Q &= 2\sqrt{\eta_1\eta_2}|O_{12}|. \end{aligned} \tag{3.7}$$

(ii) if $\eta_1 < \frac{|O_{12}|^2}{1+|O_{12}|^2}$, then

$$\begin{aligned} q_1 &= 1, \\ q_2 &= |O_{12}|^2, \\ Q &= \eta_1 + \eta_2|O_{12}|^2. \end{aligned} \tag{3.8}$$

This is a standard Von Neumann measurement with $\Pi_1 = 0$, $\Pi_2 = |\psi_1^\perp\rangle\langle\psi_1^\perp|$, $\Pi_I = |\psi_1\rangle\langle\psi_1|$.

(iii) if $\eta_1 > \frac{1}{1+|O_{12}|^2}$, then

$$\begin{aligned} q_1 &= |O_{12}|^2, \\ q_2 &= 1, \\ Q &= \eta_2 + \eta_1|O_{12}|^2. \end{aligned} \tag{3.9}$$

This is a standard Von Neumann measurement with $\Pi_1 = |\psi_2^\perp\rangle\langle\psi_2^\perp|$, $\Pi_2 = 0$, $\Pi_I = |\psi_2\rangle\langle\psi_2|$.

Therefore, to unambiguously distinguish two non-orthogonal states, the optimal measurement is a POVM bounded by standard Von Neumann measurement (See Figure 2).

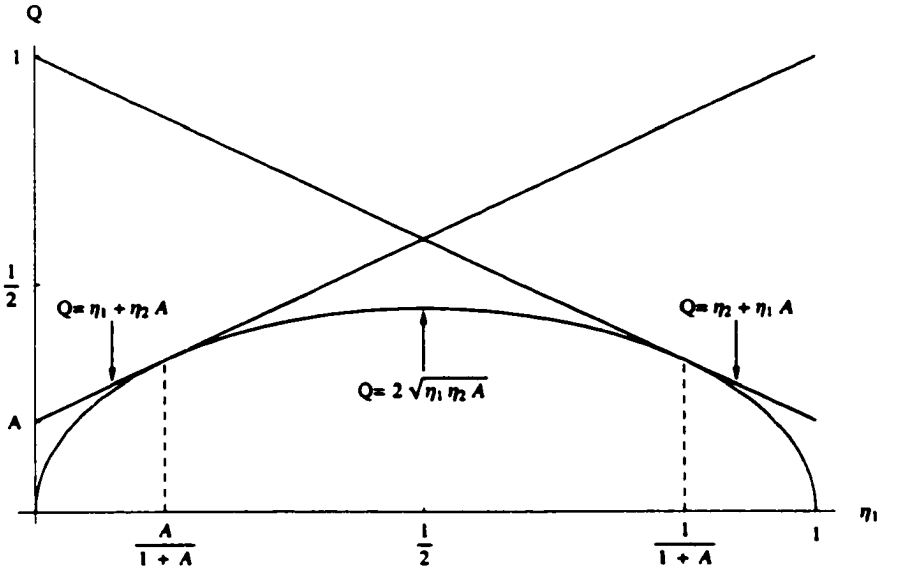

FIG. 2: Let $A = |O_{12}|^2$, when $\frac{A}{1+A} \leqslant \eta_1 \leqslant \frac{1}{1+A}$, the optimal measurement is a POVM. When η_1 is outside the boundaries, the optimal measurement is a standard Von Neumann measurement. The failure probability of the POVM is always less than the Von Neumann measurement, when $\eta_1 = \eta_2 = 1/2$, the difference between the two types of measurement is at the largest.

In particular, when the two states have equal *a priori* probabilities, $\eta_1 = \eta_2 = \frac{1}{2}$, we found that

$$q_1 = q_2 = Q = |O_{12}|. \tag{3.10}$$

This is the well known Ivanovic-Dieks-Peres (IDP) limit [17]–[19].

Now let us consider two particular input states,

$$|\psi_1\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle = \cos\theta|10\rangle_{12} + \sin\theta|01\rangle_{12}$$
$$|\psi_2\rangle = |0\rangle = |10\rangle_{12}. \tag{3.11}$$

The inner product of these two states is $\langle\psi_1|\psi_2\rangle = \cos\theta$, where $0 \leqslant \theta \leqslant \pi/2$.

We now want to use a six-port to perform nonunitary operations on a single qubit, with a certain probability of success. We send our qubit, in the dual-rail representation, into inputs 1 and 2 of the six-port, and the vacuum into the third input. The first two output ports are where we want the qubit to come out, and a detector is placed at the third output port. We only accept the output in the first two output ports if the detector does not fire.

The input state of the system can be written explicitly as

$$|\psi_1\rangle_{in} = \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix}, \quad |\psi_2\rangle_{in} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{3.12}$$

When $\eta_1 = \eta_2 = \frac{1}{2}$, from Eq. (3.10)

$$q_1 = q_2 = Q = \cos\theta, \tag{3.13}$$

which implies that $|\phi_i\rangle = (\sqrt{q_i})$ are

$$|\phi_1\rangle = \left(\sqrt{\cos\theta}\right), |\phi_2\rangle = \left(\sqrt{\cos\theta}\right). \tag{3.14}$$

$|\psi_i\rangle_{out}$ can be chosen to be

$$|\psi_1\rangle_{out} = \begin{pmatrix} \sqrt{2}\sin\frac{\theta}{2} \\ 0 \\ \sqrt{\cos\theta} \end{pmatrix}, |\psi_2\rangle_{out} = \begin{pmatrix} 0 \\ \sqrt{2}\sin\frac{\theta}{2} \\ \sqrt{\cos\theta} \end{pmatrix}. \tag{3.15}$$

The unitary transformation, U, maps the input states onto the output states, i.e. $|\psi_i\rangle_{out} = U|\psi_i\rangle_{in}$, for $i = 1, 2$. In addition, it must map the vector that is orthogonal to the two input vectors onto the vector that is orthogonal to the two output vectors,

$$\begin{pmatrix} \frac{\sqrt{\cos(\theta)}}{\sqrt{1+\cos(\theta)}} \\ \frac{\sqrt{\cos(\theta)}}{\sqrt{1+\cos(\theta)}} \\ -\frac{\sqrt{2}\sin(\frac{\theta}{2})}{\sqrt{1+\cos(\theta)}} \end{pmatrix} = U \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{3.16}$$

The action of U on these three vectors completely determines it, and we find that it is given by the matrix $M(3)$, which is

$$M(3) = \begin{pmatrix} 0 & \frac{\sec(\frac{\theta}{2})}{\sqrt{2}} & \sqrt{\frac{\cos(\theta)}{1+\cos(\theta)}} \\ \sqrt{2}\sin(\frac{\theta}{2}) & -\sqrt{2}\cot(\theta)\sin(\frac{\theta}{2}) & \sqrt{\frac{\cos(\theta)}{1+\cos(\theta)}} \\ \sqrt{\cos(\theta)} & \sqrt{\cos(\theta)}\tan(\frac{\theta}{2}) & -\frac{\sin(\frac{\theta}{2})}{\sqrt{\cos(\frac{\theta}{2})^2}} \end{pmatrix}. \tag{3.17}$$

$M(3)$ describes the action of a six-port which transforms the input states into output states. Using the method described in Sec. 2.4, $M(3)$ can be factorized as

$$M(3) = T_{1,2} \cdot T_{1,3} \cdot T_{2,3}, \tag{3.18}$$

where

$$T_{1,2} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{3.19}$$

$$T_{1,3} = \begin{pmatrix} \sqrt{1-\cos(\theta)} & 0 & \sqrt{\cos(\theta)} \\ 0 & 1 & 0 \\ \sqrt{\cos(\theta)} & 0 & -\sqrt{1-\cos(\theta)} \end{pmatrix}, \tag{3.20}$$

and

$$T_{2,3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sec(\frac{\theta}{2})}{\sqrt{2}} & \sqrt{\frac{\cos(\theta)}{1+\cos(\theta)}} \\ 0 & -\frac{\sqrt{\cos(\theta)}\sec(\frac{\theta}{2})}{\sqrt{2}} & \frac{\sec(\frac{\theta}{2})}{\sqrt{2}} \end{pmatrix}. \tag{3.21}$$

$T_{1,2}$ simply means to switch mode 1 and mode 2, each of $T_{1,3}$ and $T_{2,3}$ represents the action of a beam splitter with the transmission coefficient given by the diagonal matrix elements in the 2×2 submatrix, and the reflection coefficient given by the positive off-diagonal matrix element in the same submatrix. This implies that the six-port represented by $M(3)$ can be constructed from the two beam splitters represented by $T_{1,3}$, $T_{2,3}$ and a mirror (see Fig. 1).

3.3. Entanglement Enhancement

Entanglement shared between two parties is resource which can be used both to teleport a quantum state from one party to another or to enhance the rate at which bits can be sent between them [4, 6]. Consequently, it is useful to find ways of processing ensembles of two party states in order to concentrate the entanglement they contain [43]. Here we shall explore several simple, if not particularly efficient, ways of doing this. The first is the "Procrustean" method developed by Bennett, et. al. [43]. In order to use this method one must know the initial state. The second is similar in spirit to the "Procrustean" method. It has the advantage that the same

transformation can be used for a wide class of states making it useful for both pure and mixed ensembles, but the disadvantage that it does not produce maximally entangled states. Finally, we shall show how the transformation which takes nonorthogonal to orthogonal states can be used to increase entanglement.

In order to apply the "Procrustean" method one starts with the two-particle state, where Alice has one particle and Bob the other,

$$|\Psi\rangle = \cos\theta|0\rangle_A|1\rangle_B + \sin\theta|1\rangle_A|0\rangle_B, \tag{3.22}$$

where we shall assume that $\theta \leq \pi/4$. Alice now wants to send her particle through a device which does nothing to particles in the $|1\rangle$ state, but transmits particles in the $|0\rangle$ state with an amplitude of $\tan\theta$ and absorbs them with an amplitude of $\sqrt{1-\tan^2\theta}$. The result, with probability $\tan^2\theta$, is the maximally entangled state $|0\rangle_A|1\rangle_B + |1\rangle_A|0\rangle_B$.

If Alice's and Bob's particles are photons in the dual-rail representation, then this procedure is simple to carry out. Let Alice's photon be in modes 1 and 2 and Bob's be in modes 3 and 4. Alice now inserts a beam splitter into mode 1 (a photon in mode 1 corresponds to the state $|0\rangle$) which couples this mode to an auxiliary mode, mode 5. The beam splitter transmits an incoming photon in mode 1 with an amplitude of $\tan\theta$ and reflects it with an amplitude of $\sqrt{1-\tan^2\theta}$. A photon in mode 2 (state $|1\rangle$) is unaffected. If no photon is detected in mode 5, then the transformation has succeeded, and Alice and Bob share a maximally entangled state.

We now consider a second method of entanglement enhancement which is based

on the following transformation:

$$\alpha|0\rangle + \beta|1\rangle \rightarrow \frac{1}{(|\alpha|^2 + T|\beta|^2)^{1/2}}(\alpha|0\rangle + \sqrt{T}\beta|1\rangle). \tag{3.23}$$

It is nonunitary and has the effect of increasing the $|0\rangle$ component of a qubit and decreasing the $|1\rangle$ component. Note that the vectors $|0\rangle$ and $|1\rangle$ are left unchanged by this transformation. This means that if we consider the two qubits $|1\rangle$ and $\alpha|0\rangle + \beta|1\rangle$, then the above transformation will bring them closer to being orthogonal by decreasing the $|1\rangle$ component of the second qubit.

The six port which realizes this transformation is quite simple (see Fig. 2). The first input goes straight through and the second two are mixed by a beam splitter with transmission coefficient T and reflection coefficient R ($R + T = 1$), giving the transformation

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{T} & \sqrt{R} \\ 0 & -\sqrt{R} & \sqrt{T} \end{pmatrix}. \tag{3.24}$$

The state $|in\rangle$ is transformed by this six port into the state

$$|out\rangle = [\alpha a_1^\dagger + \beta(\sqrt{T}a_2^\dagger - \sqrt{R}a_3^\dagger)]|000\rangle. \tag{3.25}$$

After measuring mode 3 and finding no photon present the output state becomes, up to normalization

$$(\alpha a_1^\dagger + \beta\sqrt{T}a_2^\dagger)|000\rangle, \tag{3.26}$$

and the probability that this transformation will fail is just

$$p = |\beta|^2 R. \tag{3.27}$$

Clearly there is a trade off between how much the state is changed, in particular how much its $|1\rangle$ component shrinks, and the likelihood of successfully accomplishing the transformation; the bigger one is the smaller the other will be.

This transformation can be used to increase the amount of entanglement in certain kinds of two-qubit states. Consider the state

$$|\Phi\rangle = \frac{1}{\sqrt{2}(1+|\beta|^2)^{1/2}}(|1\rangle_A|\phi\rangle_B + |\phi\rangle_A|1\rangle_B), \tag{3.28}$$

where $|\phi\rangle = \alpha|0\rangle + \beta|1\rangle$. A transformation which would bring $|\phi\rangle$ closer to $|0\rangle$ and leave $|1\rangle$ unaffected would bring $|\Phi\rangle$ closer to a maximally entangled state and, thereby, increase the entanglement of the state. This is exactly the kind of transformation which was discussed in the previous two paragraphs.

We now consider a six mode system where modes 1 and 2 constitute the first qubit, modes 4 and 5 the second, and modes 3 and 6 are initially in their respective vacuum states and are subsequently measured. We first perform the above procedure on the first qubit (modes 1 and 2) with a probability of success (the probability of finding no photon in mode 3) given by

$$p_{s1} = \frac{1+T|\beta|^2}{1+|\beta|^2} - \frac{R}{2}. \tag{3.29}$$

We then apply the procedure to the second qubit where the success probability is now

$$p_{s2} = \frac{2T|\alpha|^2 + 4T^2|\beta|^2}{|\alpha|^2(1+T) + 4T|\beta|^2}, \tag{3.30}$$

and the final state becomes (up to normalization)

$$|\Phi^{out}\rangle = |1\rangle_A|\phi'\rangle_B + |\phi'\rangle_A|1\rangle_B, \tag{3.31}$$

where

$$|\phi'\rangle = \alpha|0\rangle + \beta\sqrt{T}|1\rangle. \tag{3.32}$$

This is, in fact, the desired result, and the probability of successfully accomplishing this transformation is $p_{s1}p_{s2}$.

