11 Discrimination of Quantum States

János A. Bergou[1], Ulrike Herzog[2], and Mark Hillery[1]

[1] Department of Physics and Astronomy, Hunter College of The City University of New York, 695 Park Avenue, New York, NY 10021, jbergou@hunter.cuny.edu, mhillery@hunter.cuny.edu
[2] Institut für Physik, Humboldt-Universität Berlin, Newtonstrasse 15, 12489 Berlin, Germany, ulrike.herzog@physik.hu-berlin.de

Abstract. The problem of discriminating among given nonorthogonal quantum states is underlying many of the schemes that have been suggested for quantum communication and quantum computing. However, quantum mechanics puts severe limitations on our ability to determine the state of a quantum system. In particular, nonorthogonal states cannot be discriminated perfectly, even if they are known, and various strategies for optimum discrimination with respect to some appropriately chosen criteria have been developed. In this article we review recent theoretical progress regarding the two most important optimum discrimination strategies. We also give a detailed introduction with emphasis on the relevant concepts of the quantum theory of measurement. After a brief introduction into the field, the second chapter deals with optimum unambiguous, i. e error-free, discrimination. Ambiguous discrimination with minimum error is the subject of the third chapter. The fourth chapter is devoted to an overview of the recently emerging subfield of discriminating multiparticle states. We conclude with a brief outlook where we attempt to outline directions of research for the immediate future.

11.1 Introduction

In quantum information and quantum computing the carrier of information is some quantum system and information is encoded in its state [1]. The state, however, is not an observable in quantum mechanics [2] and, thus, a fundamental problem arises: after processing the information - i.e. after the desired transformation is performed on the input state by the quantum processor - the information has to be read out or, in other words, the state of the system has to be determined. When the possible target states are orthogonal, this is a relatively simple task if the set of possible states is known. But when the possible target states are not orthogonal they cannot be discriminated perfectly, and optimum discrimination with respect to some appropriately chosen criteria is far from being trivial even if the set of the possible nonorthogonal states is known. Thus the problem of discriminating among nonorthogonal states is ubiquitous in quantum information and quantum computing, underlying many of the communication and computing schemes that have been suggested so far. It is the purpose of this article to review various theoretical schemes that have been developed for discriminating among nonorthogonal quantum states. The corresponding experimental

J.A. Bergou, U. Herzog, M. Hillery, Discrimination of Quantum States, Lect. Notes Phys. **649**, 417–465 (2004)
http://www.springerlink.com/

realizations will be mentioned only in a cursory manner as they are reviewed elsewhere in this volume by Chefles.

The field of discriminating among nonorthogonal quantum states has been around for quite some time now [3]. Stimulated by the rapid developments in quantum information theory of the 90's the question of how to discriminate between nonorthogonal quantum states *in an optimum way* has gained renewed interest. The developments until about the late 90's are reviewed in an excellent review article by Chefles [4]. Therefore, in this review we will mainly focus our attention to recent advances not contained in [4] and will cite earlier results only when they are necessary for the understanding of the newer ones.

In order to devise an optimum state-discriminating measurement, strategies have been developed with respect to various criteria. In this review article we restrict ourselves to the two most obvious criteria for optimizing a measurement scheme that is designed for discriminating between different states of a quantum system, being either pure states or mixed states. The two methods are optimum unambiguous discrimination of the states, on the one hand, and state discrimination with minimum error, on the other hand. They will be outlined in detail in the next two sections of this review. In particular, at the beginning of these sections we give a tutorial introduction to the two main strategies by considering simple examples, namely unambiguous discrimination of two pure states in Sect. 11.2, and minimum-error discrimination of two mixed states in Sect. 11.3. This will allow us to introduce the concept of generalized measurements along with the other typical theoretical tools employed in problems of this sort. Selected applications of the unambiguous discrimination strategy, one chosen from quantum communication and the other from quantum computing will be reviewed at the end of Sect. 11.2. Section 11.4 is devoted to the recently emerging sub-field of discriminating among multipartite states by means of local operations and classical communication (LOCC). We conclude with a brief outlook in Sect. 11.5. Throughout the article we assume that for each measurement only a single copy of the quantum system is available. However, the case of multiple copies could be easily accounted for, in the measurement schemes we consider, if the states are replaced by their corresponding multi-fold tensor products. We note that apart from the two optimization schemes we consider, state-distinguishing measurements can also be optimized with respect to other criteria, such as requiring a maximum of the mutual information [5] or of the fidelity [6]. In particular, the former approach is of importance for the transmission of quantum information. From a mathematical point of view, however, these other criteria pose much bigger problems, since they rely on optimizing a nonlinear functional of the given states and, therefore, are beyond the scope of the current review.

11.2 Unambiguous Discrimination

In this section we will review schemes for unambiguous discrimination. Although the simplest case of two pure states is well known and has been reviewed extensively (see, for example, [4]), from a pedagogical point of view we find it useful to include it here as well, because many of the techniques employed later can be best understood on this simple example. We also want to mention that the two main discrimination strategies evolved rather differently from the very beginning. On the one hand, unambiguous discrimination started with pure states and only very recently turned its attention to discriminating among mixed quantum states. At the end of this section we will review recent progress in this area. On the other hand, minimum-error discrimination addressed the problem of discriminating among two mixed quantum states from the very beginning and the results for two pure states followed as special cases. Each strategy has its own advantages and drawbacks. While unambiguous discrimination is relatively straightforward to generalize for more than two states it is difficult to treat mixed states. The error-minimizing approach, initially developed for two mixed states, is hard to generalize for more than two states.

Before we begin our systematic study of the optimal unambiguous discrimination of two pure states in the next subsection, we want to gain some physical insight first. To this end, let us describe the procedure of optimum unambiguous discrimination between two nonorthogonal single photon polarization states in terms of classical optics. We consider a series of weak pulses containing, on the average, much less than one photon. Each pulse is linearly polarized, with equal probability, either in the direction $\boldsymbol{e}_1$ or in the direction $\boldsymbol{e}_2$ with $\boldsymbol{e}_{1,2} = \cos\Theta\,\boldsymbol{e}_x \pm \sin\Theta\,\boldsymbol{e}_y$, where we assume that $\cos\Theta \geq \sin\Theta$. If the pulses pass through a linear optical device and undergo polarization-selective linear attenuation in the x direction, their polarization vectors can be made orthogonal. For this purpose the attenuation process has to be designed in such a way that the amplitude of the x-component is reduced by a factor of $\tan\Theta$. Due to the attenuation, the initial electric field vectors $\boldsymbol{E}_{1,2} = E_0\boldsymbol{e}_{1,2}$ of the respective pulses are then transformed into the vectors $\boldsymbol{E}'_{1,2} = E_0\sqrt{2}\sin\Theta(\boldsymbol{e}_x \pm \boldsymbol{e}_y)/\sqrt{2}$. Hence, after leaving the linear optical device, the polarization directions of the two kinds of pulses are orthogonal and, therefore, can be discriminated unambiguously even when the pulses contain only a single photon. However, the intensity that can be used for unambiguous polarization state discrimination is reduced by a factor of $2\sin^2\Theta$ relative to the total initial intensity. Since a classical intensity ratio corresponds to the probability ratio of detecting a photon, every incoming photon will yield an unambiguous result in the state discrimination process only with the probability

$$P_D = 2\sin^2\Theta = 1 - \cos(2\Theta) = 1 - |\boldsymbol{e}_1\boldsymbol{e}_2| \; . \tag{11.1}$$

Here we introduced the modulus of the scalar product since, when the direction of one of the vectors, $\boldsymbol{e}_1$ or $\boldsymbol{e}_2$, is reversed, the linear polarization state remains the same. The probability that the discrimination procedure fails, i. e. that the polarization state of the photon cannot be determined unambiguously, is therefore $Q_F = 1 - P_D = |\boldsymbol{e}_1\boldsymbol{e}_2|$. Although orthogonalization of the polarization states is also possible when polarization-selective attenuation affects a polarization direction that differs from the x-axis, it is easy to see that the symmetric procedure described above is optimal in the sense that it yields the maximum achievable value of P_D, or the minimum value of $Q_F = 1 - P_D$, respectively.

From the pioneering investigations of unambiguous discrimination between general nonorthogonal quantum states it follows that for any two states, $|\psi_1\rangle$ and $|\psi_2\rangle$, occurring with equal a priori probability, the optimum probability of obtaining an unambiguous result is given by (11.1) when the scalar product $\boldsymbol{e}_1\boldsymbol{e}_2$ is replaced by the overlap $\langle\psi_1|\psi_2\rangle$, as we will show in the following [see (11.2), in particular].

11.2.1 Unambiguous Discrimination of Two Pure States

Unambiguous discrimination started with the work of Ivanovic [7] who studied the following problem. A collection of quantum systems is prepared so that each single system is equally likely to be prepared in one of two known states, $|\psi_1\rangle$ or $\psi_2\rangle$. Furthermore, the states are not orthogonal, $\langle\psi_1|\psi_2\rangle \neq 0$. The preparer then hands the systems over to an observer one by one whose task is to determine which one of the two states has actually been prepared in each case. All the observer can do is to perform a single measurement or perhaps a series of measurements on the individual system. Ivanovic came to the conclusion that if one allows inconclusive detection results to occur then in the remaining cases the observer can conclusively determine the state of the individual system.

It is rather easy to see that a simple von Neumann measurement can accomplish this task. Let us denote the Hilbert space of the two given states by $\mathcal{H}$ and introduce the projector P_1 for $|\psi_1\rangle$ and $\bar{P}_1$ for the orthogonal subspace, such that $P_1 + \bar{P}_1 = 1$, the unity in $\mathcal{H}$. Then we know for sure that $|\psi_2\rangle$ was prepared if in the measurement of $\{P_1, \bar{P}_1\}$ a click in the $\bar{P}_1$ detector occurs. A similar conclusion for $|\psi_1\rangle$ can be reached with the roles of $|\psi_1\rangle$ and $|\psi_2\rangle$ reversed. Of course, when a click along P_1 (or P_2) occurs then we learn nothing about which state was prepared thus corresponding to inconclusive results. Ivanovic's startling observation was that a sequence of measurements can sometimes do better than a single von Neumann measurement described here. Dieks [8] then found that this sequence of measurements can be realized with a single generalized measurement (POVM) and Peres subsequently showed that this POVM is optimal in the sense that its failure probability, the probability that an inconclusive outcome occurs, is minimum [9]. The probability of the inconclusive outcome, or failure, is $Q_{IDP} = |\langle\psi_1|\psi_2\rangle|$ and the

probability of success is then given by what is called the Ivanovic-Dieks-Peres (IDP) limit,

$$P_{IDP} = 1 - Q_{IDP} = 1 - |\langle\psi_1|\psi_2\rangle| \ . \tag{11.2}$$

This result can be generalized for the case when the preparation probabilities of the states, η_1 and η_2, are different, $\eta_1 \neq \eta_2$. The preparation probabilities are also called a priori probabilities. The IDP result thus corresponds to the case of equal a priori probabilities for the two states, $\eta_1 = \eta_2 = 1/2$, and the generalization for arbitrary a priori probabilities is due to Jaeger and Shimony [10]. Here we briefly review the derivation of the general result following [10] and the very readable account by Ban [11].