Note that this procedure will increase the entanglement of a state of the form given in Eq. (3.28) for any nonzero values of α and β. This means that it can be used to increase the amount of entanglement present in ensembles as well. If the ensemble is described by the density matrix

$$\rho = \sum_j p_j |\Phi_j\rangle\langle\Phi_j|, \tag{3.33}$$

where $|\Phi_j\rangle$ is a state of the form given in Eq. (3.28), and p_j is the probability of the occurrence of $|\Phi_j\rangle$, then the entanglement present in this can be increased by performing the above nonunitary operations on its members.

Finally, let us show how the transformation discussed in section 3 can be used to produce maximally entangled states. Suppose Alice and Bob share a supply of particles in the state

$$|\Gamma\rangle = \left(\frac{1}{2 + 2\cos^2\theta}\right)^{1/2} (|0\rangle_A|\phi\rangle_B + |\phi\rangle_A|0\rangle_b), \tag{3.34}$$

where now

$$|\phi\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle. \tag{3.35}$$

Note that $|\Gamma\rangle$ is composed of the two nonorthogonal states considered in section 3. Alice and Bob now send their respective photons through six ports represented by

the matrix in Eq. (3.17) and accept the output only if they do not detect a photon in the third output port. The probability that this procedure will succeed, p_s, is the product of Alice's success probability and Bob's, and it is given by

$$p_s = \frac{(1-\cos\theta)^2}{1+\cos^2\theta}. \tag{3.36}$$

The state which results is

$$|\Gamma'\rangle = \frac{1}{\sqrt{2}}(|+\rangle_A|-\rangle_B + |-\rangle_A|+\rangle_B), \tag{3.37}$$

where $|\pm\rangle = (\pm|0\rangle + |1\rangle)/\sqrt{2}$, and it is maximally entangled.

Chapter 4

Application to three states and examples

In this section we first apply the above considerations to the problem of realizing optimal discrimination among three non-orthogonal but linearly-independent quantum states, in general. Then we illustrate the method on specific examples. For simplicity, we assume that the *a priori* probabilities are all equal, $\eta_1 = \eta_2 = \eta_3 = 1/3$.

From Eq. (2.12), the probability of failure is

$$Q = \frac{1}{3}\sum_{i=1}^{3} q_i. \tag{4.1}$$

The requirement of the linear dependence of the $|\phi_i\rangle$ vectors ($i = 1, 2, 3$) leads to the constraint given by Eq. (2.16). For the case of three vectors it can be written as

$$\begin{aligned} \Delta &= det(C) \\ &= q_1 q_2 q_3 - q_1|O_{23}|^2 - q_2|O_{13}|^2 - q_3|O_{12}|^2 \\ &\quad + O_{12}O_{23}O_{13}^* + O_{12}^* O_{23}^* O_{13} \\ &= 0, \end{aligned} \tag{4.2}$$

where $O_{ij} = \langle\psi_i|\psi_j\rangle$.

Employing the Lagrange multiplier method, we wish to minimize the quantity

$$Q' = \frac{1}{3}\sum_i q_i + \lambda\Delta, \tag{4.3}$$

which immediately leads to the conditions

$$\begin{aligned}
\frac{\partial Q'}{\partial q_1} &= \frac{1}{3} + \lambda\Delta_{23} = 0, \\
\frac{\partial Q'}{\partial q_2} &= \frac{1}{3} + \lambda\Delta_{13} = 0, \\
\frac{\partial Q'}{\partial q_3} &= \frac{1}{3} + \lambda\Delta_{12} = 0,
\end{aligned} \tag{4.4}$$

where λ is a Lagrange multiplier, and Δ_{12}, Δ_{13}, Δ_{23} are subdeterminants of C, $\Delta_{12} = q_1q_2 - |O_{12}|^2$, etc. Equation (4.4) implies that

$$\Delta_{12} = \Delta_{13} = \Delta_{23} = -\frac{1}{3\lambda}. \tag{4.5}$$

This means that all three subdeterminants are equal. Let $\delta = -\frac{1}{3\lambda}$ denote this common value and recall that all subdeterminants of C must be non-negative, so that $\delta \geq 0$.

From Eq. (4.5) we can solve for the q_i's, yielding

$$\begin{aligned}
q_1 &= \sqrt{\frac{(|O_{12}|^2+\delta)(|O_{13}|^2+\delta)}{(|O_{23}|^2+\delta)}}, \\
q_2 &= \sqrt{\frac{(|O_{12}|^2+\delta)(|O_{23}|^2+\delta)}{(|O_{13}|^2+\delta)}}, \\
q_3 &= \sqrt{\frac{(|O_{13}|^2+\delta)(|O_{23}|^2+\delta)}{(|O_{12}|^2+\delta)}}.
\end{aligned} \tag{4.6}$$

Finally, we can substitute Eq. (4.6) into Eq. (4.2) to solve for δ and then use the above equations to find the corresponding q_i values. When we solve for δ, there are

often a number of different solutions. However, we need only consider solutions that are greater than or equal to zero, and which give values of q_i that are between 0 and 1. If there are several solutions that satisfy these conditions, we must determine which one gives the actual minimum. If there are none, then we must examine the boundary of the allowed region to find the minimum. The point (q_1, q_2, q_3) lies inside or on the surface of a unit cube one whose vertices lie on the points (j, k, l), where $j, k, l = 0$ or 1. If the Lagrange multiplier approach does not yield a valid solution the minimum of Q subject to the constraint $\Delta = 0$ must lie on the surface of the cube.

Note that if the overlaps are real and positive, a situation we shall consider shortly, then $\delta = 0$ is always a solution of Eq. (4.2). In this case, if all the corresponding q_i for $\delta = 0$ are between 0 and 1, then this set of $\{q_i\}$ is a possible solution to our problem, i.e. a minimum of Q that satisfies $\Delta = 0$. If it is, in fact the solution, we see that $\Delta_{12} = \Delta_{13} = \Delta_{23} = 0$, which implies that each possible pair of the states $|\phi_i\rangle$, $i = 1, 2, 3$ is linearly dependent, so that all three states ϕ_i are in a line, i.e. the dimensionality of the auxiliary Hilbert space $\mathcal{A}$ is one. If the solution to the problem is one for which $\delta > 0$, no pair of failure states is linearly dependent. However, the three failure states together are linearly dependent, so that in this case the dimensionality of the auxiliary Hilbert space $\mathcal{A}$ is two.

Next, we shall consider specific examples involving three non-orthogonal but linearly independent state vectors, to illustrate the general considerations of the previous sections. In particular we want to determine explicitly the parameters and dimensionality for the special multiports that optimally discriminate among the three quantum

states. For simplicity, we shall assume that the *a priori* probabilities are equal in all of our examples.

Our first case is a simple one; the overlaps of the three states will be assumed to be real and equal

$$\langle\psi_1|\psi_2\rangle = \langle\psi_2|\psi_3\rangle = \langle\psi_3|\psi_1\rangle = s, \tag{4.7}$$

where $0 < s < 1$. The constraint of equation (2.16) is, in this case,

$$q_1 q_2 q_3 - s^2 \sum_i q_i + 2s^3 = 0, \tag{4.8}$$

application of the Lagrangian multiplier method implies that $q_1 = q_2 = q_3$, and that

$$q_i^3 - 3s^2 q_i + 2s^3 = 0. \tag{4.9}$$

This equation has two solutions, $q_i = s, -2s$, of which only $q_i = s$ is valid. This solution is a minimum and it implies that the optimal value of the total failure probability is $Q = s$.

Our next step is to find the failure vectors. For any 3×3 positive matrix, L, we find that we can express its matrix elements as $L_{ij} = \langle\phi_i|\phi_j\rangle$ if

$$\begin{aligned} |\phi_1\rangle &= (\sqrt{L_{11}}, 0, 0) \\ |\phi_2\rangle &= \left(\frac{L_{12}}{\sqrt{L_{11}}}, \sqrt{\frac{\Delta_{12}}{L_{11}}}, 0\right) \\ |\phi_3\rangle &= \left(\frac{L_{13}}{\sqrt{L_{11}}}, \frac{L_{23}L_{11} - L_{12}^* L_{13}}{\sqrt{L_{11}\Delta_{12}}}, \sqrt{\frac{\Delta}{\Delta_{12}}}\right), \end{aligned} \tag{4.10}$$

where $\Delta_{12} = L_{11}L_{22} - |L_{12}|^2$ and $\Delta = \det L$. Applying this to the matrix C, with $q_i = s$, $i = 1, 2, 3$, we find that the three failure vectors are identical, they all have

magnitude $\sqrt{s}$ and point in the same direction. Therefore, our failure space, $\mathcal{A}$, is one dimensional, the full Hilbert space $\mathcal{K} = \mathcal{H} \oplus \mathcal{A}$ is four dimensional, and we will need an eight port to accomplish our unitary transformation.

In order to find the necessary unitary transformation, we must first specify our input states. Let us choose our three states to be (in the full space, $\mathcal{K} = \mathcal{H} \oplus \mathcal{A}$)

$$|\psi_1^{\mathcal{K}}\rangle_{in} = \begin{pmatrix} \sqrt{\frac{2}{3}}\sqrt{1-s} \\ \frac{\sqrt{1+2s}}{\sqrt{3}} \\ 0 \\ 0 \end{pmatrix}, \quad |\psi_2^{\mathcal{K}}\rangle_{in} = \begin{pmatrix} -\frac{\sqrt{1-s}}{\sqrt{6}} \\ \frac{\sqrt{1+2s}}{\sqrt{3}} \\ \frac{\sqrt{1-s}}{\sqrt{2}} \\ 0 \end{pmatrix},$$

$$|\psi_3^{\mathcal{K}}\rangle_{in} = \begin{pmatrix} -\frac{\sqrt{1-s}}{\sqrt{6}} \\ \frac{\sqrt{1+2s}}{\sqrt{3}} \\ -\frac{\sqrt{1-s}}{\sqrt{2}} \\ 0 \end{pmatrix}, \tag{4.11}$$

where ψ_i are represented by single photon states. One can verify that $\langle\psi_1|\psi_2\rangle = \langle\psi_2|\psi_3\rangle = \langle\psi_3|\psi_1\rangle = s$. The output states can be found from Eq. (2.19), and are

explicitly given by

$$|\psi_1^{\mathcal{K}}\rangle_{out} = \begin{pmatrix} \sqrt{1-s} \\ 0 \\ 0 \\ \sqrt{s} \end{pmatrix}, \quad |\psi_2^{\mathcal{K}}\rangle_{out} = \begin{pmatrix} 0 \\ \sqrt{1-s} \\ 0 \\ \sqrt{s} \end{pmatrix},$$

$$|\psi_3^{\mathcal{K}}\rangle_{out} = \begin{pmatrix} 0 \\ 0 \\ \sqrt{1-s} \\ \sqrt{s} \end{pmatrix}. \tag{4.12}$$

The unitary transformation, U, maps the input states onto the output states, i.e. $|\psi_i^{\mathcal{K}}\rangle_{out} = U|\psi_i^{\mathcal{K}}\rangle_{in}$, for $i = 1, 2, 3$. In addition, it must map the vector that is orthogonal to the three input vectors onto the vector that is orthogonal to the three output vectors,

$$\frac{1}{\sqrt{2s+1}} \begin{pmatrix} \sqrt{s} \\ \sqrt{s} \\ \sqrt{s} \\ -\sqrt{1-s} \end{pmatrix} = U \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{4.13}$$

The action of U on these four vectors completely determines it, and we find that it

TABLE 1: Arguments of $T_{p,q}$ for case 1. The range of arccos is the interval $[0, \pi]$.

	ω	ϕ
$T_{1,2}$	$\arccos[-\frac{1}{\sqrt{5}}]$	0
$T_{1,3}$	$\arccos[-\frac{1}{\sqrt{6}}]$	0
$T_{2,3}$	$-\arccos[-\sqrt{\frac{2}{5}}]$	0
$T_{2,4}$	$\arccos[\frac{\sqrt{3s}}{\sqrt{1+2s}}]$	0

is given by the matrix $M(4)$, which is

$$M(4) = \begin{pmatrix} \sqrt{\frac{2}{3}} & \sqrt{\frac{1-s}{3(2s+1)}} & 0 & \sqrt{\frac{s}{2s+1}} \\ -\frac{1}{\sqrt{6}} & \sqrt{\frac{1-s}{3(2s+1)}} & \frac{1}{\sqrt{2}} & \sqrt{\frac{s}{2s+1}} \\ -\frac{1}{\sqrt{6}} & \sqrt{\frac{1-s}{3(2s+1)}} & -\frac{1}{\sqrt{2}} & \sqrt{\frac{s}{2s+1}} \\ 0 & \sqrt{\frac{3s}{2s+1}} & 0 & -\sqrt{\frac{1-s}{2s+1}} \end{pmatrix}. \tag{4.14}$$

Using the method described in Sec. IV, $M(4)$ can be factorized as

$$M(4) = T_{1,2} \cdot T_{1,3} \cdot T_{2,3} \cdot T_{2,4}, \tag{4.15}$$

where the parameters that determine the matrixes T_{pq} are given in table 1 (this example is referred to as case 1). Note that because these matrixes are real, the complex conjugate, which appears in Eq. (2.36) is unnecessary.

Now let us consider a more general case than the one we have been studying so far. We shall assume that two of the overlaps are the same and the third is different,

in particular that

$$\langle\psi_1|\psi_2\rangle = \langle\psi_1|\psi_3\rangle = s_1$$
$$\langle\psi_2|\psi_3\rangle = \quad s_2 \quad , \tag{4.16}$$

where we shall assume, for simplicity, that s_1 and s_2 are real and between 0 and 1. For a fixed value of s_1 there is a restriction on how large s_2 can be. The largest the angle between ψ_2 and ψ_3 can be is twice the angle between ψ_1 and ψ_2 (this maximum is achieved when the vectors are coplanar). This implies that $s_2 \geq 2s_1^2 - 1$. Application of the Lagrange multiplier method to the minimization of Q' gives us $q_2 = q_3$ and

$$q_1 = \frac{q_2^2 + s_1^2 - s_2^2}{q_2}. \tag{4.17}$$

Substituting these results into the constraint equation and defining $y = q_2/s_2$ and $\beta = s_1/s_2$, we have

$$y^4 - (2+\beta^2)y^2 + 2\beta^2 y + 1 - \beta^2 = 0. \tag{4.18}$$

The roots of this equation are $y = 1,\ -1 \pm \beta$, and two of them 1 and $\beta - 1$ yield valid solutions, the latter if $\beta \geq 1$. Substitution of these results into Q shows that if $\beta < 2$, then the solution $y = 1$ gives the minimum and if $\beta \geq 2$, then $y = \beta - 1$ gives the minimum. Summarizing, we find that if $\beta < 2$, the minimum value of Q is $[(s_1^2/s_2) + 2s_2]/3$ and (solution 1)

$$q_1 = \frac{s_1^2}{s_2}$$
$$q_2 = q_3 = s_2, \tag{4.19}$$

and if $\beta \geq 2$, then the minimum value of Q is $2(2s_1 - s_2)/3$ and (solution 2)

$$\begin{aligned} q_1 &= 2s_1 \\ q_2 &= q_3 = s_1 - s_2. \end{aligned} \tag{4.20}$$

Clearly, for these solutions to be valid, all of the probabilities have to be between 0 and 1.

The next step is to find the failure vectors. If solution 1 is the valid one, we find from Eq. (4.10) that the failure space is one dimensional, and if $|u_1^A\rangle$ is the normalized basis vector for this space, then

$$\begin{aligned} |\phi_1^A\rangle &= \frac{s_1}{\sqrt{s_2}}|u_1^A\rangle, \\ |\phi_2^A\rangle &= |\phi_3^A\rangle = \sqrt{s_2}|u_1^A\rangle. \end{aligned} \tag{4.21}$$

If solutions 2 is the valid one, then the failure space is two-dimensional. If $|u_j^A\rangle$ where $j = 1, 2$ is an orthonormal basis for this space we find that

$$\begin{aligned} |\phi_1^A\rangle &= \sqrt{2s_1}|u_1^A\rangle, \\ |\phi_2^A\rangle &= \sqrt{\frac{s_1}{2}}|u_1^A\rangle + \sqrt{\frac{s_1}{2} - s_2}|u_2^A\rangle, \\ |\phi_3^A\rangle &= \sqrt{\frac{s_1}{2}}|u_1^A\rangle - \sqrt{\frac{s_1}{2} - s_2}|u_2^A\rangle. \end{aligned} \tag{4.22}$$

Let us look at an example of each solution. If we choose our three states to be $\psi_1 = (1,0,0)$, $\psi_2 = \frac{1}{\sqrt{3}}(1,1,1)$ and $\psi_3 = \frac{1}{\sqrt{3}}(1,1,-1)$, we find that $s_1 = 1/\sqrt{3}$ and $s_2 = 1/3$, so that solution 1 is valid. The complete treatment of this case (case 2) is given in Table 2. We see that we need an eight-port which can be built up by two

$U(2)$ blocks. Note that in order to achieve minimum failure probability Q, we need to choose q_1 to be 1, which means that we sacrifice the possibility of distinguishing state $|\psi_1\rangle$.

TABLE 2: Summary of parameters and arguments of $T_{p,q}$ for case 2.

	case 2
Input states	$\|\psi_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, $\|\psi_2\rangle = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ 0 \end{pmatrix}$, $\|\psi_3\rangle = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ 0 \end{pmatrix}$
Optimal failure probability	$q_1 = 1,\ q_2 = \frac{1}{3},\ q_3 = \frac{1}{3}.$
Output states	$\|\psi_1'\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$, $\|\psi_2'\rangle = \begin{pmatrix} 0 \\ \sqrt{\frac{2}{3}} \\ 0 \\ \frac{1}{\sqrt{3}} \end{pmatrix}$, $\|\psi_3'\rangle = \begin{pmatrix} 0 \\ 0 \\ \sqrt{\frac{2}{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$
$M =$	$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$
Factorization of M	$M = T_{1,4} \cdot T_{2,3}$
Arguments of $T_{p,q}$	$T_{1,4} : \omega = 0, \phi = 0$ $T_{2,3} : \omega = \frac{\pi}{4}, \phi = 0$

If we choose our states to be $\psi_1 = (1,0,0)$, $\psi_2 = \frac{1}{3}(1,2,2)$ and $\psi_3 = \frac{1}{3}(1,2,-2)$, then we find that solution 2 is valid with $s_1 = 1/3$ and $s_2 = 1/9$. In this case (case 3) we need a ten-port, and the complete results are given in Table 3. Note that if the procedure fails, it is still possible to gain some information about the input state, because the failure space is two-dimensional [24]. This is not possible if the failure space has only one dimension.

One possibility is to attach to the failure-space outputs (outputs 4 and 5) a network that transforms states $|\phi_2^A\rangle$ and $|\phi_3^A\rangle$ into orthogonal states, which it will do only with a certain probability [5]. In particular, we can construct a network that implements the transformation

$$M(3) = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \sqrt{\frac{2}{3}} & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}, \tag{4.23}$$

where the inputs to the first two ports of this network (we shall call these ports A and B) are the outputs of ports 4 and 5 of the original network, and the input to the third port (port C) is the vacuum. This network has been designed so that if no photon is detected emerging from output C, then the input state $|\phi_2^A\rangle$ will be transformed into a photon emerging from port A, and the input state $|\phi_3^A\rangle$ will be transformed into a photon emerging from port B. If the input state is $|\phi_1^A\rangle$ and no photon is detected at output C, the probabilities of a photon emerging from either port A or port B are the same. Therefore, if the photon emerges from port A, we can conclude the input to the entire network was either ψ_1 or ψ_2, and if it emerges from port B, then the input was

TABLE 3: Summary of parameters and arguments of $T_{p,q}$ for case 3.

Input states	$\|\psi_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, $\|\psi_2\rangle = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \\ 0 \\ 0 \end{pmatrix}$, $\|\psi_3\rangle = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ -\frac{2}{3} \\ 0 \\ 0 \end{pmatrix}$
Optimal failure probability	$q_1 = \frac{2}{3}$, $q_2 = \frac{2}{9}$, $q_3 = \frac{2}{9}$
Output states	$\|\psi_1'\rangle = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ 0 \\ 0 \\ \sqrt{\frac{2}{3}} \\ 0 \end{pmatrix}$, $\|\psi_2'\rangle = \begin{pmatrix} 0 \\ \frac{\sqrt{7}}{3} \\ 0 \\ \frac{1}{\sqrt{6}} \\ \frac{1}{3\sqrt{2}} \end{pmatrix}$, $\|\psi_3'\rangle = \begin{pmatrix} 0 \\ 0 \\ \frac{\sqrt{7}}{3} \\ \frac{1}{\sqrt{6}} \\ -\frac{1}{3\sqrt{2}} \end{pmatrix}$
M	$\begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{3}} & 0 & -\sqrt{\frac{7}{12}} & 0 \\ 0 & \frac{\sqrt{7}}{4} & \frac{\sqrt{7}}{4} & -\frac{1}{4} & \frac{1}{4} \\ 0 & \frac{\sqrt{7}}{4} & -\frac{\sqrt{7}}{4} & -\frac{1}{4} & -\frac{1}{4} \\ \sqrt{\frac{2}{3}} & \frac{1}{2\sqrt{6}} & 0 & \sqrt{\frac{7}{24}} & 0 \\ 0 & 0 & \frac{1}{2\sqrt{2}} & 0 & -\sqrt{\frac{7}{8}} \end{pmatrix}$
Factorization of M	$M = T_{1,4} \cdot T_{2,3} \cdot T_{2,4} \cdot T_{3,5}$
Arguments of $T_{p,q}$	$T_{1,4} : \{\omega = \arccos[\sqrt{\frac{2}{3}}], \phi = 0\}$, $T_{2,3} : \{\omega = \frac{\pi}{4}, \phi = 0\}$ $T_{2,4} : \{\omega = \arccos[-\frac{1}{2\sqrt{2}}], \phi = 0\}$, $T_{3,5} : \{\arccos[\frac{1}{2\sqrt{2}}], \phi = 0\}$

either ψ_1 or ψ_3. Summarizing, if one of the detectors in ports 1 through 3 clicks, we know what the input state was. If the detector in either ports A or B clicks, then we gain partial information about the input state; the number of possibilities has been reduced from three to two. If the detector in port C clicks, then we have gained no information about the input state, and this happens with a probability of 1/9 if the inputs were ψ_2 or ψ_3 and 4/9 if the input was ψ_1. The addition of the second network to the failure outputs of the first significantly improves the chances of gaining some information about the input state.

Chapter 5

Optimum unambiguous discrimination between $|\psi_1\rangle$ and $\{|\psi_2\rangle, |\psi_3\rangle\}$

In this section, we consider a new variant of the problem. Instead of discriminating among all states, we ask what happens if we just want to discriminate between subsets of them. In particular, if there are three non-orthogonal states $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle\}$, what is the optimum strategy to distinguish $|\psi_1\rangle$ from the set $\{|\psi_2\rangle, |\psi_3\rangle\}$? This problem would occur if, among three states, $|\psi_1\rangle$ is the only interesting one, so that we just want to distinguish it from the other two. We refer to this problem as quantum state filtering. The analytical solution of this problem is derived in this paper.

5.1. Derivation of the optimal solution

Suppose we are given a quantum system prepared in the state $|\psi\rangle$, which is guaranteed to be a member of the set of three non-orthogonal states $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle\}$, but we do not know which one. We want to find a procedure which will tell us that $|\psi\rangle$ was prepared in $|\psi_1\rangle$, or will tell us that $|\psi\rangle$ was prepared in one of $\{|\psi_2\rangle, |\psi_3\rangle\}$. That is, the procedure can distinguish $|\psi_1\rangle$ from $\{|\psi_2\rangle, |\psi_3\rangle\}$. We also want this procedure to be error-free, i. e. the procedure may fail to give us any information about the

state, and if it fails, it must let us know that it has, but if it succeeds, it should never give us a wrong answer. We shall refer to such a procedure as quantum state filtering without error. We find that, in contrast to the unambiguous state discrimination problem, this will be possible even if $|\psi_1\rangle$ is not linearly independent from the set $\{|\psi_2\rangle, |\psi_3\rangle\}$.

If the states are not orthogonal then, they cannot be discriminated perfectly. For convenience, we can rewrite Eq. (2.12), which give the average probabilities of success and of failure to distinguish the states $|\psi_i\rangle$

$$P = \sum_i \eta_i p_i,$$
$$Q = \sum_i \eta_i q_i,$$

respectively. Our objective is to find the set of $\{p_i\}$ that maximizes the probability of success, P.

To achieve the unambiguous optimal discrimination, The desired "generalized measurement" can be described as follows. Let $\mathcal{K}$ denote a total Hilbert space, which is the direct sum of two subspaces, $\mathcal{K} = \mathcal{H} \oplus \mathcal{A}$. The space $\mathcal{H}$ is a three-dimensional space that contains the vectors $|\psi_i\rangle$, and $\mathcal{A}$ is an auxiliary space. The input state of the system is one of the vectors $|\psi_i\rangle$, which is now a vector in the subspace $\mathcal{H}$ of the total space $\mathcal{K}$, so that

$$|\psi_i^{\mathcal{K}}\rangle_{in} = |\psi_i^{\mathcal{H}}\rangle. \tag{5.1}$$

A unitary transformation, U, which acts in the entire space $\mathcal{K}$ is now applied to the

input vector, resulting in the state $|\psi_i^{\mathcal{K}}\rangle_{out}$, which is given by

$$|\psi_i^{\mathcal{K}}\rangle_{out} = |\psi_i'^{\mathcal{H}}\rangle + |\phi_i^{\mathcal{A}}\rangle = U|\psi_i^{\mathcal{K}}\rangle_{in}, \tag{5.2}$$

where, in our case, $|\psi_1'\rangle$ can always be unambiguously distinguished from the set $\{|\psi_2'\rangle, |\psi_3'\rangle\}$. Then a measurement is performed on $|\psi_i^{\mathcal{K}}\rangle_{out}$ that projects $|\psi_i^{\mathcal{K}}\rangle_{out}$ either onto $|\psi_i'\rangle$ or $|\phi_i\rangle$ (by construction, they are in orthogonal subspaces). If it projects $|\psi_i^{\mathcal{K}}\rangle_{out}$ onto $|\psi_i'\rangle$, the procedure succeeds, because $|\psi_1'\rangle$ can always be distinguished from $\{|\psi_2'\rangle, |\psi_3'\rangle\}$. The probability to get this outcome, if the input state is $|\psi_i\rangle$, is

$$p_i = \langle\psi_i'|\psi_i'\rangle. \tag{5.3}$$

If the measurement projects $|\psi_i^{\mathcal{K}}\rangle_{out}$ onto $|\phi_i\rangle$, the procedure fails. The probability of this outcome is

$$q_i = 1 - p_i = \langle\phi_i|\phi_i\rangle. \tag{5.4}$$

The nature of the problem we are trying to solve imposes a number of requirements on the output vectors. The condition that $|\psi_1'\rangle$ be distinguishable from $|\psi_2'\rangle$ and $|\psi_3'\rangle$ requires that

$$\langle\psi_1'|\psi_2'\rangle = \langle\psi_1'|\psi_3'\rangle = 0. \tag{5.5}$$

These lead to conditions on the failure vectors, $|\phi_i\rangle$. As described by Eq. (2.14), taking the scalar product between $|\psi_1^{\mathcal{K}}\rangle_{out}$ and the other two output states and using Eq. (5.5) and the fact that U is unitary leads to the conditions

$$\begin{aligned} \langle\phi_1|\phi_2\rangle &= \langle\psi_1|\psi_2\rangle, \\ \langle\phi_1|\phi_3\rangle &= \langle\psi_1|\psi_3\rangle. \end{aligned} \tag{5.6}$$

Our objective is to find the optimal $|\psi_i'\rangle$ and $|\phi_i\rangle$ which satisfy Eqs. (5.3)– (5.6) and also give the maximum success probability P.

Let us now consider the failure vectors. If they were linearly independent, we could apply a state discrimination procedure to them [25]. That means that if our original procedure fails, and we end up in the failure space, $\mathcal{A}$, then we still have some chance of determining what our input state was. This clearly implies that our original procedure, which led to the vectors $|\psi'\rangle$, was not optimal, because that process followed by another on the failure vectors would lead to a higher probability of distinguishing $|\psi_1\rangle$ from $|\psi_2\rangle$ and $|\psi_3\rangle$. Therefore, the optimal procedure should lead to failure vectors to which we cannot successfully apply a state discrimination procedure, implying that they are linearly dependent. In fact, we will now prove that for optimal discrimination they must be collinear, by demonstrating that the contrary leads to contradiction. To this end, we assume that we have achieved optimal unambiguous discrimination of $|\psi_1\rangle$ from $|\psi_2\rangle$ and $|\psi_3\rangle$ but the failure vectors are *not* collinear. Then at least one of the two failure vectors, $|\phi_2\rangle$, $|\phi_3\rangle$, will have a component in the direction that is perpendicular to $|\phi_1\rangle$. We can set up a detector projecting onto this direction and a positive outcome of the measurement (a click of the detector) will tell us that our input state was not $|\psi_1\rangle$ but one of the other two states. Thus, contrary to our assumption that our procedure has been optimal, further distinction is possible. Hence, the failure vectors must be collinear for optimal discrimination.