11.2.2 Optimal POVM and the Complete Solution

The von Neumann projective measurement described above has two outcomes. It can correctly identify one of the two states at the expense of missing the other completely and occasionally missing the identifiable one, as well. If we want to do better we would like to have a measurement with three - not just two - outcomes, $|\psi_1\rangle$, $|\psi_2\rangle$ and failure. In the two-dimensional Hilbert space $\mathcal{H}$ the number of possible outcomes for a von Neumann measurement cannot exceed two, since it is always restricted by the dimensionality of the Hilbert space. We have to turn to generalized measurements that allow greater flexibility than simple projective measurements [12]. In particular, the number of distinguishable outcomes can exceed the dimensionality of the corresponding Hilbert space. For our case this means that we replace the projector $\bar{P}_2$ by the quantum detection operator Π_1, $\bar{P}_1$ by Π_2 and introduce Π_0 for the inconclusive results in such a way that $\langle\psi_1|\Pi_1|\psi_1\rangle = p_1$ is the probability of successfully identifying $|\psi_1\rangle$, $\langle\psi_1|\Pi_0|\psi_1\rangle = q_1$ is the probability of failing to identify $|\psi_1\rangle$, (and similarly for $|\psi_2\rangle$). For unambiguous discrimination we then require $\langle\psi_2|\Pi_1|\psi_2\rangle = \langle\psi_1|\Pi_2|\psi_1\rangle = 0$. We want these possibilities to be exhaustive,

$$\Pi_1 + \Pi_2 + \Pi_0 = I \ , \tag{11.3}$$

where I is the unity operator in $\mathcal{H}$. The probabilities are always real and non-negative which implies that the quantum detection operators are Hermitean and non-negative or, in other words, positive semi-definite.

Clearly, (11.3) does not correspond to orthogonal measurements when all detection operators are different from zero. It describes an operation called *positive operator-valued measure* (POVM) or simply a generalized measurement with the detection operators as its elements.

We now want to determine these operators explicitly. Consider the operator $A_k = U_k \Pi_k^{1/2}$, where U_k is an arbitrary unitary operator ($k = 0, 1, 2$). From this expression we immediately obtain $\Pi_k = A_k^\dagger A_k$ and the detection

probability can be expressed as $\langle\psi_i|A_k^\dagger A_k|\psi_i\rangle = \|A_k\psi_i\|^2 \geq 0$ where $\|...\|$ stands for the norm. This expression also helps us to identify the so far arbitrary operator A_k. The expression $A_k|\psi_i\rangle$ corresponds to the post-detection state. Because of the positivity of the norm, the condition of unambiguous discrimination is equivalent to the requirement

$$A_1|\psi_2\rangle = A_2|\psi_1\rangle = 0 \ . \tag{11.4}$$

If we introduce $|\psi_i^\perp\rangle$ as the vector orthogonal to $|\psi_{i'}\rangle$ $(i \neq i')$ - a notation that will become obvious in Sect. 11.2.2 - then $A_1 = c_1|\bar{\psi}_1\rangle\langle\psi_1^\perp|$ and $A_2 = c_2|\bar{\psi}_2\rangle\langle\psi_2^\perp|$. Here c_i are complex coefficients to be determined from the condition of optimum and $|\bar{\psi}_1\rangle$ and $|\bar{\psi}_2\rangle$ are the post-detection states, normalized to unity. For perfect distinguishability of the post-detection states, corresponding to optimal discrimination, we have to require their orthogonality, $\langle\bar{\psi}_1|\bar{\psi}_2\rangle = 0$, so they can be represented by a pair of arbitrarily directed orthogonal vectors in $\mathcal{H}$.

With the help of these expressions we can write the detection operators as $\Pi_1 = A_1^\dagger A_1 = |c_1|^2|\psi_1^\perp\rangle\langle\psi_1^\perp|$ and $\Pi_2 = A_2^\dagger A_2 = |c_2|^2|\psi_2^\perp\rangle\langle\psi_2^\perp|$. Inserting these expressions in the definition of p_1 and p_2 gives $|c_1|^2 = p_1/|\langle\psi_1|\psi_1^\perp\rangle|^2$ and a similar expression for $|c_2|^2$. Finally, introducing $\cos\Theta = |\langle\psi_1|\psi_2\rangle|$ and $\sin\Theta = |\langle\psi_1|\psi_1^\perp\rangle|$, we can write the detection operators as

$$\begin{aligned} \Pi_1 &= \frac{p_1}{\sin^2\Theta}|\psi_1^\perp\rangle\langle\psi_1^\perp| \ , \\ \Pi_2 &= \frac{p_2}{\sin^2\Theta}|\psi_2^\perp\rangle\langle\psi_2^\perp| \ . \end{aligned} \tag{11.5}$$

Now, Π_1 and Π_2 are positive semi-definite operators by construction. However, there is one additional condition for the existence of the POVM which is the positivity of the inconclusive detection operator,

$$\Pi_0 = I - \Pi_1 - \Pi_2 \ . \tag{11.6}$$

This is a simple 2 by 2 matrix in $\mathcal{H}$ and the corresponding eigenvalue problem can be solved analytically. Non-negativity of the eigenvalues leads, after some tedious but straightforward algebra, to the condition

$$q_1 q_2 \geq |\langle\psi_1|\psi_2\rangle|^2 \ , \tag{11.7}$$

where $q_1 = 1 - p_1$ and $q_2 = 1 - p_2$ are the failure probabilities for the corresponding input states.

Equation (11.7) represents the constraint imposed by the positivity requirement on the optimum detection operators. The task we set out to solve can now be formulated as follows. Let

$$Q = \eta_1 q_1 + \eta_2 q_2 \tag{11.8}$$

denote the average failure probability for unambiguous discrimination. We want to minimize this failure probability subject to the constraint, (11.7).

Due to the relation, $P = \eta_1 p_1 + \eta_2 p_2 = 1 - Q$, the minimum of Q also gives us the maximum probability of success. Clearly, for optimum the product $q_1 q_2$ should be at its minimum allowed by (11.7), and we can then express q_2 with the help of q_1 as $q_2 = \cos^2\Theta/q_1$. Inserting this expression in (11.8) yields

$$Q = \eta_1 q_1 + \eta_2 \frac{\cos^2\Theta}{q_1} \ , \tag{11.9}$$

where q_1 can now be regarded as the independent parameter of the problem. Optimization of Q with respect to q_1 gives $q_1^{POVM} = \sqrt{\eta_2/\eta_1}\cos\Theta$ and $q_2^{POVM} = \sqrt{\eta_1/\eta_2}\cos\Theta$. Finally, substituting these optimal values into (11.8) gives the optimum failure probability,

$$Q^{POVM} = 2\sqrt{\eta_1\eta_2}\cos\Theta \ . \tag{11.10}$$

For $\eta_1 = \eta_2 = 1/2$ this reproduces the IDP result, (11.2), as it should.

Let us next see how this result compares to the average failure probabilities of the two possible unambiguously discriminating von Neumann measurements that were described at the beginning of this section. The average failure probability for the first von Neumann measurement, with its failure direction along $|\psi_1\rangle$, can be written by simple inspection as

$$Q_1 = \eta_1 + \eta_2 |\langle\psi_1|\psi_2\rangle|^2 \ , \tag{11.11}$$

since $|\psi_1\rangle$ gives a click with probability 1 in this direction but it is only prepared with probability η_1 and $|\psi_2\rangle$ gives a click with probability $|\langle\psi_1|\psi_2\rangle|^2$ but it is only prepared with probability η_2. The corresponding detector set-up, yielding Q_1 for the failure probability, is depicted in Fig. 11.1.

By entirely similar reasoning, the average failure probability for the second von Neumann measurement, with its failure direction along $|\psi_2\rangle$, is given by

$$Q_2 = \eta_1 |\langle\psi_1|\psi_2\rangle|^2 + \eta_2 \ . \tag{11.12}$$

The corresponding detector set-up, yielding Q_2 for the failure probability, is depicted in Fig. 11.2.

What we can observe is that Q_1 and Q_2 are given as the arithmetic mean of two terms and Q^{POVM} is the geometric mean of the same two terms for either case. So, one would be tempted to say that the POVM performs better always. This, however, is not quite the case, it does so only when it exists. The obvious condition for the POVM solution to exists is that both $q_1^{POVM} \leq 1$ and $q_2^{POVM} \leq 1$. Using $\eta_2 = 1 - \eta_1$, a little algebra tells us that the POVM exists in the range $\cos^2\Theta/(1+\cos^2\Theta) \leq \eta_1 \leq 1/(1+\cos^2\Theta)$. If η_1 is smaller than the lower boundary, the POVM goes over to the first von Neumann measurement and if η_1 exceeds the upper boundary the POVM goes over to the second von Neumann measurement. This can be easily seen from (11.5) and (11.6) since $p_1 = 1 - q_1 = 0$ for $q_1 = 1$ and Π_0 becomes a projection along

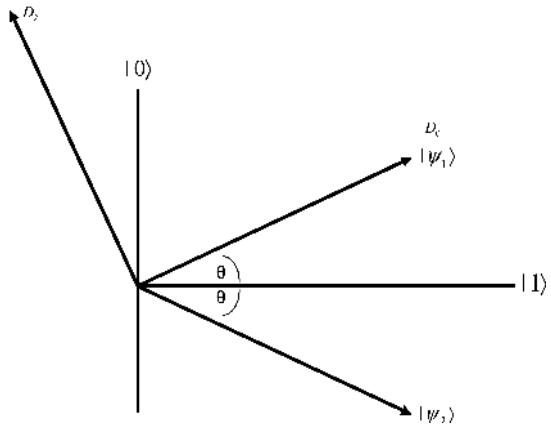

Fig. 11.1. A von Neumann measurement that discriminates $|\psi_2\rangle$ unambiguously. The detector $D_0 = P_1$ is set up along $|\psi_1\rangle$ and the detector $D_2 = \bar{P}_1$ is set up along the orthogonal direction. When a click in the D_2 detector occurs we know for certain that $|\psi_2\rangle$ was prepared as the input state since it is the only one that has a component along this direction. When a click in the detector D_0 occurs we learn nothing about which state was prepared as the input since both have a component along this direction. This measurement outcome then corresponds to the inconclusive result so D_0 is the failure detector and the probability that it clicks is Q_1.

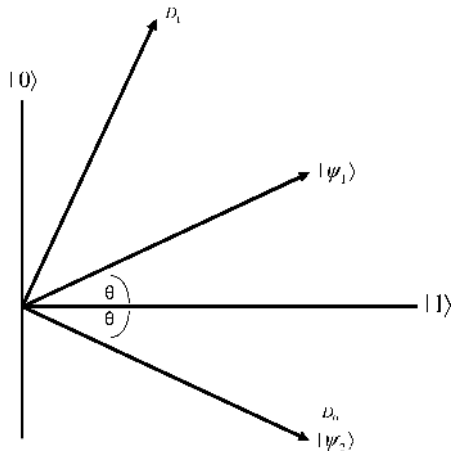

Fig. 11.2. A von Neumann measurement that discriminates $|\psi_1\rangle$ unambiguously. The detector $D_0 = P_2$ is set up along $|\psi_2\rangle$ and the detector $D_1 = \bar{P}_2$ is set up along the orthogonal direction. When a click in the D_1 detector occurs we know for certain that $|\psi_1\rangle$ was prepared as the input state since it is the only one that has a component along this direction. When a click in the detector D_0 occurs we learn nothing about which state was prepared as the input since both have a component along this direction. This measurement outcome then corresponds to the inconclusive result so D_0 is the failure detector and the probability that it clicks is Q_2.

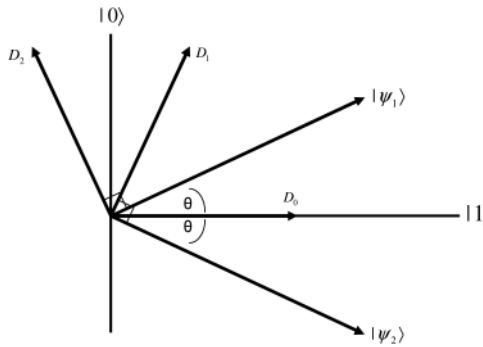

Fig. 11.3. Optimal POVM that discriminates $|\psi_1\rangle$ and $|\psi_2\rangle$ unambiguously. The detector $D_0 = \Pi_0$ is set up symmetrically between $|\psi_1\rangle$ and $|\psi_2\rangle$, the detector $D_1 = \Pi_1$ is set up along $\bar{P}_2$ and the detector $D_2 = \Pi_2$ is set up along $\bar{P}_1$. When a click in the D_i detector occurs we know for certain that $|\psi_i\rangle$ was prepared ($i = 1, 2$) as the input state since it is the only one that has a component along this direction. When a click in the detector D_0 occurs we learn nothing about which state was prepared as the input since both have a component along this direction. This measurement outcome then corresponds to the inconclusive result so D_0 is the failure detector and the probability that it clicks is Q^{POVM}. (The figure is for illustrative purposes only. Arrows representing POVM detection operators *do not correspond* to simple projections along their respective directions. They are drawn shorter than arrows representing state vectors which, in turn, have unit length. For an implementation of the POVM, see Sect. 11.2.3.)

$|\psi_1\rangle$ (and correspondingly for $p_2 = 0$). The set-up of the detection operators, yielding Q^{POVM} for the failure probability, is depicted in Fig. 11.3.

These findings can be summarized as follows. The optimal failure probability, Q^{opt}, is given as

$$Q^{opt} = \begin{cases} Q^{POVM} & \text{if } \frac{\cos^2\Theta}{1+\cos^2\Theta} \leq \eta_1 \leq \frac{1}{1+\cos^2\Theta} \ , \\ Q_1 & \text{if } \eta_1 < \frac{\cos^2\Theta}{1+\cos^2\Theta} \ , \\ Q_2 & \text{if } \frac{1}{1+\cos^2\Theta} < \eta_1 \ . \end{cases} \tag{11.13}$$

Figure 11.4 displays the failure probabilities vs. η_1 for a fixed value of the overlap, $\cos^2\Theta$.

The above result is very satisfying from a physical point of view. The POVM delivers a lower failure probability in its entire range of existence than either of the two von Neumann measurements. At the boundaries of this range it merges smoothly with the von Neumann measurement that has a lower failure probability at that point. Outside this range the state preparation is dominated by one of the states and the optimal measurement becomes a von Neumann projective measurement, using the state that is prepared less frequently as its failure direction.

Next we turn our attention to the physical implementation of the POVM.

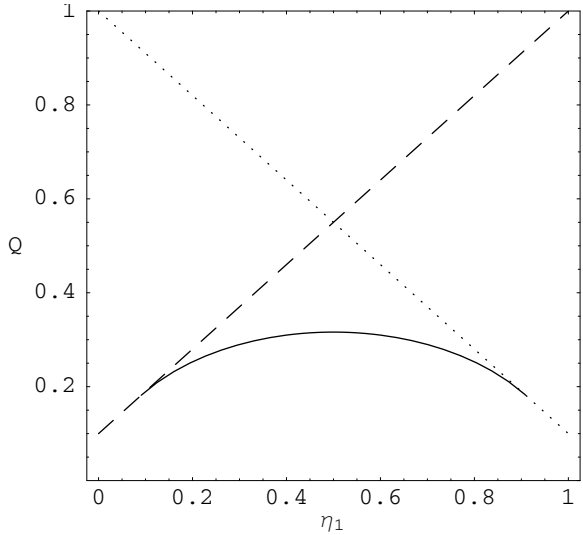

Fig. 11.4. Failure probability, Q, vs. the prior probability, η_1. Dashed line: Q_1, dotted line: Q_2, solid line: Q_{POVM}. For the figure we used the following representative value: $|\langle\psi_1|\psi_2\rangle|^2 = 0.1$. For this the optimal failure probability, Q_{opt} is given by Q_1 for $0 < \eta_1 < 0.09$, by Q_{POVM} for $0.09 \leq \eta_1 \leq 0.9$ and by Q_2 for $0.9 < \eta_1$.

11.2.3 Neumark's Theorem and the Realization of the POVM

The problem with the POVM is that it is very hard to realize in the original system Hilbert space since it involves detection operators that do not correspond to orthogonal projectors. In fact, they are not projectors at all. Fortunately, there is an essential result, known as Neumark's theorem [13]. It states that a POVM can be realized by the following constructive procedure, also known as generalized measurement. The system is embedded in a larger Hilbert space where the extra degrees of freedom are customarily called the ancilla. Then a unitary transformation entangles the system degrees of freedom with those of the ancilla. Finally, after this interaction, projective von Neumann measurements can be carried out on this larger system. As a consequence of the entanglement between the original system and the ancilla, a projective measurement on the larger system will also transform the system state in the original Hilbert space. One can choose the unitary transformation and the subsequent von Neumann measurement in such a way that if outcome k is found in the von Neumann measurement on the larger system, the resulting transformation on the original system state is $A_k|\psi\rangle$, i. e. it corresponds to an element of the POVM in the original system Hilbert space.

We illustrate the power of this theorem on an alternative derivation of the condition on the individual failure probabilities, (11.7). The joint Hilbert space $\mathcal{K}$ of the 'system plus ancilla' is a tensor product of the two Hilbert

spaces, $\mathcal{H}$ of the system and $\mathcal{A}$ of the ancilla, $\mathcal{K} = \mathcal{H} \otimes \mathcal{A}$. This means that a state in $\mathcal{K}$ is a superposition of product states where, in each product, the first member is from $\mathcal{H}$ and the second is from $\mathcal{A}$. Specifically, the two inputs now correspond to $|\psi_1\rangle|\phi_0\rangle$ and $|\psi_2\rangle|\phi_0\rangle$, where $|\phi_0\rangle$ describes the initial state of the ancilla. We choose the unitary transformation as

$$U(|\psi_1\rangle|\phi_0\rangle) = \sqrt{p_1}|\psi_1'\rangle|\phi_0\rangle + \sqrt{q_1}|\psi_0\rangle|\phi_1\rangle \ , \tag{11.14}$$

$$U(|\psi_2\rangle|\phi_0\rangle) = \sqrt{p_2}|\psi_2'\rangle|\phi_0\rangle + \sqrt{q_2}e^{i\theta}|\psi_0\rangle|\phi_1\rangle \ , \tag{11.15}$$

where $|\phi_1\rangle$ is chosen to be orthogonal to $|\phi_0\rangle$, and $|\psi_1'\rangle$ and $|\psi_1'\rangle$ correspond to orthogonal vectors in the original Hilbert space. If we now perform a von Neumann measurement on the ancilla then a click along the $|\phi_1\rangle$ direction collapses both inputs onto the same output, $|\psi_0\rangle$, and all information about the inputs is lost. The probability that this happens for input i is q_i ($i = 1, 2$). Obviously, this outcome corresponds to the inconclusive result so q_i are the failure probabilities of the corresponding input states i. On the other hand, a click along the $|\phi_0\rangle$ direction transforms the original inputs into orthogonal outputs in the system Hilbert space. The probability that this happens for input i is p_i ($i = 1, 2$). Obviously, this outcome corresponds to full distinguishability in the original system Hilbert space, so p_i are the probabilities of success for discriminating the corresponding input states i. From unitarity we obtain $p_i + q_i = 1$ for $i = 1, 2$ and by taking the inner product of (11.14) and (11.15) we obtain (11.7). Thus Neumark's theorem delivers the positivity condition in a few lines and from here the rest of the derivation of the optimal failure probability goes along the same lines as for the POVM method. An optical implementation, based on the tensor product extension of the Hilbert space and employing linear optical elements (beam splitters and phase shifters) only, has been proposed in [14]. The states to be discriminated are represented by a single photon that can be in a superposition between two input ports of a six port interferometer with three input and three output ports. The third input corresponds to the ancilla, initially in the vacuum state. The desired unitary transformation is carried out by this six port constructed from the appropriate linear optical elements. At the output side detectors are placed in front of each output port. If the photon emerges from the ancilla port 3, the measurement failed. We can construct the device so that if the photon emerges from port i ($i = 1, 2$, and we know that the input state was i. Thus, detector clicks in the first two output ports correspond to successful measurements. Measurements of this type can easily be generalized higher dimensional systems, using more general multiports. An explicit example will be given in Sect. 3.2.

Before moving on to more uncharted territory we should note that the example of unambiguous discrimination between two nonorthogonal polarization states by using a polarization-sensitive absorber that was considered at the beginning of this section does not conform to the scheme described here, although it clearly accomplishes its goal. Rather, it corresponds to another

kind of possible extension of the original system Hilbert space via the direct sum method. In this case we append the ancilla space, $\mathcal{A}$, to the system space $\mathcal{H}$ to form the joint Hilbert space $\mathcal{K}$ of the 'system plus ancilla'. $\mathcal{K}$ is now a direct sum, $\mathcal{K} = \mathcal{H} \oplus \mathcal{A}$, of $\mathcal{H}$ and $\mathcal{A}$, meaning that a state in $\mathcal{K}$ is a superposition of two terms where the first one lies entirely in $\mathcal{H}$ and the second one in $\mathcal{A}$. In particular, the two input states, $|\psi_1\rangle$ and $|\psi_2\rangle$, have no components in the ancilla space. In the example with the two single photon polarization states spanning the two-dimensional system Hilbert space this means that we simply include a third state, the vacuum, as the ancilla space. We can say that our two-dimensional objects, the qubits, secretly live in the three-dimensional space of qutrits. The described unambiguous discrimination procedure between two single-photon polarization states has been successfully performed experimentally, implementing the polarization-selective attenuation either by polarization-dependent absorption in a fiber [15] or with the help of polarizing beamsplitters [16]. The transformation corresponding to the linear absorber redistributes the population among the three basis states in such a way that the parts of the input qubit states that remain in the original qubit Hilbert space become orthogonal there. In this method we do not keep track of what happens to the photon in the absorption process, whereas in the tensor product method, including a complete description of the ancilla (environment), we do. For the simple examples that we consider, one can introduce equivalent notations for the two alternative methods. Therefore, in the next section we will employ the somewhat simpler notations of the direct sum method.