We shall now explore the consequences of this conclusion. Since $|\phi_i\rangle$ $(i = 1, \ldots, n)$ are collinear, the failure space, $\mathcal{A}$, is one dimensional. If $|u\rangle$ is the basis vector

spanning this Hilbert space we can write the failure vectors as $|\phi_i\rangle = \sqrt{q_i}e^{x_i}|u\rangle$. Substituting this representation of the failure vectors into Eq. (5.6), we find that

$$\begin{aligned} q_1 q_2 &= |\langle\psi_1|\psi_2\rangle|^2, \\ q_1 q_3 &= |\langle\psi_1|\psi_3\rangle|^2. \end{aligned} \tag{5.7}$$

These two conditions are a consequence of unitarity and imply that only one of the three failure probabilities can be chosen independently. If we chose q_1 as the independent one we can express the other two as $q_2 = |\langle\psi_1|\psi_2\rangle|^2/q_1$ and $q_3 = |\langle\psi_1|\psi_3\rangle|^2/q_1$. If we introduce the notation $O_{ij} = \langle\psi_i|\psi_j\rangle$ then, with the help of these two equations, the average failure probability can be written explicitly as

$$\begin{aligned} Q &= \sum_i \eta_i q_i \\ &= \eta_1 q_1 + \frac{\eta_2|O_{12}|^2 + \eta_3|O_{13}|^2}{q_1}. \end{aligned} \tag{5.8}$$

If we further introduce the notation $A = \eta_2|O_{12}|^2 + \eta_3|O_{13}|^2$ for the frequently occuring average overlap then, from the condition

$$\frac{dQ}{dq_1} = 0, \tag{5.9}$$

we find the optimal value of q_1 to be

$$q_1 = \sqrt{A/\eta_1}. \tag{5.10}$$

This value, however, cannot always be realized. For it to be true, there must be a unitary transformation, from Eq. (5.2), that takes $|\psi_j\rangle$ to $|\psi_j\rangle_{out}$ which, together

with the one-dimensionality of the failure space yields

$$|\psi_j\rangle_{out} = |\psi_j'\rangle + \sqrt{q_j}|e^{ix_j}|u\rangle. \tag{5.11}$$

Here we have that $\langle\psi_j'|u\rangle = 0$, $\langle\psi_1'|\psi_j'\rangle = 0$ for $j = 2, 3$, and the phase factors are fixed by the requirement (cf. Eq. (5.6)) that

$$\langle\psi_1|\psi_j\rangle = \sqrt{q_1 q_j}e^{i(x_j - x_1)} \tag{5.12}$$

for $j = 2, 3$. These equations imply that

$$\langle\psi_j'|\psi_k'\rangle = \langle\psi_j|\psi_k\rangle - \sqrt{q_j q_k}e^{i(x_k - x_j)}. \tag{5.13}$$

This set of equations can only be true if the matrix M, where

$$M_{jk} = \langle\psi_j|\psi_k\rangle -,, \sqrt{q_j q_k}e^{i(x_k - x_j)}, \tag{5.14}$$

is positive semidefinite, as discussed in detail in Section 2.2

Using again $O_{jk} = \langle\psi_j|\psi_k\rangle$, M can be expressed as

$$M = \begin{pmatrix} 1 - q_1 & 0 & 0 \\ 0 & 1 - \frac{|O_{12}|^2}{q_1} & O_{23} - \frac{O_{21}O_{13}}{q_1} \\ 0 & O_{32} - \frac{O_{31}O_{12}}{q_1} & 1 - \frac{|O_{13}|^2}{q_1} \end{pmatrix}. \tag{5.15}$$

Clearly, this matrix will be positive semidefinite if $0 \leq q_1 \leq 1$, and if the 2×2 submatrix is also positive semidefinite. This will be true if both the trace and determinant of the submatrix are greater than or equal to zero. Positivity requires that the diagonal matrix elements of the submatrix be non-negative, so that it must be true that $q_1 \geq |O_{12}|$ and $q_1 \geq |O_{13}|$. Without loss of generality, we can assume

that $|O_{12}| \geq |O_{13}|$ by simply arranging the states in set 2 in the order of decreasing overlaps with $|\psi_1\rangle$. Doing so and imposing the condition that $q_1 \geq |O_{12}|$ guarantees that the condition $q_1 \geq |O_{13}|$ is also satisfied, and together they imply that the trace is greater than or equal to zero.

The condition that the determinant be non-negative gives us a lower bound on q_1,

$$q_1 \geq \frac{|O_{12}|^2 + |O_{13}|^2 - (O_{12}O_{23}O_{31} + O_{13}O_{32}O_{21})}{1 - |O_{23}|^2}. \tag{5.16}$$

We want to interpret this inequality, in particular, we want to find what the right-hand side is equal to. In order to do so, we shall find the projection operator, P_{23}, that projects onto the subspace spanned by ψ_2 and ψ_3. One of the basis vectors in this subspace can be chosen to be $|\psi_2\rangle$ and, using the Gram-Schmidt orthogonalization method, the other is defined as the (normalized) orthogonal component of $|\psi_3\rangle$,

$$|\tilde{\psi}_3\rangle = \frac{1}{\sqrt{1 - |O_{23}|^2}}(|\psi_3\rangle - O_{23}|\psi_2\rangle). \tag{5.17}$$

This leads to

$$P_{23} = |\psi_2\rangle\langle\psi_2| + |\tilde{\psi}_3\rangle\langle\tilde{\psi}_3|. \tag{5.18}$$

Let us represent the input state, $|\psi_1\rangle$, as $|\psi_1\rangle = |\psi_1^{\perp}\rangle + |\psi_1^{\parallel}\rangle$, where $|\psi_1^{\perp}\rangle = (1 - P_{23})|\psi_1\rangle$ is the component of the input vector that is perpendicular to the subspace spanned by ψ_2 and ψ_3 and $|\psi_1^{\parallel}\rangle = P_{23}|\psi_1\rangle$ is the component in that subspace. Then, using Eqs. (5.17) and (5.18), the explicit expression for the parallel component is given by

$$|\psi_1^{\parallel}\rangle = \frac{O_{21} - O_{23}O_{31}}{1 - |O_{23}|^2}|\psi_2\rangle + \frac{O_{31} - O_{32}O_{21}}{1 - |O_{23}|^2}|\psi_3\rangle]. \tag{5.19}$$

Calculating the norm of this expression yields

$$\langle\psi_1^{\parallel}|\psi_1^{\parallel}\rangle = \frac{|O_{12}|^2 + |O_{13}|^2 - (O_{12}O_{23}O_{31} + O_{13}O_{32}O_{21})}{1 - |O_{23}|^2}, \tag{5.20}$$

which is identical to the right-hand side of Eq. (5.16).

Thus, Eq. (5.16) tells us that the failure probability, q_1, has a lower bound which is given by the weight of $|\psi_1\rangle$ in the other subspace, $\|P_{23}\psi_1\|^2 = \langle\psi_1|P_{23}|\psi_1\rangle = \langle\psi_1^{\parallel}|\psi_1^{\parallel}\rangle$, a result that is intuitively obvious. Clearly, this expression is larger than (or at most equal to) $|O_{12}|^2$. This implies that, because $q_2 = |O_{12}|^2/q_1$, we have

$$\begin{aligned} q_2 &\leq \frac{|O_{12}|^2}{\langle\psi_1^{\parallel}|\psi_1^{\parallel}\rangle} \\ &= \frac{|O_{12}|^2}{|O_{12}|^2 + |\langle\bar{\psi}_3|\psi_1\rangle|^2} \leq 1, \end{aligned} \tag{5.21}$$

and similarly for q_3.

We can then distinguish three different regimes of the parameters. If the r.h.s. of Eq. (5.10) is greater than 1 then $q_1 = 1$, if it is less than $\langle\psi_1^{\parallel}|\psi_1^{\parallel}\rangle$ then $q_1 = \langle\psi_1^{\parallel}|\psi_1^{\parallel}\rangle$, and in the intermediate range the optimum given by Eq. (5.10) is realized. This can be summarized as follows.

(i) If $\eta_1|\langle\psi_1^{\parallel}|\psi_1^{\parallel}\rangle|^2 \leq A \leq \eta_1$, then

$$\begin{aligned} q_1 &= \sqrt{A/\eta_1}, \\ q_2 &= \sqrt{\eta_1/A}|O_{12}|^2, \\ q_3 &= \sqrt{\eta_1/A}|O_{13}|^2, \end{aligned} \tag{5.22}$$

yielding the average failure probability

$$Q = 2\sqrt{\eta_1 A}. \tag{5.23}$$

(ii) If $A \geq \eta_1$, then

$$\begin{aligned} q_1 &= 1, \\ q_2 &= |O_{12}|^2, \\ q_3 &= |O_{13}|^2. \end{aligned} \tag{5.24}$$

yielding the average failure probability

$$Q = \eta_1 + A. \tag{5.25}$$

(iii) If $A \leq \eta_1 |\langle \psi_1^{\perp} | \psi_1^{\parallel} \rangle|^2$, then

$$\begin{aligned} q_1 &= \langle \psi_1^{\parallel} | \psi_1^{\parallel} \rangle, \\ q_2 &= \frac{|O_{12}|^2}{\langle \psi_1^{\parallel} | \psi_1^{\parallel} \rangle}, \\ q_3 &= \frac{|O_{13}|^2}{\langle \psi_1^{\parallel} | \psi_1^{\parallel} \rangle}, \end{aligned} \tag{5.26}$$

yielding the average failure probability

$$Q = \eta_1 \langle \psi_1^{\parallel} | \psi_1^{\parallel} \rangle + \frac{A}{\langle \psi_1^{\parallel} | \psi_1^{\parallel} \rangle}. \tag{5.27}$$

Equations (5.22)–(5.27) summarize our main results. In the intermediate range of the average overlap, A, the optimal failure probability, Eq. (5.23), is achieved by a generalized measurement or POVM. Outside this region, for very large average overlap, $A \geq \eta_1$, or very small average overlap, $A \leq \eta_1 |\langle \psi_1^{\perp} | \psi_1^{\parallel} \rangle|^2$, the optimal failure probabilities, Eqs. (5.25) and (5.27), are realized by standard von Neumann measurements. For very large A the optimal von Neumann measurement consists of projections onto $|\psi_1\rangle$ and two orthogonal directions whose directionality needs not be

specified further. A click along $|\psi_1\rangle$ corresponds to failure because it can have its origin in any of the two subsets and a click in the orthogonal directions uniquely assigns the input state to the set $\{|\psi_2\rangle, |\psi_3\rangle\}$. For very small A the optimal von Neumann measurement consists of projections onto $|\psi_1^{\parallel}\rangle$ and two orthogonal directions that are uniquely determined by the requirement that they correspond to two mutually exclusive alternatives. One of them is onto $|\psi_1^{\perp}\rangle$ and the other onto the remaining orthogonal direction in the subspace of $\{|\psi_2\rangle, |\psi_3\rangle\}$. A click along $|\psi_1^{\parallel}\rangle$ corresponds to failure because it can originate from any of the input states while a click in any of the alternative directions unambiguously assigns the input to one or the other of the two mutually exclusive subsets. It is interesting to observe that the failure space is one dimensional for each of the three different optimal measurements in the three different regions. At the boundaries of their respective regions of validity, the optimal measurements transform into one another continuously. Furthermore, each of the two von Neumann expressions can be written as the arithmetic mean of two terms and the POVM result as the geometric mean of the same two terms. Therefore, in its range of validity the POVM performs better than any von Neumann measurement.

In closing this Section we want to point out an interesting feature of the solution. The results hold true even when there is no perpendicular component of the first input state, $|\psi_1^{\perp}\rangle = 0$, i.e. it lies entirely in the Hilbert space spanned by the other two vectors or, in other words, the two sets are linearly dependent. In this case the two von Neumann measurements coincide and the range of validity of the POVM solution shrinks to zero. A click in the detector along the first input vector corresponds to

failure - it might originate from either of the two subsets - and a click in the detector along the single direction orthogonal to it unambiguously identifies the set of the other two vectors.

5.2. Derivation of the optimal solution via the method of Lagrange multipliers

In this section, we shall show that by using the method of Lagrange multipliers, we can derive the conclusions contained in Eqs. (5.22)-(5.27) rigorously, starting from the fact that for optimal discrimination, the vectors $|\phi_i\rangle$ must be linearly dependent. To express this statement in a compact form we use the constraint given by Eq. (2.16), which says that for the positive semidefinite matrix C, where $C_{ij} = \langle\phi_i|\phi_j\rangle$, if $|\phi_i\rangle$ $(i = 1, \ldots, n)$ are linearly dependent, the determinant of matrix C must vanish, $\Delta = \det(C) = 0$. With the help of Eqs. (5.4) and (5.6), we can eliminate two of the three overlaps from the matrix C and obtain explicitly

$$\begin{aligned}\Delta &= \begin{vmatrix} q_1 & O_{12} & O_{13} \\ O_{12}^* & q_2 & re^{i\theta} \\ O_{13}^* & re^{-i\theta} & q_3 \end{vmatrix} \\ &= q_1q_2q_3 - r^2q_1 - |O_{13}|^2q_2 - |O_{12}|^2q_3 \\ &\quad +2|O_{12}||O_{13}|r\cos(\theta - \alpha) = 0. \end{aligned} \tag{5.28}$$

Here O_{ij} again denotes $\langle\psi_i|\psi_j\rangle$, $re^{i\theta} = \langle\phi_2|\phi_3\rangle$ is the remaining overlap where r and θ are to be determined from the conditions for optimum, and $\alpha = -\arg(O_{12}O_{13}^*)$. Since C is positive semidefinite, all the diagonal subdeterminants of Δ must be non-

negative.