This completes our tutorial review of the unambiguous discrimination of two pure states and illustrates the two possible approaches, the direct POVM method and the method of generalized measurements based on Neumark's theorem. We now turn our attention to more complicated problems and briefly review recent progress in dealing with them.

11.2.4 More than Two Pure States

So far we have only considered discriminating between two states, and we have seen that in that case a complete solution can be given. For more than two states, however, there are only a few general results, and explicit solutions exist only for special cases. Here we shall review what is known.

Two general results apply to the case of unambiguous discrimination. The first, due to Chefles, is that only linearly independent states can be unambiguously discriminated [17]. This can be seen as follows. Let the POVM for discriminating the N states $|\psi_1\rangle$, ... $|\psi_N\rangle$ be given by the operators A_1, ... A_N, and A_I, where the operators act on the vectors in the space $\mathcal{H}$, which is the span of the vectors $|\psi_1\rangle$, ... $|\psi_N\rangle$, and

$$A_I^\dagger A_I + \sum_{j=1}^{N} A_j^\dagger A_j = I \ , \tag{11.16}$$

which is an obvious generalization of (11.3) to N states. The operator A_I again corresponds to the inconclusive outcome, and the operator A_j, for $j = 1 \ldots N$ corresponds to identifying the state as $|\psi_j\rangle$. Because there must be no errors, we must have

$$\langle\psi_k|A_j^\dagger A_j|\psi_k\rangle = p_j\delta_{jk}, \tag{11.17}$$

where $0 \leq p_j \leq 1$ is the probability of successfully identifying $|\psi_j\rangle$. Now suppose that the states are linearly dependent so that they can be expressed in terms of each other

$$|\psi_j\rangle = \sum_{k=1}^{N} c_{jk}|\psi_k\rangle. \tag{11.18}$$

Substituting this into the above equation we have that

$$\sum_{m,n=1}^{N} c_{km}^* c_{kn}\langle\psi_m|A_j^\dagger A_j|\psi_n\rangle = p_j\delta_{jk}. \tag{11.19}$$

This can be simplified by noting that

$$|\langle\psi_m|A_j^\dagger A_j|\psi_n\rangle|^2 \leq \langle\psi_m|A_j^\dagger A_j|\psi_m\rangle\langle\psi_n|A_j^\dagger A_j|\psi_n\rangle, \tag{11.20}$$

which gives us that

$$\langle\psi_m|A_j^\dagger A_j|\psi_n\rangle = p_j\delta_{mn}\delta_{jm}. \tag{11.21}$$

Substituting this into (11.19) we find that $|c_{kj}|^2 = \delta_{jk}$, which implies that the states are not linear combinations of each other and are, hence, linearly independent.

It is also possible to quickly draw some conclusions about the form of the operator A_j. Because we have that

$$A_j|\psi_k\rangle = 0, \tag{11.22}$$

for $j \neq k$, it annihilates the subspace $\mathcal{H}_k$, which is the span of the vectors $|\psi_1\rangle$, … $|\psi_N\rangle$ with $|\psi_k\rangle$ omitted. Let $|\psi_k^\perp\rangle$ be the unit vector orthogonal to $\mathcal{H}_k$ which, by the way, explains the notation introduced in the previous section. We can then choose

$$A_j = \frac{\sqrt{p_j}}{\langle\psi_j^\perp|\psi_j\rangle}|\psi_j'\rangle\langle\psi_j^\perp| \ , \tag{11.23}$$

where $|\psi_j'\rangle$, $j = 1, \ldots N$ are arbitrary orthogonal unit vectors. The remaining problem is to find the values of p_j. Let us denote the *a priori* probability of the state $|\psi_j\rangle$ by η_j. The values of p_j should be chosen to maximize the

average success probability, P, where

$$P = \sum_{j=1}^{N} \eta_j p_j \ . \tag{11.24}$$

In addition, they must be chosen so that the operator

$$\begin{aligned} A_I^\dagger A_I &= I - \sum_{j=1}^{N} A_j^\dagger A_j \\ &= I - \sum_{j=1}^{N} \frac{p_j |\psi_j^\perp\rangle\langle\psi_j^\perp|}{|\langle\psi_j|\psi_j^\perp\rangle|^2}, \end{aligned} \tag{11.25}$$

is positive. This is a nontrivial problem.

The second general result states that there do exist upper and lower bounds on the success probability. Here we shall just present the results; the interested reader can consult the original papers for the derivations. An upper bound is given by [18]

$$P \leq 1 - \frac{1}{N-1} \sum_{j=1}^{N} \sum_{k=1, k\neq j}^{N} \sqrt{\eta_j \eta_k} |\langle\psi_j|\psi_k\rangle|. \tag{11.26}$$

Building on work by Duan and Guo [19], X. Sun, *et al.* derived a lower bound [20]. Consider the $N \times N$ matrix whose elements are $\langle\psi_j|\psi_k\rangle$, and let λ_N be the smallest eigenvalue of this matrix. They showed that $P \geq \lambda_N$.

The problem of discriminating among three nonorthogonal states was first considered by Peres and Terno [21]. They developed a geometric approach and applied it numerically to several examples. A different method was considered by Duan and Guo [19] and Y. Sun and ourselves [22]. We considered the three vectors to be discriminated, $|\psi_j\rangle$, $j = 1, 2, 3$, to lie in the space $\mathcal{H}$. To this a "failure" space, $\mathcal{A}$, is appended so that the whole problem takes place in the space obtained by the direct sum extension, $\mathcal{K} = \mathcal{H} \oplus \mathcal{A}$. If the procedure fails, the vector $|\psi_j\rangle$ is mapped into a vector in the failure space, $|\phi_j\rangle$, and if it succeeds it is mapped onto a vector in the original space, $\sqrt{p_j}|\psi_j'\rangle$, where $\|\psi_j'\| = 1$, and $0 \leq p_j \leq 1$. The vectors $|\psi_j'\rangle$ are mutually orthogonal, so that they can be perfectly distinguished. Chefles showed that the set of failure vectors must be linearly dependent for the optimal procedure [17], so that the dimension of $\mathcal{A}$ will be one or two. One way of understanding this result is that if the failure vectors were linearly independent, then we could perform a further unambiguous state discrimination procedure on them and, with some probability, tell which state we were originally given. This would imply that the original procedure was not optimal. Therefore, the optimal procedure produces linearly dependent failure vectors, which cannot be further discriminated.

Making this more explicit, we assume that there is a unitary operator, U, acting on $\mathcal{K}$, such that

$$U|\psi_j\rangle = \sqrt{p_j}|\psi'_j\rangle + |\phi_j\rangle \ . \tag{11.27}$$

It should be noted that, unlike in (11.15), the vector $|\phi_j\rangle$ is not normalized to unity. Instead, we use $\langle\phi_j|\phi_j\rangle = q_j$ here. After U has been performed, we measure the projection operator onto the space $\mathcal{H}$. If we obtain 1, the procedure has succeeded, and we know what the input state was. If the input was $|\psi_j\rangle$, the procedure succeeds with probability p_j. If we obtain 0, the procedure has failed, and this happens with probability $q_j = 1 - p_j = \|\phi_j\|^2$, if the input state was $|\psi_j\rangle$. The above equation implies that

$$\langle\phi_j|\phi_k\rangle = \langle\psi_j|\psi_k\rangle - \delta_{jk}p_j. \tag{11.28}$$

Defining the matrix $C_{jk} = \langle\phi_j|\phi_k\rangle$, we see by its definition that it must be positive definite. Therefore, the problem of finding the optimal unambiguous state discrimination procedure reduces to finding the values of p_j that optimize the success probability

$$P = \sum_{j=1}^{3} \eta_j p_j, \tag{11.29}$$

subject to the constraint that the 3×3 matrix, whose elements are $\langle\psi_j|\psi_k\rangle - \delta_{jk}p_j$, is positive.

This can be solved in some special cases. We shall assume that all of the a priori probabilities are the same, so that they are all 1/3. If all of the overlaps are the same, i.e.

$$\langle\psi_1|\psi_2\rangle = \langle\psi_1|\psi_3\rangle = \langle\psi_2|\psi_3\rangle = s, \tag{11.30}$$

where s is real and positive, then $q_j = s$, for $j = 1, 2, 3$, and $Q = 1 - P = s$ as well.

There is also an explicit solution if

$$\begin{aligned} \langle\psi_1|\psi_2\rangle = \langle\psi_1|\psi_3\rangle &= s_1 \ , \\ \langle\psi_2|\psi_3\rangle &= s_2 \ , \end{aligned} \tag{11.31}$$

where both s_1 and s_2 are real and positive. We first note that for a fixed value of s_1 there is a restriction on how large s_2 can be. The largest the angle between $|\psi_2\rangle$ and $|\psi_3\rangle$ can be is twice the angle between $|\psi_1\rangle$ and $|\psi_2\rangle$ (this maximum is achieved when the vectors are coplanar). This implies that $s_2 \geq 2s_1^2 - 1$. The solution to the state discrimination problem depends on whether $s_1/s_2 < 2$ or not. If it is, we have

$$\begin{aligned} q_1 &= \frac{s_1^2}{s_2} \ , \quad q_2 = q_3 = s_2 \ , \\ Q &= \frac{1}{3}\left[\frac{s_1^2}{s_2} + 2s_2\right] \ . \end{aligned} \tag{11.32}$$

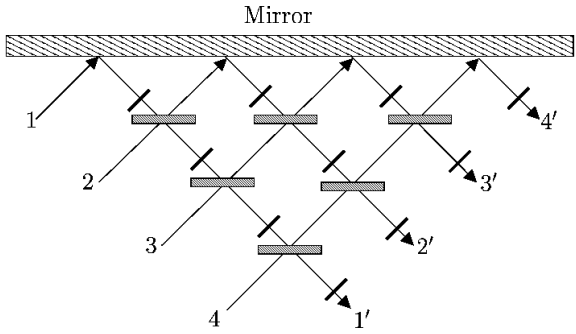

Fig. 11.5. An optical eight-port. The beams are straight lines, a suitable beam splitter is placed at each point where two beams intersect, phase shifters are at one input of each beam splitter and at each output.

If $s_1/s_2 \geq 2$, we have

$$q_1 = 2s_1 \ , \quad q_2 = q_3 = s_1 + s_2 \ ,$$
$$Q = \frac{2}{3}(2s_1 - s_2) \ . \tag{11.33}$$

This approach lends itself naturally to an optical implementation [22]. The states to be discriminated are represented by a single photon that can be in one of three modes. Additional modes (one or two, depending on the dimension of the failure space) that represent the ancilla, are initially in the vacuum state. Figure 11.5 displays the simpler case, that of an eight port. In general, the unitary transformation is carried out by an optical N port, where N is either 8 or 10, depending on the number of vacuum ports needed. It should be noted that an N port is a linear optical device with $N/2$ inputs and $N/2$ outputs, that can be constructed from beam splitters, phase shifters, and mirrors. At the output detectors are placed to determine from which of the ports the photon emerges. If it emerges from one of the failure ports, the measurement has failed, but if it emerges from one of the other three, we know what the input state was.

In particular, we can construct the device so that the three output ports that correspond to a successful measurement are numbered 1 through 3, and a photon emerging from port j means that the input state was $|\psi_j\rangle$. Measurements of this type have been carried out by Mohseni, *et al.* [23].

In the case that the failure space is two-dimensional, it is sometimes possible to obtain some information about the input state even if the initial measurement fails [21,22]. Sun, *et al.* presented an example of an optical network that does the following [22]. It consists of an optical 10 port followed by a 6 port. The first three inputs of the 10 port are where the state $|\psi_j\rangle$, $j = 1, 2, 3$ is sent in, and the other two are in the vacuum state The failure space for this particular situation is two-dimensional, and if the photon emerges from outputs 4 or 5, the measurement has failed. If it emerges from

outputs 1, 2, or 3, we know what the input state was. The 6 port has as its three inputs the two failure outputs from the 10 port, and the vacuum. It is constructed so that if the photon comes out of the first output, we know that the input state was not $|\psi_3\rangle$, if it comes out of the second output, the input was not $|\psi_2\rangle$, and if it comes out of the third output, the measurement has failed. Therefore, even if the initial measurement (the 10 port) fails, there is some possibility of gaining information about the input state by further processing. Note that this depends on the failure space having two dimensions; with a one-dimensional failure space, no further processing is possible.

There is another case where the exact solution to the unambiguous discrimination problem is known, and to which we now turn our attention. First we note that a set of N states is called symmetric [3,5] if there exists a unitary operator, V, such that, for $j = 1, \ldots N-1$,

$$V|\psi_j\rangle = |\psi_{j+1}\rangle \ , \quad V|\psi_N\rangle = |\psi_1\rangle \ . \tag{11.34}$$

This implies that $|\psi_j\rangle = V^{j-1}|\psi_1\rangle$. The case of unambiguous state discrimination for N symmetric states was analyzed by Chefles and Barnett [24]. They found an analytical expression for the optimal success probabilities for the case when the *a priori* probabilities of the states are the same. The vectors $|\psi_j\rangle$, $j = 1, \ldots N$ are now assumed to be linearly independent, and to span the entire space. Because the states $|\psi_j\rangle$ form a basis for the space, (11.34) now implies that $V^N = I$. This, in turn, means that V can be expressed as

$$V = \sum_{k=0}^{N-1} e^{2\pi i k/N} |\gamma_k\rangle\langle\gamma_k| \ , \tag{11.35}$$

where $|\gamma_k\rangle$ is the eigenstate of V with eigenvalue $e^{2\pi i k/N}$. The states we are trying to distinguish among can now be expanded as

$$|\psi_j\rangle = \sum_{k=0}^{N-1} e^{2\pi i k(j-1)/N} c_k |\gamma_k\rangle \ . \tag{11.36}$$

The optimal success probability was found to be

$$P = N \min |c_k|^2 \ , \tag{11.37}$$

where the minimum is taken over k, and throughout the derivation it was assumed that the *a priori* probabilities of the states were the same.

11.2.5 The Discrimination of Mixed States

We begin this part by introducing some terminology first. This will greatly facilitate the presentation of the material to follow and help us to interpret some of the existing statements about this topics in the literature.

The *support* of a mixed state, described by a density matrix, is the space spanned by its eigenvectors with nonzero eigenvalues. The *rank* of a mixed state is the dimension of its support. The *kernel* of a mixed state is the space orthogonal to its support. With these definitions at hand, we are ready to interpret the statement that "... one cannot unambiguously discriminate mixed states (the reason is that the IDP scheme does not work for linearly dependent states)" [25]. This is not the only statement of this sort but is a representative one that is obviously correct. But we have to elaborate it further. It clearly refers to mixed states that have the same support and excludes mixed states that have a nonzero overlap with the kernels of the others. In the following we will focus our attention on precisely those cases, admittedly a small subset of all mixed states of a system, where the density operators to be discriminated have different supports. It is only within this subset that unambiguous discrimination of mixed states is possible.

By the time of the publication of [25], several other works have been published that can be interpreted as special instances of unambiguous discrimination of mixed states. Below we give a brief overview of recent progress in this area.

Unambiguous Filtering

In [26], we introduced the problem of unambiguous discrimination between sets of states. The set of N given states is divided into subsets, and we want to determine to which subset a particular input state, known to be prepared in one of the N given states, belongs. In the simplest case, the division is into two sets only, the first containing the first M states and the second the remaining $N - M$ states. We want to unambiguously assign a given input state to one of these two subsets. Clearly, the discrimination of two pure states corresponds to $N = 2$ and $M = 1$, there is one state in each set. The next simplest case is the one with $N = 3$ and $M = 1$. This is the case of unambiguously discriminating whether a state is $|\psi_1\rangle$ or whether it is in the set $\{|\psi_2\rangle, |\psi_3\rangle\}$ with a priori probabilities η_1, η_2, and η_3, respectively. Since in this case all we are interested in is whether a particular input is $|\psi_1\rangle$ or not we termed this case unambiguous quantum state filtering. (Actually, we introduced the term quantum state filtering first in the context of minimum-error discrimination [63] and this work will be discussed in Sect. 11.3.4.) First, it is straightforward to see that filtering is a particular instance of mixed state discrimination and, by a simple extension of the following considerations, set discrimination, in general, is equivalent to the discrimination of mixed states. Indeed, since we do not want to resolve the states within a set, we can introduce the density operators

$$\rho_\alpha = |\psi_1\rangle\langle\psi_1|, \qquad \rho_\beta = \frac{\eta_2}{\eta_2 + \eta_3}|\psi_2\rangle\langle\psi_2| + \frac{\eta_3}{\eta_2 + \eta_3}|\psi_3\rangle\langle\psi_3|, \tag{11.38}$$

with a priori probabilities $\eta_\alpha = \eta_1$, $\eta_\beta = \eta_2 + \eta_3$. Filtering is, thus, the discrimination between these two density operators, one a rank one mixed state which is, in fact, a pure state, and the other a rank two mixed state.

Key to the solution is the observation that the ancilla space is one dimensional in the Neumark implementation of the optimal POVM. This immediately yields the condition analogous to (11.7),

$$q_1 q_i = |\langle\psi_1|\psi_i\rangle|^2 \ , \tag{11.39}$$

where q_1 and q_i $(i = 2, 3)$ are the failure probabilities for the corresponding input states. The derivation of the optimum average failure probability is then very similar to the derivation of the IDP result for two pure states. The optimal failure probability, Q^{opt}, is given as

$$Q^{opt} = \begin{cases} Q^{POVM} & \text{if } \frac{F^2}{1+F^2} \leq \eta_\alpha \leq \frac{1}{1+F^2} \ , \\ Q_\alpha & \text{if } \eta_\alpha < \frac{F^2}{1+F^2} \ , \\ Q_\beta & \text{if } \frac{1}{1+F^2} < \eta_\alpha \ . \end{cases} \tag{11.40}$$

Here $F = \sqrt{\langle\psi_1|\rho_\beta|\psi_1\rangle}$ is the fidelity between a pure and a mixed state [1], and

$$\begin{aligned} Q^{POVM} &= 2\sqrt{\eta_\alpha \eta_\beta} F \ , \\ Q_\alpha &= \eta_\alpha + \eta_\beta F^2 \ , \\ Q_\beta &= \eta_\alpha \|\psi_1^{\|}\|^2 + \eta_\beta \frac{F^2}{\|\psi_1^{\|}\|^2} \ . \end{aligned} \tag{11.41}$$

In the last line $|\psi_1^{\|}\rangle$ is the component of $|\psi_1\rangle$ in the support of ρ_β. Written in terms of the fidelity between a pure and a mixed state, the solution remains valid for an arbitrary number of states in the second set and, so, it represents the solution for a rank one vs. rank N unambiguous discrimination problem where N is arbitrary [27]. An interesting application of this result for a probabilistic quantum algorithm will be presented in Sect. 11.2.6. We note that unambiguous discrimination between multiple sets of pure states is a straightforward generalization and was independently investigated by Zhang and Ying in [28]. It, too, can be recast in a form of discriminating between mixed states. The problem of filtering a mixed state out of many was addressed by Takeoka, *et al.* in [29].

State Comparison

In [30], Barnett, Chefles, and Jex introduced the problem of state comparison that can be stated as follows. Given two systems, each of which is in one of two, in general, nonorthogonal states $\{|\psi_1\rangle, |\psi_2\rangle\}$, what is the optimum probability to determine whether the two systems were prepared in the same

state or in different states? One of the surprising aspects of their result is that one can decide with a certain probability whether the two systems are not in the same state even if there is no prior knowledge of the possible states. In this context we note that discrimination between non perfectly known states has also been addressed by Ježek in [31].

More importantly for our purpose, the comparison of states can be cast to a form that corresponds to the discrimination between mixed states described by rank two density matrices. Obviously, we now want to discriminate between the sets $\{|\psi_1\rangle|\psi_1\rangle, |\psi_2\rangle|\psi_2\rangle\}$ and $\{|\psi_1\rangle|\psi_2\rangle, |\psi_2\rangle|\psi_1\rangle\}$. The first set contains the combined states where the two systems are in the same individual state with a priori probabilities η_1 and η_2 for the respective combined states and the second set contains the states where the two systems are in different individual states with a priori probabilities η_3 and η_4 for the respective combined states. Introducing density operators for the sets, by employing a slight extension of the method in (11.38), we see that state comparison is equivalent to the discrimination of two rank two mixed states [32]. In [33], we derived the optimal POVM failure probability solution for the unambiguous discrimination between a rank two and a rank N density matrix,

$$Q^{POVM} = 2\sqrt{\eta_\alpha \eta_\beta} F \ , \tag{11.42}$$

where F is the fidelity between the two mixed states ρ_α and ρ_β [1]. In the case of state comparison, the intersection between the two supports is one dimensional and the fidelity can easily be calculated to give $F = |\langle\psi_1|\psi_2\rangle|$. Thus, for the optimal failure probability of state comparison, Q_{SC}, we get

$$Q_{SC} = 2\sqrt{\eta_\alpha \eta_\beta}|\langle\psi_1|\psi_2\rangle| \ , \tag{11.43}$$

where $\eta_\alpha = \eta_1 + \eta_2$ and $\eta_\beta = \eta_3 + \eta_4$ are the a priori probabilities for the respective mixed states. This result slightly generalizes the result of [30], based on a previously unpublished result of our own [33]. In a subsequent work [34], Jex, Andersson, and Chefles extended the state comparison results from two systems to the comparison of many. By employing similar methods to the ones here, this can be shown to be equivalent to the discrimination between density matrices of higher rank.

Two Arbitrary Mixed States: General Considerations

Generalizing the ideas of [26]- [34], two important works have appeared recently. In the first, Rudolph, *et al.* [35] established lower and upper bounds on the minimum failure probability for the unambiguous discrimination of two mixed states. Based on Uhlmann's theorem [1] they found that the lower bound is given by the fidelity, and the upper bound is based on the geometrical invariants between the kernels. They showed that for all known solutions the upper and lower bounds coincide and found numerically in other cases that the two bounds are very close. As an application of their method, they

also provided the general solution for the one-dimensional kernel problem and applied it to the special case of two rank $N-1$ density matrices in an N-dimensional Hilbert space.