We now wish to minimize the average probability of failure Q, Eq. (5.1), subject to the constraint in Eq. (5.28). This can be done by minimizing the quantity

$$\tilde{Q} = \sum_{i}^{3} \eta_i q_i + \lambda \Delta, \tag{5.29}$$

where λ is a Lagrange multiplier. The conditions for minimum with respect to r and θ, $\partial\tilde{Q}/\partial r = 0$ and $\partial\tilde{Q}/\partial\theta = 0$, lead immediately to

$$|O_{12}||O_{13}|\cos(\theta - \alpha) - q_1 r = 0, \tag{5.30}$$

$$r|O_{12}||O_{13}|\sin(\theta - \alpha) = 0. \tag{5.31}$$

The solutions of these equations, corresponding to the minimum of Q, are

$$\theta = \alpha, \tag{5.32}$$

and

$$q_1 r = |O_{12}||O_{13}|. \tag{5.33}$$

Next, we perform the optimization with respect to the remaining variables. Notice that the derivative of $\tilde{Q}$ with respect to λ returns Eq. (5.28). Therefore, we use the optimal values of r and θ in Eq. (5.28) and in the conditions for minimum with respect to the failure probabilities, $\partial\tilde{Q}/\partial q_i = 0$ for $i = 1, 2, 3$. After some algebra we

obtain the following set of equations

$$q_1\Delta = \Delta_{12}\Delta_{13} = 0, \tag{5.34}$$

$$q_1^2\frac{\partial\tilde{Q}}{\partial q_1} = \eta_1 q_1^2 + \lambda(\Delta_{12}\Delta_{13} + |O_{12}|^2\Delta_{13} + |O_{13}|^2\Delta_{12}) = 0, \tag{5.35}$$

$$\frac{\partial\tilde{Q}}{\partial q_2} = \eta_2 + \lambda\Delta_{13} = 0, \tag{5.36}$$

$$\frac{\partial\tilde{Q}}{\partial q_3} = \eta_3 + \lambda\Delta_{12} = 0, \tag{5.37}$$

where Δ_{12} and Δ_{13} are the diagonal subdeterminants of Δ,

$$\Delta_{12} = q_1 q_2 - |O_{12}|^2, \tag{5.38}$$

$$\Delta_{13} = q_1 q_3 - |O_{13}|^2. \tag{5.39}$$

We now have four variables q_1, q_2, q_3, and λ, and four equations, Eqs. (5.34)–(4.6), to find them. Eq. (5.34) tells us that at least one of the diagonal subdeterminants vanishes. With no loss of generality we can assume this to be $\Delta_{12} = 0$. Comparing this to Eq. (4.6) we see that λ must be singular. The singularity, however, is tractable since the same equation tells us that the product $\lambda\Delta_{12}$ is finite. Then it follows from the singular behavior of λ and Eq. (4.5) that the other diagonal subdeterminant also vanishes, $\Delta_{13} = 0$, but the product $\lambda\Delta_{12}$ also remains finite. Using these finite values from Eqs. (5.36)–(5.37) in Eq. (5.35), we can summarize our findings as follows

$$\Delta_{12} = \Delta_{13} = 0, \tag{5.40}$$

which is just equation (5.7), and

$$\eta_1 q_1^2 - \eta_2|O_{12}|^2 - \eta_3|O_{13}|^2 + \lambda\Delta_{12}\Delta_{13} = 0. \tag{5.41}$$

Multiplying Eq. (4.5) by Δ_{12} (or Eq. (4.6) by Δ_{13}) and taking into account Eq. (5.40) gives that the singularity in λ is such that $\lambda\Delta_{12}\Delta_{13} = 0$. Using this in Eq. (5.41) we finally obtain

$$\eta_1 q_1^2 - \eta_2|O_{12}|^2 - \eta_3|O_{13}|^2 = 0. \tag{5.42}$$

This is the solution found in Section II, Eq. (5.10), and the rest of Section II follows from here and Eq. (5.40).

For the sake of completeness we also give the expression for $1/\lambda$,

$$\frac{1}{\lambda} = -\sqrt{\frac{\Delta_{12}\Delta_{13}}{\eta_2\eta_3}}, \tag{5.43}$$

which exhibits no singularity. In fact, $1/\lambda = 0$ when $\Delta_{12} = \Delta_{13} = 0$, as expected. Finally, let us note that Eq. (5.40), which is identical to Eq. (5.7), implies that all of the failure vectors, $|\phi_i\rangle$, are parallel to each other, i. e. they lie in a space, $\mathcal{A}$, of dimension one.

5.3. Comparison to the case when all states are discriminated

In this section we want to compare the average probability of failure Q of the filtering problem to that of distinguishing all three states. Let Q' denote the average probability of failure for distinguishing all the states $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle\}$. We can see immediately, that the probability of failure to distinguish $|\psi_1\rangle$ from $\{|\psi_2\rangle, |\psi_3\rangle\}$, Q, should be no larger than Q'. For the latter problem, the necessary condition for

achieving optimal discrimination is

$$\begin{vmatrix} q_1 & O_{12} & O_{13} \\ O_{12}^* & q_2 & O_{23} \\ O_{13}^* & O_{23}^* & q_3 \end{vmatrix} = 0. \tag{5.44}$$

When comparing this equation to Eq. (5.28), we see that, instead of a given constant O_{23} that appears in Eq. (5.44), there are the variables r and θ in Eq. (5.28). These variables are chosen to minimize the average probability of failure Q. Therefore, Q should be no larger than Q', $Q \leq Q'$.

To illustrate this point, we use a simple symmetric case, where all of the overlaps between the states are real and equal,

$$\langle\psi_1|\psi_2\rangle = \langle\psi_1|\psi_3\rangle = \langle\psi_2|\psi_3\rangle = s, \tag{5.45}$$

with $0 < s < 1$. We shall also assume that the *a priori* probabilities are equal for all the examples in this paper. From previous work we know that in this case, the optimal values of the failure probabilities when we wish to distinguish among all of the states $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle\}$ are $q_i = s$, which implies that $Q' = s$.

For the problem of distinguishing $|\psi_1\rangle$ from $\{|\psi_2\rangle, |\psi_3\rangle\}$, from the results of Eqs. (5.22) and (5.24), we have (i) if $0 < s \leq \frac{\sqrt{2}}{2}$, then

$$\begin{aligned} q_1 &= \sqrt{2}s, \\ q_2 = q_3 &= \frac{\sqrt{2}}{2}s, \\ Q &= \frac{2\sqrt{2}}{3}s. \end{aligned} \tag{5.46}$$

So the average probability of failure Q is less than $Q' = s$. (ii) if $\frac{\sqrt{2}}{2} < s < 1$, then

$$q_1 = 1,$$

$$q_2 = q_3 = s^2,$$

$$Q = \frac{1}{3} + \frac{2}{3}s^2. \tag{5.47}$$

These solutions are illustrated and compared to Q' in Figure 3. Note that in both cases we have that $Q < s = Q'$.

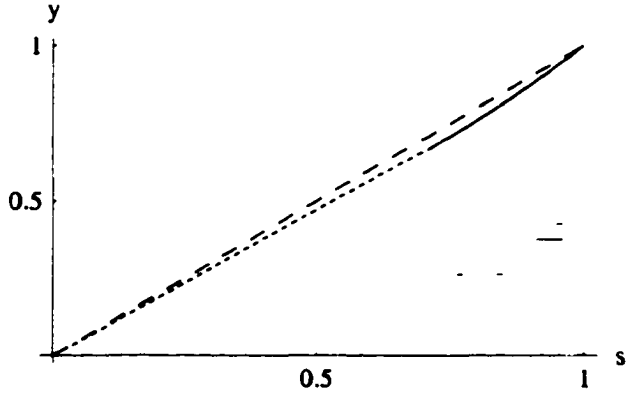

FIG. 3: We compare Q and Q'. For $0 < s \leq \frac{\sqrt{2}}{2}$ we have that $Q' = s$ and $Q = \frac{2\sqrt{2}}{3}s$. For $\frac{\sqrt{2}}{2} < s \leq 1$, we still have that $Q' = s$, but $Q = \frac{1}{3} + \frac{2}{3}s^2$. Note that Q is always smaller than Q'.

Now we shall compare filtering to the problem of distinguishing two states $\{|\psi_1\rangle, |\psi_2\rangle\}$, when all the *a priori* probabilities are equal. If we denote by Q'' the average probability of failure when distinguishing between the two states $\{|\psi_1\rangle$ and $|\psi_2\rangle\}$, we know that $Q'' = |O_{12}|$ as indicated by Eq. (3.10). For the case we are considering, $|O_{12}| = |O_{13}| = s$, and we see that $Q < Q''$.

A second example is more illuminating. The overlaps are now given by

$$\begin{aligned} \langle\psi_1|\psi_2\rangle &= \langle\psi_1|\psi_3\rangle = s_1, \\ \langle\psi_2|\psi_3\rangle &= s_2, \end{aligned} \tag{5.48}$$

where, for simplicity, s_1 and s_2 are real, $0 < s_1, s_2 < 1$, and

$$0 < s_1 < \frac{\sqrt{2}}{2}, \; s_1^2 < s_2, \text{ and } s_1 < 2s_2.$$

The probabilities of failure for discriminating $|\psi_1\rangle$ from $\{|\psi_2\rangle, |\psi_3\rangle\}$ are

$$\begin{aligned} q_1 &= \sqrt{2}s_1, \\ q_2 &= q_3 = \frac{\sqrt{2}}{2}s_1, \end{aligned} \tag{5.50}$$

and the average failure probability is

$$Q = \frac{2\sqrt{2}}{3}s_1. \tag{5.51}$$

The optimal probabilities of failure for discriminating among all three states $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle\}$ are given by

$$\begin{aligned} q_1' &= \frac{s_1^2}{s_2} \\ q_2' &= q_3' = s_2 \\ Q' &= \frac{1}{3}[(s_1^2/s_2) + 2s_2]. \end{aligned} \tag{5.52}$$

Q can be compared to Q' by examining the ratio

$$\frac{Q}{Q'} = \frac{2\sqrt{2}s_1 s_2}{s_1^2 + 2s_2^2} \leq 1. \tag{5.53}$$

From the above equation, we see that when s_1 is much smaller than s_2, Q is much smaller than Q'. For example, when $s_1 = \frac{\sqrt{2}}{5}$, $s_2 = \frac{4}{5}$, $Q/Q' = 0.47$.

5.4. Optical realization

Now we shall present a scheme for a possible experimental realization of the optimal discrimination between $|\psi_1\rangle$ and $\{|\psi_2\rangle, |\psi_3\rangle\}$. We can follow the method described by Section 2.4. Recall that the dimension of the total Hilbert space is four, so we shall require four modes, and the input states $|\psi_i\rangle$ will be represented by single photon states as

$$|\psi_i\rangle = \sum_{j=1}^{4} d_{ij}\hat{a}_j^\dagger|0\rangle, \tag{5.54}$$

where $\sum_{j=1}^{4} |d_{ij}|^2 = 1$, and $\hat{a}_j^\dagger$ is the creation operator for the jth mode. We shall require $d_{i4} = 0$ for $i = 1, 2, 3$, that is, the initial single photon state is sent to the first three input ports, and the vacuum into the fourth input port. The first three modes correspond to the space, $\mathcal{H}$, containing the states to be distinguished and the fourth mode to the failure space, $\mathcal{A}$. Since the dimension of the input and output states is four, here we shall use an eight-port (see Figure 4).

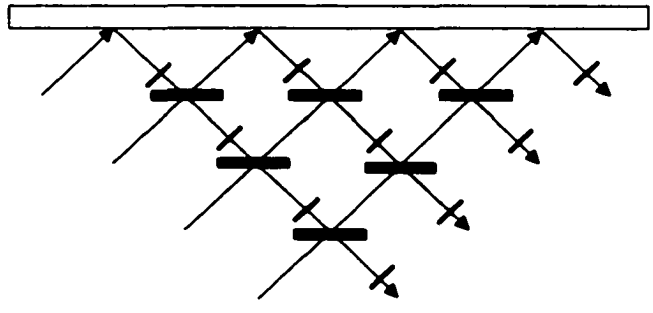

FIG. 4: An optical eight-port. The beams are straight lines, a suitable beam splitter is placed at each point where two beams intersect, phase shifters are at one input of each beam splitter and at each output.

Now we rewrite Eq. (2.23) as

$$a_{j\text{out}} = U^{-1} a_j U = \sum_{k=1}^{4} M_{jk} a_k, \tag{5.55}$$

where M_{jk} are the elements of a 4×4 unitary matrix $M(4)$. In the Schrödinger picture, the *in* and *out* states are related by

$$|\psi\rangle_{out} = U|\psi\rangle_{in}. \tag{5.56}$$

It is shown in Section 2.4, Eq. (2.27), that when using single photon states representation, the matrix element M_{il} is the same as the matrix element of U between the single-particle states $|i\rangle = a_i^\dagger|0\rangle$ and $|l\rangle = a_l^\dagger|0\rangle$, i.e.,

$$\langle i|U|l\rangle = M_{il}. \tag{5.57}$$

To design the desired eight-port, we first calculate the optimal value of q_i. Then from Eq. (5.4) and the fact that our failure space is one-dimensional, the vectors $|\phi_i\rangle$ are given by

$$|\phi_i\rangle = \sqrt{q_i}|1^{\mathcal{A}}\rangle = \sqrt{q_i} a_4^\dagger|0\rangle, \tag{5.58}$$

where the state $|1^{\mathcal{A}}\rangle$ denotes one photon state in the failure space, which is just one photon in mode 4. Once the vectors $|\phi_i\rangle$ are determined, the inner products $\langle\psi_i'|\psi_j'\rangle$ $(i, j = 1, 2, 3)$ are given by

$$\langle\psi_i'|\psi_j'\rangle = \langle\psi_i|\psi_j\rangle_{in} - \langle\phi_i|\phi_j\rangle. \tag{5.59}$$

We then have to find vectors $|\psi_i'\rangle$ that satisfy this equation. The answer is not unique, and one way of proceeding is the following. If we define the hermitian matrix L to be

$$L_{ij} = \langle\psi_i|\psi_j\rangle_{in} - \langle\phi_i|\phi_j\rangle, \tag{5.60}$$

then we note from Eq. (5.6) that $L_{12} = L_{13} = 0$. This implies that the simplest choice for $|\psi_1'\rangle$ is a vector with only one nonzero component. Then the vectors $|\psi_2'\rangle$ and $|\psi_3'\rangle$ will have nonzero components in only their other two places. The obvious choice is

$$|\psi_1'\rangle = \begin{pmatrix} \sqrt{p_1} \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{5.61}$$

In this column vector, the first entry is the amplitude of the photon to be in mode 1, the second is the amplitude to be in mode 2, etc. Mode 4 corresponds to the failure space, $\mathcal{A}$. The vectors $|\psi_2'\rangle$ and $|\psi_3'\rangle$ will have nonzero components in only their second and third places, and if their overlap is real, we can choose

$$|\psi_2'\rangle = \begin{pmatrix} 0 \\ \sqrt{p_2}\cos\theta \\ \sqrt{p_2}\sin\theta \\ 0 \end{pmatrix}, \quad |\psi_3'\rangle = \begin{pmatrix} 0 \\ \sqrt{p_3}\cos\theta \\ -\sqrt{p_3}\sin\theta \\ 0 \end{pmatrix}, \tag{5.62}$$

where

$$\theta = \frac{1}{2}\cos^{-1}\left(\frac{L_{23}}{\sqrt{p_2 p_3}}\right). \tag{5.63}$$

This simple choice works for the last example in this section (see Eq. (5.74), below). For the first, somewhat more general, example we are forced to choose the second component of $|\psi_1'\rangle$ to be nonzero and then the first and third components of the other two success vectors are different from zero. They can be obtained by simply

interchanging the first and second components in the above expressions of the vectors $|\psi_i'\rangle$ (see Eq. (5.65), below).