Raynal, *et al.* [36], introduced reduction theorems for the problem of optimal unambiguous discrimination of two general density matrices of rank N and M. In particular, they showed that the problem can be reduced to the discrimination of two density matrices that have the same rank N_0 where N_0 is bounded by $N_0 \leq \min(N, M)$. Necessary and sufficient conditions for optimality were discussed also in [25] and [37], along with some numerical methods based on linear programming.

The upper and lower bounds on the failure probabilities by [35] and the reduction theorems by [36] bring us very close to a full solution of the unambiguous discrimination problem between two arbitrary mixed states and it is fully expected that a full solution will be found in the near future.

11.2.6 Selected Applications

We shall conclude this section with a discussion of two applications of unambiguous state discrimination. The first is a quantum cryptographic protocol, while the second is a quantum algorithm.

Bennett proposed using the unambiguous discrimination of two nonorthogonal states as the basis of a quantum-key-distribution protocol [38]. Alice and Bob want to establish a secure key that they can use to send encrypted messages to each other. To do so, Alice sends Bob a sequence of particles, where each particle is either in the state $|\psi_1\rangle$ or $|\psi_2\rangle$. The state $|\psi_1\rangle$ corresponds to a bit value of 0 and $|\psi_2\rangle$ corresponds to a bit value of 1. These states are known to both Alice and Bob, and they are not orthogonal. Upon receiving a particle, Bob applies the optimal two-state unambiguous measurement procedure to it. He then tells Alice over a public channel whether the measurement succeeded or failed. If it succeeded, they keep the bit, and if it failed, they discard the bit. In this way they can establish a key.

An eavesdropper, Eve, who wants to determine the key without being discovered has a problem. Let us assume that she can intercept the particles Alice is sending to Bob, and that she knows the states $|\psi_1\rangle$ and $|\psi_2\rangle$. Because these states are not orthogonal, she cannot tell with certainty which state a particular particle is in. One possibility is for her to apply the same procedure used by Bob, optimal two-state unambiguous state discrimination. If her measurement succeeds, all is well. She simply notes which state she found, and prepares another particle in this state and sends it on to Bob. She then knows this key bit, and Alice and Bob do not know that she knows. However, if her measurement fails, she does not know which state Alice sent, and she has to make a guess which state to send on to Bob. That means that she will introduce discrepancies between the state that Alice sent and the state that Bob received. In some of the cases in which this happens, Bob's measurement will succeed, and this will cause errors in the key to appear. If Alice and Bob

publicly share a subset of their good key bits (these bits have to then be discarded), and if they see discrepancies, then they know an eavesdropper was present, and that the key is insecure. If they find none, then with high probability (Eve could get lucky and have all of her measurements succeed, but this is very improbable) the key is secure.

Our second example is a probabilistic quantum algorithm to discriminate between sets of Boolean functions [27]. A Boolean function on n bits is one that returns either 0 or 1 as output for every possible value of the input x, where $0 \leq x \leq 2^n - 1$. The function is constant if it returns the same output on all of its arguments, i.e. either all 0's or all 1's; it is balanced if it returns 0's on half of its arguments and 1's on the other half; and it is biased if it returns 0's on m_0 of its arguments and 1's on the remaining $m_1 = 2^n - m_0$ arguments ($m_0 \neq m_1 \neq 0$ or $2^n - 1$). Classically, if one is given an unknown function and told that it is either balanced or constant, one needs $2^{(n-1)} + 1$ measurements to decide which. Deutsch and Jozsa [39] developed a quantum algorithm that can accomplish this task in one step. To discriminate a biased Boolean function from an unknown balanced one, $2^{(n-1)} + m_1 + 1$ measurements are needed classically, where, without loss of generality, we have assumed that $m_1 < m_0$. There is a probabilistic quantum algorithm, based on quantum state filtering, that can unambiguously discriminate a known biased Boolean function from a given set of balanced ones. There is a significant chance that only one function evaluation will be necessary.

The algorithm distinguishes between sets of Boolean functions. Let $f(x)$, where $0 \leq x \leq 2^n - 1$, be a Boolean function, i.e. $f(x)$ is either 0 or 1. One of the sets we want to consider is a set of balanced functions. The second set has only two members, and we shall call it $\mathcal{W}_k$. A function is in $\mathcal{W}_k$ if $f(x) = 0$ for $0 \leq x < [(2^k - 1)/2^k]2^n$ and $f(x) = 1$ for $[(2^k - 1)/2^k]2^n \leq x \leq 2^n - 1$, or if $f(x) = 1$ for $0 \leq x < [(2^k - 1)/2^k]2^n$ and $f(x) = 0$ for $[(2^k - 1)/2^k]2^n \leq x \leq 2^n - 1$. We now wish to distinguish between the balanced functions and functions in $\mathcal{W}_k$, that is, we are given an unknown function that is in one of the two sets, and we want to find out which set it is in. We note that the two functions in $\mathcal{W}_k$ are biased functions, so that this is a special case of a more general problem of distinguishing a set of biased functions from balanced functions.

The Deutsch-Jozsa algorithm makes use of the unitary operation

$$|x\rangle|y\rangle \rightarrow |x\rangle|y + f(x)\rangle, \tag{11.44}$$

where the first state, $|x\rangle$, is an n-qubit state, the second state, $|y\rangle$, is a single qubit state, and the addition is modulo 2. The state $|x\rangle$, where x is an n-digit binary number, is a member of the computational basis for n qubits, and the state $|y\rangle$, where y is either 0 or 1, is a member of the computational basis for a single qubit. In solving the Deutsch-Jozsa problem, this mapping is employed in the following way

$$\sum_{x=0}^{D-1} |x\rangle(|0\rangle - |1\rangle) \to \sum_{x=0}^{D-1} (-1)^{f(x)}|x\rangle(|0\rangle - |1\rangle), \tag{11.45}$$

where $D = 2^n$. This has the effect of mapping Boolean functions to vectors in the D-dimensional Hilbert space, $\mathcal{H}_D$, and we shall do the same. The final qubit is not entangled with the remaining n qubits and can be discarded. The vectors $\sum_{x=0}^{D-1}(-1)^{f(x)}|x\rangle$ that are produced by balanced functions are orthogonal to those produced by constant functions. This is why the Deutsch-Jozsa problem is easy to solve quantum mechanically. In our case, the vectors produced by functions in $\mathcal{W}_k$ are not orthogonal to those produced by balanced functions. However, unambiguous quantum state filtering provides a probabilistic quantum algorithm for the optimal solution of this problem.

In order to apply the filtering solution, we note that both functions in $\mathcal{W}_k$ are mapped, up to an overall sign, to the same vector in $\mathcal{H}_D$, which we shall call $|w_k\rangle$. The vectors that correspond to balanced functions are contained in the subspace, $\mathcal{H}_b$, of $\mathcal{H}_D$, where $\mathcal{H}_b = \{|v\rangle \in \mathcal{H}_D | \sum_{x=0}^{D-1} v_x = 0\}$, and $v_x = \langle x|v\rangle$. This subspace has dimension $2^n - 1 = D - 1$, and it is possible to choose an orthonormal basis, $\{|v_i\rangle | i = 2, \ldots D\}$, for it in which each basis element corresponds to a particular balanced Boolean function [40].

Let us first see how the filtering procedure performs when applied to the problem of distinguishing $|w_k\rangle (= |\psi_1\rangle)$ from the set of the $D-1$ orthonormal basis states, $|v_i\rangle (= |\psi_i\rangle)$, in $\mathcal{H}_b$. We assume their a priori probabilities to be equal, i.e. $\eta_i = \eta = (1 - \eta_1)/(D - 1)$ for $i = 2, \ldots D$, where η_1 is the a priori probability for $|w_k\rangle$. For $\|\psi_1^{\|}\|^2 = \|w_k^{\|}\|^2 \equiv f_k$ we obtain $f_k = (2^k - 1)/2^{2k-2}$. Then the average overlap, S_k, between $|w_k\rangle$ and the set of balanced basis vectors can be written as

$$S_k = \frac{1 - \eta_1}{D - 1} f_k \ , \tag{11.46}$$

in terms of f_k [40]. The failure probabilities are given by the filtering result, using $S = S_k$ and, to good approximation, the POVM result holds when $1/2^{k-2} \leq D\eta_1 \leq 2^{k-2}$. For example, in the case in which all of the a priori probabilities are equal, i.e. $\eta_1 = 1/D$, we find that $Q_1 = Q_2 \equiv Q^{SQM} = (1 + f_k)/D$ where SQM stands for Standard Quantum Measurement (or von Neumann projective measurement). To good approximation, $Q^{POVM}/Q^{SQM} = 4/2^{k/2}$, which, for $k \gg 1$, shows that the POVM can perform significantly better than the von Neumann measurements.

Now that we know how this procedure performs on the basis vectors in $\mathcal{H}_b$, we shall examine its performance on any balanced function, i.e. we apply it to the problem of distinguishing $|w_k\rangle$ from the set of all states in $\mathcal{H}_b$ that correspond to balanced functions. The number of such states is $N = D!/(D/2)!^2$ and we again assume their a priori probabilities to be equal, $\eta = (1 - \eta_1)/N$. It can be shown [40] that the average overlap between $|w_k\rangle$ and the set $\{|v\rangle\}$ is given by the same expression, (11.46), as in the previous case. Therefore, much of what was said in the previous paragraph remains

valid for this case, as well, with one notable difference. The case $\eta_1 = 1/D$ now does not correspond to equal a priori probability for the states but, rather, to a priori weight of the sets that is proportional to their dimensionality. In this case it is the POVM that performs best. In the case of equal a priori probability for all states, $\eta_1 = 1/(N+1)$, we are outside of the POVM range of validity and it is the first standard quantum measurement (SQM1) that performs best. Both the POVM and the SQM1 are good methods for distinguishing functions in $\mathcal{W}_k$ from balanced functions. Which one is better depends on the a priori probabilities of the functions.

Classically, in the worst case, one would have to evaluate a function $2^n[(1/2)+(1/2^k)]+1$ times to determine if it is in $\mathcal{W}_k$ or if it is an even function. Using quantum information processing methods, one has a very good chance of determining this with only one function evaluation. This shows that Deutsch-Jozsa-type algorithms need not be limited to constant functions; certain kinds of biased functions can be discriminated as well.

11.3 State Discrimination with Minimum Error

11.3.1 Introductory Remarks

In the previous chapter we have required that, whenever a definite answer is returned after a measurement on the state, the result should be unambiguous, at the expense of allowing inconclusive outcomes to occur. For many applications in quantum communication, however, one wants to have conclusive results only. This means that errors are unavoidable when the states are non-orthogonal. Based on the outcome of the measurement, in each single case then a guess has to be made as to what the state of the quantum system was. This procedure is known as *quantum hypothesis testing*. The problem consists in finding the optimum measurement strategy that minimizes the probability of errors.

Let us state the optimization problem a little more precisely. In the most general case, we want to distinguish, with minimum probability of error, between N given states of a quantum system ($N \geq 2$), being characterized by the density operators ρ_j ($j = 1, 2, \ldots, N$) and occurring with the given a priori probabilities η_j which sum up to unity. The measurement can be formally described with the help of a set of detection operators Π_j that refer to the possible measurement outcomes [3, 4]. They are defined in such a way that $\mathrm{Tr}(\rho\Pi_j)$ is the probability to infer the system is in the state ρ_j if it has been prepared in a state ρ. Since the probability is a real non-negative number, the detection operators have to be Hermitean and positive-semidefinite. In the error-minimizing measurement scheme the measurement is required to be exhaustive and conclusive in the sense that in each single case one of the N possible states is identified with certainty and inconclusive results do not

occur. This leads to the requirement

$$\sum_{j=1}^{N} \Pi_j = I_{D_S}, \tag{11.47}$$

where I_{D_S} denotes the unit operator in the D_S-dimensional physical state space of the quantum system. The overall probability $P_{\rm err}$ to make an erroneous guess for any of the incoming states is then given by

$$P_{\rm err} = 1 - P_{\rm corr} = 1 - \sum_{j=1}^{N} \eta_j {\rm Tr}(\rho_j \Pi_j) \tag{11.48}$$

with $\sum_j \eta_j = 1$. Here we introduced the probability $P_{\rm corr}$ that the guess is correct. In order to find the minimum-error measurement strategy, one has to determine the specific set of detection operators that minimizes the value of $P_{\rm err}$ under the constraint given by (11.47). By inserting these optimum detection operators into (11.48), the minimum error probability $P_{\rm err}^{\rm min} \equiv P_E$ is determined. The explicit solution to the error-minimizing problem is not trivial and analytical expressions have been derived only for a few special cases.

11.3.2 Distinguishing Two Quantum States with Minimum Error

The Helstrom Formula

For the case that only two states are given, either pure or mixed, the minimum error probability, P_E, was derived in the mid 70s by Helstrom [3] in the framework of quantum detection and estimation theory. We find it instructive to start by analyzing the two-state minimum-error measurement with the help of an alternative method (cf. [41, 42]) that allows us to gain immediate insight into the structure of the optimum detection operators, without applying variational techniques. Starting from (11.48) and making use of the relations $\eta_1 + \eta_2 = 1$ and $\Pi_1 + \Pi_2 = I_{D_S}$ that have to be fulfilled by the a priori probabilities and the detection operators, respectively, we see that the total probability to get an erroneous result in the measurement is given by

$$P_{\rm err} = 1 - \sum_{j=1}^{2} \eta_j {\rm Tr}(\rho_j \Pi_j) = \eta_1 {\rm Tr}(\rho_1 \Pi_2) + \eta_2 {\rm Tr}(\rho_2 \Pi_1). \tag{11.49}$$

This can be alternatively expressed as

$$P_{\rm err} = \eta_1 + {\rm Tr}(\Lambda \Pi_1) = \eta_2 - {\rm Tr}(\Lambda \Pi_2), \tag{11.50}$$

where we introduced the Hermitean operator

$$\Lambda = \eta_2 \rho_2 - \eta_1 \rho_1 = \sum_{k=1}^{D_S} \lambda_k |\phi_k\rangle\langle\phi_k|. \tag{11.51}$$

Here the states $|\phi_k\rangle$ denote the orthonormal eigenstates belonging to the eigenvalues λ_k of the operator Λ. The eigenvalues are real, and without loss of generality we can number them in such a way that

$$\begin{aligned} \lambda_k &< 0 \qquad \text{for} \qquad 1 \le k < k_0, \\ \lambda_k &> 0 \qquad \text{for} \qquad k_0 \le k \le D, \\ \lambda_k &= 0 \qquad \text{for} \qquad D < k \le D_S. \end{aligned} \tag{11.52}$$

By using the spectral decomposition of Λ, we get the representations

$$P_{\rm err} = \eta_1 + \sum_{k=1}^{D_S} \lambda_k \langle\phi_k|\Pi_1|\phi_k\rangle = \eta_2 - \sum_{k=1}^{D_S} \lambda_k \langle\phi_k|\Pi_2|\phi_k\rangle. \tag{11.53}$$

Our optimization task now consists in determining the specific operators Π_1, or Π_2, respectively, that minimize the right-hand side of (11.53) under the constraint that

$$0 \le \langle\phi_k|\Pi_j|\phi_k\rangle \le 1 \qquad (j = 1, 2) \tag{11.54}$$

for all eigenstates $|\phi_k\rangle$. The latter requirement is due to the fact that $\mathrm{Tr}(\rho\Pi_j)$ denotes a probability for any ρ. From this constraint and from (11.53) it immediately follows that the smallest possible error probability, $P_{\rm err}^{\rm min} \equiv P_E$, is achieved when the detection operators are chosen in such a way that the equations $\langle\phi_k|\Pi_1|\phi_k\rangle = 1$ and $\langle\phi_k|\Pi_2|\phi_k\rangle = 0$ are fulfilled for eigenstates belonging to negative eigenvalues, while eigenstates corresponding to positive eigenvalues obey the equations $\langle\phi_k|\Pi_1|\phi_k\rangle = 0$ and $\langle\phi_k|\Pi_2|\phi_k\rangle = 1$. Hence the optimum detection operators can be written as

$$\Pi_1 = \sum_{k=1}^{k_0-1} |\phi_k\rangle\langle\phi_k|, \qquad \Pi_2 = \sum_{k=k_0}^{D_S} |\phi_k\rangle\langle\phi_k|, \tag{11.55}$$

where the expression for Π_2 has been supplemented by projection operators onto eigenstates belonging to the eigenvalue $\lambda_k = 0$, in such a way that $\Pi_1 + \Pi_2 = I_{D_S}$. Obviously, provided that there are positive as well as negative eigenvalues in the spectral decomposition of Λ, the minimum-error measurement for discriminating two quantum states is a von Neumann measurement that consists in performing projections onto the two orthogonal subspaces spanned by the set of states $\{|\phi_1\rangle, \ldots, |\phi_{k_0-1}\rangle\}$, on the one hand, and $\{|\phi_{k_0}\rangle, \ldots, |\phi_{D_S}\rangle\}$, on the other hand. An interesting special case arises when negative eigenvalues do not exist. In this case it follows that $\Pi_1 = 0$ and $\Pi_2 = I_{D_S}$ which means that the minimum error probability can be achieved by always guessing the quantum system to be in the state ρ_2, without performing any measurement at all. Similar considerations hold true in the absence of positive eigenvalues. We note that these findings are in

agreement with the recently gained insight [44] that measurement does not always aid minimum-error discrimination. By inserting the optimum detection operators into (11.50) the minimum error probability is found to be [42]

$$P_E = \eta_1 - \sum_{k=1}^{k_0-1} |\lambda_k| = \eta_2 - \sum_{k=k_0}^{D} |\lambda_k|. \tag{11.56}$$

Taking the sum of these two alternative representations and using $\eta_1+\eta_2 = 1$, we arrive at

$$P_E = \frac{1}{2}\left(1 - \sum_k |\lambda_k|\right) = \frac{1}{2}\left(1 - \mathrm{Tr}|\Lambda|\right), \tag{11.57}$$

where $|\Lambda| = \sqrt{\Lambda^\dagger \Lambda}$. Together with (11.48) this immediately yields the well-known Helstrom formula [3] for the minimum error probability in discriminating ρ_1 and ρ_2,

$$P_E = \frac{1}{2}\left(1 - \mathrm{Tr}|\eta_2\rho_2 - \eta_1\rho_1|\right) = \frac{1}{2}\left(1 - \|\eta_2\rho_2 - \eta_1\rho_1\|\right). \tag{11.58}$$

In the special case that the states to be distinguished are the pure states $|\psi_1\rangle$ and $|\psi_2\rangle$, this expression reduces to [3]

$$P_E = \frac{1}{2}\left(1 - \sqrt{1 - 4\eta_1\eta_2|\langle\psi_1|\psi_2\rangle|^2}\right). \tag{11.59}$$

The set-up of the detectors that achieve the optimum error probabilities is particularly simple for the case of equal a priori probabilities. Two orthogonal detectors, placed symmetrically around the two pure states, will do the task, as shown in Fig. 11.6. The simplicity is particularly striking when one compares this set-up to that of Fig. 11.3, that displays the corresponding POVM set-up for optimal unambiguous discrimination.

In order to provide a simple and intuitive physical picture of an error-minimizing state discrimination measurement, let us consider the experiment performed by Barnett and Rijs [43] for distinguishing two equiprobable non-orthogonal single-photon polarization states, produced with the help of strongly attenuated light pulses. Imagine, we are given a series of very weak light pulses and we know in advance that, by preparation, each of the pulses is linearly polarized, with equal probability either in the direction $\boldsymbol{e}_1$ or in the direction $\boldsymbol{e}_2$ with $\boldsymbol{e}_{1,2} = \cos\theta\,\boldsymbol{e}_x \pm \sin\theta\,\boldsymbol{e}_y$, where $0 \le \theta \le \pi/4$. Both kinds of pulses are assumed to have equal intensity and to contain on the average much less than one photon, so that the probability to obtain two photodetector clicks from a single pulse is negligible. Using a linear polarizer, we want to determine for each photon which polarization it had, by making the guess that it had polarization $\boldsymbol{e}_1$ when the photon is transmitted by the polarizer and had polarization $\boldsymbol{e}_2$ when the photon is absorbed in the

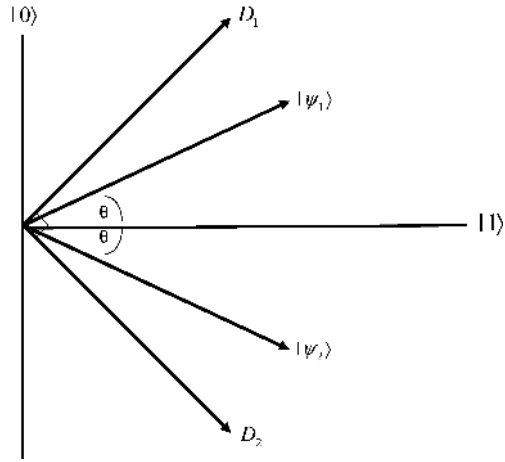

Fig. 11.6. Detector configuration for the optimum minimum-error discrimination of two pure states with equal a priori probabilities. A von Neumann measurement with two orthogonal detectors placed symmetrically around $|\psi_1\rangle$ and $|\psi_2\rangle$ will achieve the optimum.

polarizer. The problem of finding the optimum measurement strategy then amounts to determining the optimum orientation of the polarizer that minimizes the probability to make a wrong guess. It is easy to calculate that the error probability is smallest when the polarizer is oriented symmetrically with respect to the two given polarization directions, i. e. when its transmission direction is given by the unit vector $\boldsymbol{e}_p = (\boldsymbol{e}_x + \boldsymbol{e}_y)/\sqrt{2}$. In this case the fraction of the intensity that is transmitted from the first pulse, which is also the probability, that the photon is indeed transmitted provided it is of the first kind, follows from the projection of $\boldsymbol{e}_1$ onto $\boldsymbol{e}_p$ and is given by $\cos^2(\pi/4 - \theta)$. The same quantity describes the probability that the photon is absorbed, provided it is of the second kind, as can be found by projecting $\boldsymbol{e}_2$ onto the vector perpendicular to $\boldsymbol{e}_p$. Hence the resulting error probability is determined by

$$P_E = 1 - \cos^2(\frac{\pi}{4} - \theta) = \frac{1}{2}(1 - \sin 2\theta) = \frac{1}{2}(1 - \sqrt{1 - |\boldsymbol{e}_1 \boldsymbol{e}_2|^2}). \qquad (11.60)$$

which reproduces (11.59) with $\eta_1 = \eta_2 = 1/2$ when the scalar product $\boldsymbol{e}_1\boldsymbol{e}_2$ is replaced by the overlap $\langle\psi_1|\psi_2\rangle$.