Once we have the input and output vectors, the unitary transformation, U, which maps the input states onto the output states then can be chosen, and this, as shown by Eq. (5.57), gives the explicit form of $M(4)$. Furthermore, as described by Eq. (2.36), $M(4)$ can be factorized as a product of two-dimensional $U(2)$ transformations which can be implemented by a lossless beam splitter and a phase shifter with appropriate parameters. Therefore, we can use appropriate beam splitters, phase shifters and a mirror to construct the desired eight-port.

Finally, photon detection is performed at the four output ports. We can design the total transformation in such a way that if the photon is detected at the first output port, we claim with certainty that the initial state was $|\psi_1\rangle$, if the photon is detected at the second or the third output port, we claim with certainty that the initial state was either $|\psi_2\rangle$ or $|\psi_3\rangle$, but we do not know which of these two states it was. If the photon is detected at the fourth output port, we obtain no information about the input state.

We shall now consider two examples. The first is more general than the second, but the second has the advantage that it is simple and the eight-port that it requires consists of only two $50-50$ beam splitters. In the first example, all of the input vectors have the same overlap, which is given by s, and we shall consider the case $0 < s \leq 1/\sqrt{2}$. The optimal failure probabilities for this case are given in Eq. (5.46).

For the input vectors we shall take

$$|\psi_1\rangle_{in} = \begin{pmatrix} \frac{1}{\sqrt{3}}(1+2s)^{1/2} \\ \sqrt{\frac{2}{3}}(1-s)^{1/2} \\ 0 \\ 0 \end{pmatrix},$$

$$|\psi_2\rangle_{in} = \begin{pmatrix} \frac{1}{\sqrt{3}}(1+2s)^{1/2} \\ -\frac{1}{\sqrt{6}}(1-s)^{1/2} \\ \frac{1}{\sqrt{2}}(1-s)^{1/2} \\ 0 \end{pmatrix},$$

$$|\psi_3\rangle_{in} = \begin{pmatrix} \frac{1}{\sqrt{3}}(1+2s)^{1/2} \\ -\frac{1}{\sqrt{6}}(1-s)^{1/2} \\ -\frac{1}{\sqrt{2}}(1-s)^{1/2} \\ 0 \end{pmatrix}. \tag{5.64}$$

The output vectors, $|\psi_i\rangle_{out} = |\psi_i'\rangle + |\phi_i\rangle$, can be computed by the method outlined

above. Doing so gives us

$$|\psi_1\rangle_{out} = \begin{pmatrix} 0 \\ (1-\sqrt{2}s)^{1/2} \\ 0 \\ (s\sqrt{2})^{1/2} \end{pmatrix},$$

$$|\psi_2\rangle_{out} = \begin{pmatrix} ((1+s-s\sqrt{2})/2)^{1/2} \\ 0 \\ ((1-s)/2)^{1/2} \\ (s/\sqrt{2})^{1/2} \end{pmatrix},$$

$$|\psi_3\rangle_{out} = \begin{pmatrix} ((1+s-s\sqrt{2})/2)^{1/2} \\ 0 \\ -((1-s)/2)^{1/2} \\ (s/\sqrt{2})^{1/2} \end{pmatrix}. \tag{5.65}$$

Our next step is to determine the transformation, U, that describes the eight-port, or, more specifically, the matrix $M(4)$ that describes its action in the one-photon subspace. It must satisfy $|\psi_i\rangle_{out} = U|\psi\rangle_{in}$, and, in addition, it must map the vector that is orthogonal to all three input vectors, onto the vector that is orthogonal to all three output vectors,

$$\frac{1}{A}\begin{pmatrix} -(s\sqrt{2})^{1/2}B \\ -(s\sqrt{2})^{1/2}C \\ 0 \\ BC \end{pmatrix} = M(4)\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \tag{5.66}$$

where

$$\begin{aligned} A &= [(1-s)(1+2s)]^{1/2}, \\ B &= (1-s\sqrt{2})^{1/2}, \\ C &= (1+s-s\sqrt{2})^{1/2}. \end{aligned} \tag{5.67}$$

These equations determine M(4) and it is given by $M(4) =$

$$\begin{pmatrix} \sqrt{\frac{2}{3}}\frac{C}{\sqrt{1+2s}} & -\frac{C}{\sqrt{3(1-s)}} & 0 & -\frac{B}{A}(s\sqrt{2})^{1/2} \\ \frac{B}{\sqrt{3(1+2s)}} & \sqrt{\frac{2}{3}}\frac{B}{\sqrt{1-s}} & 0 & -\frac{C}{A}(s\sqrt{2})^{1/2} \\ 0 & 0 & 1 & 0 \\ \frac{(\sqrt{2}+1)(s\sqrt{2})^{1/2}}{\sqrt{3(1+2s)}} & \frac{(\sqrt{2}-1)(s\sqrt{2})^{1/2}}{\sqrt{3(1-s)}} & 0 & \frac{BC}{A} \end{pmatrix}. \tag{5.68}$$

This matrix can be expressed as the product of three matrixes each of which corresponds to a beam splitter. In particular, we have that

$$M(4) = T_{2,4}T_{1,4}T_{1,2} \tag{5.69}$$

where the matrix $T_{p,q}$ represents the action of a beam splitter that mixes only modes p and q. The 4×4 matrix for $T_{p,q}$ can be obtained from that of a 4×4 identity matrix, I, by replacing the matrix elements I_{pp} and I_{qq} by the transmissivity of the beam splitter, t, replacing I_{pq} by the reflectivity, r, and replacing I_{qp} by $-r$. The transmissivities and reflectivities for beam splitters in Eq. (5.69) are

$$\begin{aligned} T_{2,4}: &\quad t = B && r = -(s\sqrt{2})^{1/2} \\ T_{1,4}: &\quad t = \tfrac{C}{A} && r = -(s\sqrt{2})^{1/2}\tfrac{B}{A} \\ T_{1,2}: &\quad t = \sqrt{\tfrac{2(1-s)}{3}} && r = -\sqrt{\tfrac{1+2s}{3}}. \end{aligned} \tag{5.70}$$

This constitutes a complete description of the optical network that optimally discriminates between $|\psi_1\rangle_{in}$ and $\{|\psi_2\rangle_{in}, |\psi_3\rangle_{in}\}$, where these input states are given in Eq. (5.64), and it is shown schematically in Figure 5.

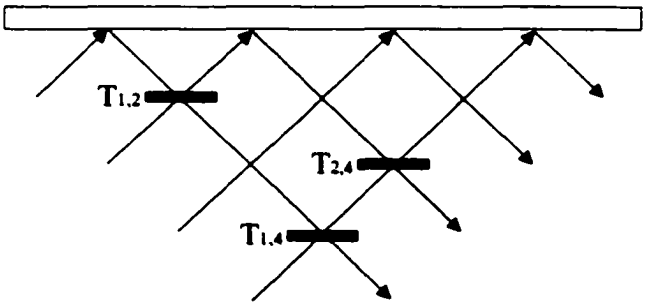

FIG. 5: The eight-port described by Eq. (5.68) can be constructed from three beam splitters and a mirror.

An especially simple network will suffice for our second example. The input vectors are

$$|\psi_1\rangle_{in} = \begin{pmatrix} \sqrt{2/3} \\ 0 \\ 1/\sqrt{3} \\ 0 \end{pmatrix}, \quad |\psi_2\rangle_{in} = \begin{pmatrix} 0 \\ 1/\sqrt{3} \\ \sqrt{2/3} \\ 0 \end{pmatrix}, \quad |\psi_3\rangle_{in} = \begin{pmatrix} 0 \\ -1/\sqrt{3} \\ \sqrt{2/3} \\ 0 \end{pmatrix}. \tag{5.71}$$

These input states have the property that

$$\begin{aligned} {}_{in}\langle\psi_1|\psi_2\rangle_{in} &= {}_{in}\langle\psi_1|\psi_3\rangle_{in} = \frac{\sqrt{2}}{3}, \\ {}_{in}\langle\psi_2|\psi_3\rangle_{in} &= \frac{1}{3}. \end{aligned} \tag{5.72}$$

The optimal failure probabilities are found to be $q_1 = 2/3$ and $q_2 = q_3 = 1/3$. Using

Eqs. (5.50) and (5.51) this gives

$$Q = \frac{4}{9}, \tag{5.73}$$

for the minimum average failure probability of this kind of generalized measurement. This is to be compared to 5/9, the average failure probability of a von Neumann type projective measurement.

The output vectors, $|\psi_i\rangle_{out} = |\psi_i'\rangle + |\phi_i\rangle$, can again be computed by the method outlined previously. Doing so gives us

$$|\psi_1\rangle_{out} = \begin{pmatrix} 1/\sqrt{3} \\ 0 \\ 0 \\ \sqrt{2/3} \end{pmatrix}, \; |\psi_2\rangle_{out} = \begin{pmatrix} 0 \\ 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}, \; |\psi_3\rangle_{out} = \begin{pmatrix} 0 \\ -1/\sqrt{3} \\ 1/\sqrt{3} \\ (1/\sqrt{3} \end{pmatrix}. \tag{5.74}$$

The matrix $M(4)$ can be chosen to be

$$M(4) = \begin{pmatrix} 1/\sqrt{2} & 0 & 0 & -1/\sqrt{2} \\ 0 & 1 & 0 & 0 \\ -1/2 & 0 & 1/\sqrt{2} & -1/2 \\ 1/2 & 0 & 1/\sqrt{2} & 1/2 \end{pmatrix}, \tag{5.75}$$

and it can be expressed as

$$M(4) = T_{3,4}T_{1,4}. \tag{5.76}$$

In this case, both $T_{1,4}$ and $T_{3,4}$ represent $50-50$ beam splitters, and they are given

explicitly by

$$\begin{aligned} T_{1,4}: \quad & t = \tfrac{1}{\sqrt{2}} \quad r = -\frac{1}{\sqrt{2}} \\ T_{3,4}: \quad & t = \tfrac{1}{\sqrt{2}} \quad r = -\frac{1}{\sqrt{2}}. \end{aligned} \tag{5.77}$$

This last example constitutes what is probably the simplest choice of the set of parameters for a possible experimental realization.

Chapter 6

Unambiguous quantum states comparison

6.1. Introduction

In this section we consider an interesting problem. Assume there are two similar quantum systems A and B, each of which could be in either quantum state $|\Psi_1\rangle$ or $|\Psi_2\rangle$, which, in general, are nonorthogonal. We want to determine whether the two systems A and B are in the same states or different.

In classical physics, such a problem can be easily solved by measuring the two systems separately and then comparing the results. In quantum physics, however, since it is about to extract information regarding two nonorthogonal states $|\Psi_1\rangle$ and $|\Psi_2\rangle$, the comparison is more subtle. Actually, such a problem can be treated as that a composite system is known to be in one of the four composite states

$$\begin{aligned} |\psi_1\rangle &= |\Psi_1\rangle_A|\Psi_2\rangle_B, \\ |\psi_2\rangle &= |\Psi_2\rangle_A|\Psi_1\rangle_B, \\ |\psi_3\rangle &= |\Psi_1\rangle_A|\Psi_1\rangle_B, \\ |\psi_4\rangle &= |\Psi_2\rangle_A|\Psi_2\rangle_B, \end{aligned} \tag{6.1}$$

we want to determine whether the state of the composite system belongs to

$\{|\psi_1\rangle, |\psi_2\rangle\}$ or $\{|\psi_3\rangle, |\psi_4\rangle\}$. That is, we want to distinguish between two sets of quantum states. Our aim is to derive the optimal result of this discrimination and to show that it is always better than to distinguish all the four states, because it requires less information. In section 3, it is found that to unambiguously distinguish two nonorthogonal states $|\Psi_1\rangle$ and $|\Psi_2\rangle$, with η_1 and η_2 as the *a priori* probabilities, the minimum probability of obtaining inclusive result is [5, 20]

$$Q = 2\sqrt{\eta_1\eta_2}|\langle\Psi_1|\Psi_2\rangle|. \tag{6.2}$$

Later, we will compare the result of states comparison to this.

6.2. Distinguish all the states

First we shall derive the optimal result to distinguish all the four states. To optimally distinguish $|\psi_i\rangle$, we shall use is a "generalized measurement". [28] Let $\mathcal{K}$ denote a total Hilbert space, which is the direct sum of two subspaces, $\mathcal{K} = \mathcal{H} \oplus \mathcal{A}$. The space $\mathcal{H}$ is the space that contains the vectors $|\psi_i\rangle$, and $\mathcal{A}$ is an auxiliary space. The input state of the system is one of the vectors $|\psi_i\rangle$, which is now a vector in the subspace $\mathcal{H}$ of the total space $\mathcal{K}$, so that

$$|\psi_i^{\mathcal{K}}\rangle_{in} = |\psi_i^{\mathcal{H}}\rangle. \tag{6.3}$$

A unitary transformation, U, which acts in the entire space $\mathcal{K}$ is now applied to the input vector, resulting in the state $|\psi_i^{\mathcal{K}}\rangle_{out}$, which is given by

$$|\psi_i^{\mathcal{K}}\rangle_{out} = |\psi_i'^{\mathcal{H}}\rangle + |\phi_i^{\mathcal{A}}\rangle = U|\psi_i^{\mathcal{K}}\rangle_{in}, \tag{6.4}$$

where $|\psi_i'\rangle$ is the desired state, it always fulfill the condition imposed by the desired unambiguous discrimination. Then a measurement is performed on $|\psi_i^{\mathcal{K}}\rangle_{out}$ that projects $|\psi_i^{\mathcal{K}}\rangle_{out}$ either onto $|\psi_i'\rangle$ or $|\phi_i\rangle$ (by construction, they are in orthogonal subspaces). If it projects $|\psi_i^{\mathcal{K}}\rangle_{out}$ onto $|\psi_i'\rangle$, the procedure succeeds, the probability to get this outcome, if the input state is $|\psi_i\rangle$, is

$$p_i = \langle\psi_i'|\psi_i'\rangle. \tag{6.5}$$

If the measurement projects $|\psi_i^{\mathcal{K}}\rangle_{out}$ onto $|\phi_i\rangle$, the procedure fails. The probability of this outcome is

$$q_i = 1 - p_i = \langle\phi_i|\phi_i\rangle. \tag{6.6}$$

The above procedure is a nonunitary transformation that transforms $|\psi_i\rangle$ into $|\psi_i'\rangle$ with a certain probability of failure. The requirements that $|\psi_i'\rangle$ can always be distinguishable implies that

$$\langle\psi_i'|\psi_j'\rangle = 0. \qquad \text{for i} \neq \text{j}. \tag{6.7}$$

These lead to conditions on the failure vectors, $|\phi_i\rangle$. Taking the scalar product between $\langle\psi_i^{\mathcal{K}}|\psi_j^{\mathcal{K}}\rangle_{out}$ and using Eq. (6.7) and the fact that U is unitary leads to the conditions

$$\langle\phi_i|\phi_j\rangle = 0. \qquad \text{for i} \neq \text{j}. \tag{6.8}$$

Our objective is to find the optimum $|\psi_i'\rangle$ and $|\phi_i\rangle$ which satisfy Eqs. (6.5)– (6.8) and also give the maximum success probability P.

Starting from the fact that for optimum discrimination, the vectors $|\phi_i\rangle$ must be linearly dependent [25], we have

$$\det(C) = \begin{vmatrix} \langle\phi_1|\phi_1\rangle & \langle\phi_1|\phi_2\rangle & \langle\phi_1|\phi_3\rangle & \langle\phi_1|\phi_4\rangle \\ \langle\phi_2|\phi_1\rangle & \langle\phi_2|\phi_2\rangle & \langle\phi_2|\phi_3\rangle & \langle\phi_2|\phi_4\rangle \\ \langle\phi_3|\phi_1\rangle & \langle\phi_3|\phi_2\rangle & \langle\phi_3|\phi_3\rangle & \langle\phi_3|\phi_4\rangle \\ \langle\phi_4|\phi_1\rangle & \langle\phi_4|\phi_2\rangle & \langle\phi_4|\phi_3\rangle & \langle\phi_4|\phi_4\rangle \end{vmatrix} = 0. \tag{6.9}$$

Substituting Eqs. (6.6) and (6.8) into Eq. (6.9), let $S = \langle\Psi_1|\Psi_2\rangle$, we obtain

$$\Delta = \det(C) = \begin{vmatrix} q_1 & |S|^2 & S^* & S \\ |S|^2 & q_2 & S^* & S \\ S & S & q_3 & |S|^2 \\ S^* & S^* & |S|^2 & q_4 \end{vmatrix} = 0, \tag{6.10}$$

Since C is positive semidefinite, all the subdeterminants of Δ must be non-negative.

Denote the average probability of failure as

$$Q = \sum_{i=1}^{4} \eta_i q_i. \tag{6.11}$$

where η_i are the *a priori* probabilities. Taking the constraint, Eq. (6.10), into account, using Lagrangian multiplier method, we can rewrite Q as

$$Q = \sum_i \eta_i q_i + \lambda\Delta, \tag{6.12}$$

where λ is a Lagrangian multiplier.

We wish to minimize Q, this leads to the conditions

$$\frac{\partial Q}{\partial q_i} = \eta_i + \lambda\frac{\partial \Delta}{\partial q_i} = 0, \qquad i = \{1 \ldots 4\}, \tag{6.13}$$

To simplify Eq. (6.13), we can use the following theorem on the derivative of a determinant [44]

Theorem 1 *Let a_{ij} be differentiable functions of $\lambda (i, j = 1, \ldots, n)$. Then*

$$(\frac{d}{d\lambda})\Delta(\lambda) = \sum_{k=1}^{n} \Delta_k(\lambda) \tag{6.14}$$

where $\Delta(\lambda) = det[a_{ij}(\lambda)]$, and where $\Delta_k(\lambda)$ is the determinant formed by replacing the kth row $a_{kj}(\lambda)$ $(j = 1, \ldots, n)$ by the row of derivatives $a'_{kj}(\lambda)$ $(j = 1, \ldots, n)$.

Using the above theorem, Eq. (6.13) leads to

$$\eta_i + \lambda \Delta_{ii}(3) = 0, \qquad i = \{1 \ldots 4\}, \tag{6.15}$$

where $\Delta_{ii}(3)$ is the 3×3 determinant formed by eliminating row and column i from Δ.

Eq. (6.10) can be rewritten explicitly as

$$(q_1 q_2 - |S|^4)(q_3 q_4 - |S|^4) - |S|^2(q_1 + q_2 - 2|S|^2)(q_3 + q_4 - 2|S|^2). \tag{6.16}$$

Eq. (6.15) can be rewritten explicitly as

$$q_2(q_3 q_4 - |S|^4) - |S|^2(q_3 + q_4 - 2|S|^2) = -\frac{\eta_1}{\lambda}, \tag{6.17}$$

$$q_1(q_3 q_4 - |S|^4) - |S|^2(q_3 + q_4 - 2|S|^2) = -\frac{\eta_2}{\lambda}, \tag{6.18}$$

$$q_4(q_1 q_2 - |S|^4) - |S|^2(q_1 + q_2 - 2|S|^2) = -\frac{\eta_3}{\lambda}, \tag{6.19}$$

$$q_3(q_1 q_2 - |S|^4) - |S|^2(q_1 + q_2 - 2|S|^2) = -\frac{\eta_4}{\lambda}. \tag{6.20}$$

Let Eq. (6.17), Eq. (6.18) multiply by $(q_1q_2 - |S|^4)$, and Eq. (6.19), Eq. (6.20) multiply by $(q_3q_4 - |S|^4)$, taking Eq. (6.16) into account, we have that

$$|S|^2(q_3 + q_4 - 2|S|^2)(q_2 - |S|^2)^2 = -\frac{\eta_1}{\lambda}(q_1q_2 - |S|^4), \tag{6.21}$$

$$|S|^2(q_3 + q_4 - 2|S|^2)(q_1 - |S|^2)^2 = -\frac{\eta_2}{\lambda}(q_1q_2 - |S|^4), \tag{6.22}$$

$$|S|^2(q_1 + q_2 - 2|S|^2)(q_3 - |S|^2)^2 = -\frac{\eta_3}{\lambda}(q_3q_4 - |S|^4), \tag{6.23}$$

$$|S|^2(q_1 + q_2 - 2|S|^2)(q_4 - |S|^2)^2 = -\frac{\eta_4}{\lambda}(q_3q_4 - |S|^4), \tag{6.24}$$

which implies that

$$\eta_1(q_1 - |S|^2)^2 = \eta_2(q_2 - |S|^2)^2, \tag{6.25}$$

$$\eta_3(q_3 - |S|^2)^2 = \eta_4(q_4 - |S|^2)^2. \tag{6.26}$$

The constraint that matrix C is positive semidefinite requires that

$$q_i > |S|^2, \tag{6.27}$$

which leads to

$$\sqrt{\eta_1}(q_1 - |S|^2) = \sqrt{\eta_2}(q_2 - |S|^2), \tag{6.28}$$

$$\sqrt{\eta_3}(q_3 - |S|^2) = \sqrt{\eta_4}(q_4 - |S|^2). \tag{6.29}$$

Using the relations given by Eq. (6.28), (6.29), we can solve the equations and find the solutions for q_i, among which only one satisfies the condition given by Eq. (6.27),

it is

$$q_1 = \frac{(\sqrt{\eta_3} + \sqrt{\eta_4})|S| - \sqrt{\eta_2}|S|^2}{\sqrt{\eta_1}},$$
$$q_2 = \frac{(\sqrt{\eta_3} + \sqrt{\eta_4})|S| - \sqrt{\eta_1}|S|^2}{\sqrt{\eta_2}},$$
$$q_3 = \frac{(\sqrt{\eta_1} + \sqrt{\eta_2})|S| - \sqrt{\eta_4}|S|^2}{\sqrt{\eta_3}},$$
$$q_4 = \frac{(\sqrt{\eta_1} + \sqrt{\eta_2})|S| - \sqrt{\eta_3}|S|^2}{\sqrt{\eta_4}}. \tag{6.30}$$

Correspondingly, the probabilities of failure and success are

$$Q = 2(\sqrt{\eta_1} + \sqrt{\eta_2})(\sqrt{\eta_3} + \sqrt{\eta_4})|S|$$
$$- 2(\sqrt{\eta_1}\sqrt{\eta_2} + \sqrt{\eta_3}\sqrt{\eta_4})|S|^2,$$
$$P = 1 - 2(\sqrt{\eta_1} + \sqrt{\eta_2})(\sqrt{\eta_3} + \sqrt{\eta_4})|S|$$
$$+ 2(\sqrt{\eta_1}\sqrt{\eta_2} + \sqrt{\eta_3}\sqrt{\eta_4})|S|^2. \tag{6.31}$$

If we denote $\{\alpha_1, \alpha_2\}$ the *a priori* probabilities that system A was in the state $\{\psi_1, \psi_2\}$, and $\{\beta_1, \beta_2\}$ the *a priori* probabilities that system B was in the state $\{\psi_1, \psi_2\}$, observing that

$$\eta_1 = \alpha_1\beta_2,$$
$$\eta_2 = \alpha_2\beta_1,$$
$$\eta_3 = \alpha_1\beta_1,$$
$$\eta_4 = \alpha_2\beta_2, \tag{6.32}$$

the probability of success can be written as

$$P = (1 - 2|S|\sqrt{\alpha_1\alpha_2})(1 - 2|S|\sqrt{\beta_1\beta_2}). \tag{6.33}$$

From Eq. (6.2) we see that P is a product of two optimal successful probabilities corresponding to two states discrimination in system A and B. This implies that the optimal discrimination of the four states can be achieved by doing optimal discrimination in the two system A and B separately.

6.3. States Comparison

In this section we shall derive the optimal solution to unambiguously distinguish $\{|\psi_1\rangle, |\psi_2\rangle\}$ and $\{|\psi_3\rangle, |\psi_4\rangle\}$. This imposes that

$$\begin{aligned} \langle\psi_1'|\psi_3'\rangle &= \langle\psi_1'|\psi_4'\rangle = 0, \\ \langle\psi_2'|\psi_3'\rangle &= \langle\psi_2'|\psi_4'\rangle = 0. \end{aligned} \tag{6.34}$$

Correspondingly, for optimum discrimination, Eq. (6.10) shall be modified to be

$$C = \begin{vmatrix} q_1 & x & S^* & S \\ x^* & q_2 & S^* & S \\ S & S & q_3 & y \\ S^* & S^* & y^* & q_4 \end{vmatrix} = 0, \tag{6.35}$$

where $x = r_x e^{i\theta_x}$ and $y = r_y e^{i\theta_y}$ are parameters needed to be determined.

We can write the Lagrangian multiplier equation on the average probability of failure Q as

$$Q = \sum_i \eta_i q_i + \lambda C, \tag{6.36}$$

and get the following equations to minimize Q,

$$\frac{\partial Q}{\partial q_i} = 0, \qquad i = \{1 \ldots 4\}, \tag{6.37}$$

$$\frac{\partial Q}{\partial r_x} = \frac{\partial C}{\partial r_x} = 0, \tag{6.38}$$

$$\frac{\partial Q}{\partial \theta_x} = \frac{\partial C}{\partial \theta_x} = 0, \tag{6.39}$$

$$\frac{\partial Q}{\partial r_y} = \frac{\partial C}{\partial r_y} = 0, \tag{6.40}$$

$$\frac{\partial Q}{\partial \theta_y} = \frac{\partial C}{\partial \theta_y} = 0. \tag{6.41}$$

Let us denote Δ_{lm} the 3×3 determinant formed by eliminating row l and column m from C. Using Theorem 1, Eq. (6.38), (6.39) lead to

$$e^{i\theta_x}\Delta_{12} + e^{-i\theta_x}\Delta_{21} = 0, \tag{6.42}$$

$$ir_x e^{i\theta_x}\Delta_{12} - ir_x e^{-i\theta_x}\Delta_{21} = 0, \tag{6.43}$$

which implies that when $r_x \neq 0$,

$$\Delta_{12} = \Delta_{21} = 0. \tag{6.44}$$

By observing $\Delta_{lm} = \Delta^*_{ml}$, Eq. (6.44) is actually one independent equation.

Similarly, from Eq. (6.40), (6.41), we can get

$$\Delta_{34} = \Delta_{43} = 0. \tag{6.45}$$

Again, Eq. (6.45) is just one independent equation.

Now we shall use the following theorem on determinant

Theorem 2 *For a $n \times n$ determinant $\Delta(n)$ and a $(n-1) \times (n-1)$ determinant $\Delta(n-1)_{lm}$, where $\Delta(n-1)_{lm}$ is formed by eliminating row l and column m from Δ, if given*

$$\Delta(n) = 0, \tag{6.46}$$

$$\Delta(n-1)_{lm} = 0. \tag{6.47}$$

Then, for $k = 1, \ldots, n$,

$$\textit{either} \qquad \Delta(n-1)_{lk} = 0, \tag{6.48}$$

$$\textit{or} \qquad \Delta(n-1)_{km} = 0. \tag{6.49}$$

To prove the theorem, we can write Eq. (6.46) explicitly as

$$\Delta = \begin{vmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{vmatrix} = 0. \tag{6.50}$$

Define vectors

$$\vec{r}_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix}, \qquad (i = 1, \ldots, n), \tag{6.51}$$

and

$$\vec{r_i}' = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{l-1i} \\ a_{l+1i} \\ \vdots \\ a_{ni} \end{pmatrix}, \qquad (i = 1, \ldots, n), \tag{6.52}$$

that is, $\vec{r_i}$ is formed by the ith column of $\Delta(n)$, and $\vec{r_i}'$ is formed by the ith column of $\Delta(n)$ without its lth row element.

Eq. (6.46) implies that for some constants μ_i

$$\sum_{i=1}^{n} \mu_i \vec{r_i} = 0, \qquad (\text{not all } \mu_i = 0). \tag{6.53}$$

So that

$$\sum_{i=1}^{n} \mu_i \vec{r_i}' = 0, \qquad (\text{not all } \mu_i = 0). \tag{6.54}$$

Eq. (6.47) implies that for some constants μ'_k, $k \neq m$,

$$\sum_{k=1}^{n} \mu'_k \vec{r_i}' = 0, \qquad (\text{not all } \mu'_k = 0). \tag{6.55}$$

For arbitrary k, if $\mu'_k = 0$, then from Eq. (6.55) we can get that $\vec{r_1}'$, ..., $\vec{r_{k-1}}'$, $\vec{r_{k+1}}'$,...,$\vec{r_n}'$ are linearly dependent. Therefore,

$$\Delta(n-1)_{lk} = 0. \tag{6.56}$$

If $\mu'_k \neq 0$, we can rewrite Eq. (6.55) as

$$\vec{r_k}' = -(\frac{\mu'_1}{\mu'_k}\vec{r_1}' + \ldots + \frac{\mu'_{m-1}}{\mu'_k}\vec{r_{m-1}}' + \frac{\mu'_{m+1}}{\mu'_k}\vec{r_{m+1}}' + \frac{\mu'_n}{\mu'_k}\vec{r_n}') \tag{6.57}$$

and substitute it into Eq. (6.54) to get

$$\nu_1\vec{r_1}' + \ldots + \nu_{k-1}\vec{r_{k-1}}' + \nu_{k+1}\vec{r_{k+1}}' + \ldots + \nu_{m-1}\vec{r_{m-1}}'$$
$$+ \mu_m\vec{r_m}' + \nu_{m+1}\vec{r_{m+1}}' + \ldots + \nu_n\vec{r_n}' = 0, \tag{6.58}$$

where $\nu_i = (\mu_i - \mu_k\mu_i'/\mu_k')$.

If $\mu_m \neq 0$, Eq. (6.58) implies that $\vec{r_1}', \ldots, \vec{r_{k-1}}', \vec{r_{k+1}}', \ldots, \vec{r_n}'$ are linearly dependent, which leads to the same as Eq. (6.56).

If $\mu_m = 0$, then Eq. (6.53) becomes for $\mu_i \neq m$

$$\sum_{i=1}^{n} \mu_i\vec{r_i} = 0, \qquad \text{(not all } \mu_i = 0), \tag{6.59}$$

which leads to for all $k = 1, \ldots, n$

$$\Delta(n-1)_{km} = 0. \tag{6.60}$$

Eq. (6.56) and Eq. (6.60) complete the proof.

In our case, using theorem 2 and the fact that $\Delta_{ij} = \Delta_{ji}^*$, combining Eq. (6.35) with Eq. (6.44) leads to either

$$\Delta_{11} = \Delta_{12} = \Delta_{13} = \Delta_{14} = 0, \tag{6.61}$$

or

$$\Delta_{21} = \Delta_{22} = \Delta_{23} = \Delta_{24} = 0. \tag{6.62}$$

Combining Eq. (6.35) with Eq. (6.45) leads to either

$$\Delta_{31} = \Delta_{32} = \Delta_{33} = \Delta_{34} = 0, \tag{6.63}$$

or

$$\Delta_{41} = \Delta_{42} = \Delta_{43} = \Delta_{44} = 0. \tag{6.64}$$

Both Eq. (6.61) and Eq. (6.62) impose that row 3 of Δ is proportional to row 4 by a factor of $e^{-2i\theta}$, where $\theta = \arg(S)$. This leads to

$$q_3 = q_4 = |y|, \tag{6.65}$$

$$\arg(y) = 2\theta. \tag{6.66}$$

Similarly, both Eq. (6.63) and Eq. (6.64) impose that row 1 of Δ is identical to row 2. This leads to

$$q_1 = q_2 = x = |x| \tag{6.67}$$

Hence, the average probability of failure is

$$Q = (\eta_1 + \eta_2)|x| + (\eta_3 + \eta_4)|y|. \tag{6.68}$$

The constraint that C is positive semidefinite, which means all the subdeterminants of Δ must be non-negative, implies that to get the minimum Q,

$$\begin{vmatrix} q_2 & S^* \\ S & q_3 \end{vmatrix} = \begin{vmatrix} |x| & S^* \\ S & |y| \end{vmatrix} = 0. \tag{6.69}$$

So that

$$|x||y| = |S|^2. \tag{6.70}$$

Substitute Eq. (6.70) into Eq. (6.68), we get

$$Q = (\eta_1 + \eta_2)|x| + (\eta_3 + \eta_4)\frac{|S|^2}{|x|}. \tag{6.71}$$

To get the minimum Q, let $dQ/d|x| = 0$, and we get the optimum solution of q_i as

$$q_1 = q_2 = \sqrt{\frac{\eta_3 + \eta_4}{\eta_1 + \eta_2}}|S|,$$
$$q_3 = q_4 = \sqrt{\frac{\eta_1 + \eta_2}{\eta_3 + \eta_4}}|S|. \tag{6.72}$$

Correspondingly, the average probability of failure is

$$Q = \sum_{i=1}^{4} \eta_i q_i = 2\sqrt{(\eta_1 + \eta_2)(\eta_3 + \eta_4)}|S|. \tag{6.73}$$

An interesting point is that if we note that $\eta_1 + \eta_2$ is the *a priori* probability that the two systems are in different states, and $\eta_3 + \eta_4$ is the *a priori* probability that the two systems are in the same states, then the above result reduces to Eq. (6.2).

Now we shall show that the average probability of failure of Eq. (6.73) is always smaller than that of to distinguish all the states. That is, using Eq. (6.31) and Eq. (6.73), we shall prove that

$$[2(\sqrt{\eta_1} + \sqrt{\eta_2})(\sqrt{\eta_3} + \sqrt{\eta_4}) - \sqrt{(\eta_1 + \eta_2)(\eta_3 + \eta_4)}]|S|$$
$$- 2(\sqrt{\eta_1}\sqrt{\eta_2} + \sqrt{\eta_3}\sqrt{\eta_4})|S|^2 > 0. \tag{6.74}$$

Since

$$|S| > |S|^2, \tag{6.75}$$

all we need to show is that

$$(\sqrt{\eta_1} + \sqrt{\eta_2})(\sqrt{\eta_3} + \sqrt{\eta_4}) - \sqrt{(\eta_1 + \eta_2)(\eta_3 + \eta_4)}$$
$$- (\sqrt{\eta_1}\sqrt{\eta_2} + \sqrt{\eta_3}\sqrt{\eta_4}) > 0. \tag{6.76}$$

To prove the above inequality, actually we show that

$$\begin{aligned} Diff =&[(\sqrt{\eta_1}+\sqrt{\eta_2})(\sqrt{\eta_3}+\sqrt{\eta_4})-(\sqrt{\eta_1}\sqrt{\eta_2}+\sqrt{\eta_3}\sqrt{\eta_4})]^2 \\ &-(\eta_1+\eta_2)(\eta_3+\eta_4)>0. \end{aligned} \tag{6.77}$$

Using the notation from Eq. (6.32), let

$$\begin{aligned} \alpha_1 &= \sin^2\omega_1, \\ \alpha_2 &= \cos^2\omega_1, \\ \beta_1 &= \sin^2\omega_2, \\ \beta_2 &= \cos^2\omega_2, \end{aligned} \tag{6.78}$$

where $\omega_{1,2}\in[0,\pi/2]$, and substitute this into Eq. (6.77), we can get

$$Diff=\frac{1}{2}\sin 2\omega_1\sin 2\omega_2(1-\sin 2\omega_1)(1-\sin 2\omega_2). \tag{6.79}$$

So $Diff$ is always positive.

We have considered the problem to unambiguously compare two quantum systems, each of which could be in one of two nonorthogonal quantum states. We have derived the optimal analytical solution of this problem. In our case, there is no entanglement between the two systems, so the composite states are pure product states. An interesting extension is when there is entanglement between the systems, this deserves further investigation.

Chapter 7

Conclusions

Discrimination of quantum states is a fundamental problem in quantum physics. If the quantum states are nonorthogonal, they can not be discriminated with unit success probability. Two different strategies exist: minimum-error discrimination and unambiguous discrimination. The later one is error-free, but inconclusive result is inevitable. Optimum unambiguous discrimination means to have the minimum probability of getting inconclusive result.

In order to achieve optimum unambiguous discrimination, we need to use generalized measurement, which can be described by POVM. We have shown the general POVM formulism to derive optimum unambiguous discrimination. A necessary condition is given. It can be used as a constraint to derive the minimum probability of getting inconclusive result. We have also provided an optical implementation of any desired unambiguous discrimination. We have shown that nonorthogonal quantum states, each realized as a photon split among several modes, can be conditionally distinguished by means of a linear optical network.

We have discussed in detail on how to discriminate two and three nonorthogonal quantum states. We have given explicit examples together with the optical networks,

which give the maximum success probabilities for several sets of states. An application of unambiguous discrimination, quantum entanglement enhancement, was discussed. In addition, it was shown that the addition of a second network to the outputs corresponding to a failure of the initial network to distinguish the states, can sometimes provide partial information about the input state.

We have also considered the problem of unambiguous discrimination between subsets. The set of possible states is divided into two subsets, and we only want to know to which subset the quantum state of our given system belongs. As this is a less ambitious task than actually identifying the state, we expect that our probability to be successful will be greater for attaining this more limited goal.

We considered the simplest instance of this problem, the situation in which we are trying to discriminate between a set containing one quantum state and another containing two. A method for finding the optimal strategy for discriminating between these two sets was presented, and analytical solutions for particular cases were given. In addition, we have shown that if the quantum states are single-photon states, where the photon can be split among several modes, the optimal discrimination strategy can be implemented by using a linear optical network.

The above problem is a perfect example of unambiguous discrimination between two mixed states. Another case: state comparison has also been studied. We have presented the optimum solution and showed its relation to the IDP solution.

There are still many interesting problems in this field, such as optimum unambiguous discrimination between mixed states, the applications of mixed states discrimi-

nation. More detailed consideration of these problems remains for future research.

Bibliography

[1] C. H. Bennett , phys. Rev. lett. **68**, 3121 (1992).

[2] M. Hillery and V. Bužek, *Phys. Rev. A* **56** 1212 (1997).

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

[4] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A **53**, 2046 (1996).

[5] J. A. Bergou, M. Hillery and Y. Sun, J. Mod. Opt. **47**, 487 (2000).

[6] C. H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres, and W. K. Wooters, Phys. Rev. Lett. **70**, 1895 (1993).

[7] M. A. Nielsen and I.L. Chuang, *Quantum Computation and Information* (Cambridge University Press, 2000), See page 87.

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

[9] A. S. Holevo, J. Multivar. Anal. **3** 337 (1973).

[10] H. P. Yuen, R. S. Kennedy, and M. Lax, IEEE Trans. Inf. Theory bf IT-21, 125 (1975).

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

[12] M. Sasaki, K. Kato, M. Izutsu, and O. Hirota, Phys. Rev. A **58**, 146 (1998).

[13] S. M. Barnett, Phys. Rev. A **64**, 030303 (2001).

[14] E.Andersson, S. M. Barnett, C. B. Gilson, and K. Hunter, quant-ph/0201074.

[15] U. Herzog and J. A. Bergou, Phys. Rev. A **65**, 050305 (2002).

[16] R. B. M. Clarke, V. M. Kendon, A. Chefles, S. M. Barnett, E. Riis, and M. Sasaki, Phys. Rev. A **64**, 012303 (2001).

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

[18] D. Dieks, Phys. Lett. A **126**, 303 (1988).

[19] A. Peres, Phys. Lett. A **128**, 19 (1988).

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

[21] B. Huttner, A. Muller, J. D. Gautier, H. Zbinden, and N. Gisin, Phys. Rev. A **54**, 3783 (1996).

[22] R. B. M. Clarke, A. Chefles, S. M. Barnett, and E. Riis, LANL arXiv, quant-ph/0007063.

[23] A. Chefles and S. M. Barnett, J. Mod. Opt. **45**, 1295 (1998).

[24] A. Peres and D. Terno, J. Phys. A **31**, 7105 (1998).

[25] A. Chefles, Phys. Lett. **A239**, 339 (1998).

[26] A. Chefles and S. M. Barnett, Phys. Lett. **A250**, 223 (1998).

[27] L-M. Duan and G-C. Guo, Phys. Lett. **A261**, 25 (1999).

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

[29] Y. Sun, J. A. Bergou, and M. Hillery, Phys. Rev. A **66**, 032315 (2002), quant-ph/0112051.

[30] J. A. Bergou, U. Herzog and M. Hillery, Phys. Rev. Lett. **90**, 257901 (2002), quant-ph/020907.

[31] T. Rudolph, R. W. Spekkens and P. S. Turner, arXiv:quant-ph/0303071 (2003)

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

[33] Y.C. Eldar, Phys. Rev. A **67**, 042309 (2003)

[34] P. Raynal and N. Ltkenhaus, arXiv:quant-ph/0304179 (2003)

[35] J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, (Princeton Univercity Press, 1935).

[36] K. Kraus, *States, Effects and Operations: Fundamental Notions of Quantum Theory* (Springer-Verlag, Berlin, 1983).

[37] M. A. Naimark, *Izv. Akad. Nauk. SSSR Ser. Mat.* **4** (1940).

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

[39] G. Milburn, Phys. Rev. Lett. **62**, 2124 (1988).

[40] Isaac L. Chuang and Yoshihisa Yamamoto, Phys. Rev. A **52**, 3489 (1995).

[41] Isaac L. Chuang and Yoshihisa Yamamoto, Phys. Rev. Lett **76**, 4281 (1996).

[42] J. A. Bergou, M. Hillery and Y. Sun, Fortschritte der Physik **49**, 915 (2001).

[43] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A **53**, 2046 (1996).

[44] Joel N. Franklin, *Matrix Theory* (Dover 0-486-41179-6), p. 23.