Minimum-Error Discrimination Versus Unambiguous Discrimination for Two Mixed States

In the error-minimizing scheme for discriminating two different mixed states ρ_1 and ρ_2 of a quantum system, a non-zero probability of making a correct

guess can always be achieved. However, it is clear that we can only distinguish the two mixed states unambiguously when, in the D_S-dimensional physical state space of the quantum system, there exists at least one state vector that has a non-zero probability of occurring in the first of the mixed states but will never be found in the second mixed state. Recently it has been shown that the maximum success probability for unambiguously discriminating the two given mixed states does not exceed a certain upper bound [35], depending on the states and their a priori probabilities. Consequently, the minimum failure probability, Q_F, cannot be smaller than a certain lower bound, Q_L. We now investigate the relation between this lower bound on the failure probability, on the one hand, and the minimum error probability for discriminating the same two states, on the other hand. The procedure that we shall apply is closely related to the derivation of inequalities between the fidelity and the trace distance [1].

Let us first consider the failure probability Q_F for unambiguously discriminating the two mixed states. From the result derived in [35] it follows that

$$Q_F \geq \begin{cases} 2\sqrt{\eta_1\eta_2}\, F(\rho_1,\rho_2) \equiv Q_L & \text{if } F(\rho_1,\rho_2) < \sqrt{\frac{\eta_{\min}}{\eta_{\max}}} \\ \eta_{\min} + \eta_{\max}[F(\rho_1,\rho_2)]^2 \geq Q_L & \text{otherwise,} \end{cases} \tag{11.61}$$

where in the second line we have used the fact that the arithmetic mean cannot be smaller than the geometric mean. Here $\eta_{\min}(\eta_{\max})$ is the smaller (larger) of the two a priori probabilities η_1 and η_2, and F is the fidelity defined as $F(\rho_1,\rho_2) = \mathrm{Tr}\, O$, where $O = \left(\sqrt{\rho_2}\,\rho_1\sqrt{\rho_2}\right)^{1/2}$. To allow for a comparison between P_E and Q_F, or P_E and Q_L, respectively, we need a suitable orthonormal basis. It has been shown [1,41] that when the basis states are chosen to be the eigenstates $\{|l\rangle\}$ of the operator $\rho_2^{-1/2} O \rho_2^{-1/2}$, the fidelity takes the form

$$F(\rho_1,\rho_2) = \sum_l \sqrt{\langle l|\rho_1|l\rangle\langle l|\rho_2|l\rangle} = \sum_l \sqrt{r_l\, s_l}. \tag{11.62}$$

Here $\sum_l |l\rangle\langle l| = I$, and we introduced the abbreviations $r_l = \langle l|\rho_1|l\rangle$ and $s_l = \langle l|\rho_2|l\rangle$. The lower bound on the failure probability then obeys the equation

$$1 - Q_L = 1 - 2\sqrt{\eta_1\eta_2} \sum_l \sqrt{r_l\, s_l} = \sum_l \left(\sqrt{\eta_1 r_l} - \sqrt{\eta_2 s_l}\right)^2, \tag{11.63}$$

where the second equality sign is due to the relation $\eta_1 + \eta_2 = 1$ and to the normalization conditions $\mathrm{Tr}\rho_1 = \sum_l r_l = 1$ and $\mathrm{Tr}\rho_2 = \sum_l s_l = 1$.

Now we estimate the minimum error probability P_E, using the same set of basis states $\{|l\rangle\}$. From (11.57) and from the fact that $\langle\phi_k|\phi_k\rangle = \sum_l |\langle\phi_k|l\rangle|^2 = 1$ it follows that

$$\begin{aligned} 1 - 2P_E &= \sum_k |\lambda_k| = \sum_l \sum_k |\lambda_k||\langle\phi_k|l\rangle|^2 \\ &\geq \sum_l \left|\sum_k \lambda_k |\langle\phi_k|l\rangle|^2\right| = \sum_l |\langle l|\Lambda|l\rangle|, \end{aligned} \tag{11.64}$$

where the last equality sign is due to the spectral decomposition (11.51). After expressing Λ in terms of the density operators describing the given states, we arrive at

$$1 - 2P_E \geq \sum_l |\langle l|\eta_1\rho_1 - \eta_2\rho_2|l\rangle| = \sum_l |\sqrt{\eta_1 r_l} - \sqrt{\eta_2 s_l}|\,|\sqrt{\eta_1 r_l} + \sqrt{\eta_2 s_l}|\,. \tag{11.65}$$

By comparing the expressions on the right-hand sides of (11.63) and (11.65), respectively, it becomes immediately obvious that $1 - 2P_E \geq 1 - Q_L$, or

$$P_E \leq \frac{1}{2} Q_L\ . \tag{11.66}$$

This means that for two arbitrary mixed states, occurring with arbitrary a priori probabilities, the smallest possible failure probability in unambiguous discrimination is at least twice as large as the smallest probability of errors in minimum-error discrimination of the same states.

11.3.3 The General Strategy for Minimum-Error State Discrimination

Formal Solution for N Mixed States

We return now to the general problem of discriminating with minimum error between N given mixed states, where N is an arbitrary number. As has been outlined already in the introductory remarks, the problem amounts to finding the specific optimum detection operators that minimize the expression (11.48) under the constraint (11.47). It has been shown by Holevo [45] and Yuen *et* al. [46] that the set of detection operators $\{\Pi_j\}$ determining the optimum measurement strategy must satisfy the necessary and sufficient conditions

$$\Pi_k(\eta_k\rho_k - \eta_j\rho_j)\Pi_j = 0 \tag{11.67}$$

$$\sum_j \eta_j\rho_j\Pi_j - \eta_k\rho_k \geq 0, \tag{11.68}$$

$(1 \leq j, k \leq N)$, where the last equation expresses the fact that the eigenvalues of the operator on the left-hand side are non-negative. These conditions can be understood from the following considerations. In order to take the constraint (11.47) into account, it is possible to introduce a Lagrange operator Γ, in analogy to a Lagrangian multiplier. The optimization task is then equivalent to maximizing the operator functional

$$P_{\rm corr}(\{\Pi_j\}, \Gamma) = \sum_j {\rm Tr}[(\eta_j\rho_j - \Gamma)\Pi_j] + {\rm Tr}\Gamma, \tag{11.69}$$

with Γ and each of the Π_j varying independently. Since $P_{\rm corr}$ is real, Γ has to be Hermitean. The necessary condition for an extremum, $\delta P_{\rm corr} = 0$, yields for any $j = 1, \ldots, N$ the requirement that $\mathrm{Tr}[(\eta_j \rho_j - \Gamma)\delta\Pi_j] = 0$. The detection operators can be varied by starting from their eigenstate expansion $\Pi_j = \sum_l \pi_{jl} |\mu_{jl}\rangle\langle\mu_{jl}|$ and performing variations in the state vectors, yielding $\delta\Pi_j = \sum_l \pi_{jl} |\mu_{jl}\rangle\langle\delta\mu_{jl}| + H.A.$, where $\pi_{jl} \geq 0$. This leads to the necessary extremal condition [25]

$$(\eta_j \rho_j - \Gamma)\Pi_j = 0, \tag{11.70}$$

from which it follows that for any j, k the equations $\eta_j \rho_j \Pi_j = \Gamma \Pi_j$ and $\eta_k \Pi_k \rho_k = \Pi_k \Gamma^\dagger$ have to be fulfilled, where Γ is Hermitean. By multiplying the first equation from the left with Π_k and the second from the right with Π_j and taking the difference, the representation (11.67) for the necessary condition becomes immediately obvious. Now we discuss the second condition. First we note that

$$\Gamma = \sum_j \eta_j \rho_j \Pi_j, \tag{11.71}$$

which follows from (11.70) and from the constraint $\sum_j \Pi_j = I$. As can be seen from (11.69) and (11.70), the extremal value is given by $P_{\rm corr} = \mathrm{Tr}\Gamma$. Certainly for the extremum of $P_{\rm corr}$ to be a maximum, its value cannot be smaller than the probability of correct guesses that would follow from other choices of the set of detection operators, in particular from the choice $\{\Pi_j\} = \{I\ \delta_{j,k}\}$, leading to $P_{\rm corr} = \mathrm{Tr}(\eta_k \rho_k)$. This means that for the extremum to be a maximum the relation $\mathrm{Tr}\Gamma \geq \mathrm{Tr}(\eta_k \rho_k)$ has to be fulfilled for any k, which is indeed guaranteed by the requirement (11.68). The conditions (11.67) and (11.68) implicitly determine the solution of the general minimum-error discrimination problem.

Survey of Explicitly Solvable Special Cases

In the following we give a brief overview that summarizes the cases where explicit analytical expressions for the optimum detection operators, and hence for the minimum error probability, have been determined from the implicit general solution. The case that only two states are given, $N = 2$, has been already extensively discussed in the previous section. When the two given states are pure and occur with equal prior probability, the Helstrom bound derived there is a special case of a more general result, referring to the class of equiprobable and symmetric pure states. For these states the solution of the minimum-error discrimination problem, as derived by Ban *et* al. [47], will be discussed in a separate section. Recently Eldar *et al.* [48] and Chou and Hsu [49] obtained an extension of this solution to the case of N equiprobable states that are symmetric and mixed. A few other cases have been solved analytically, too. They include certain classes of linearly independent states [50]

and also equiprobable states the projectors of which sum up to the identity [46]. Moreover, Barnett [51] found the minimum-error strategy for multiple symmetric pure states, and Andersson *et* al. [52] solved the case of three mirror symmetric pure states. With respect to the general problem of discriminating an arbitrary number of mixed states, Hunter [44] found the condition and gave the solution for those cases when the best strategy consists in making no measurement at all, but simply to guess always the state with the highest a priori probability. We encountered an example of such a case when considering the discrimination between two mixed states.

Apart from the analytical solutions, the minimum-error strategy has been also investigated numerically. In particular, Ježek *et al.* [53] proposed an algorithm for finding the optimum measurement by applying the theory of semi-definite programming.

Generalized Measurements

When the N given states of the quantum system are linearly independent, the minimum-error strategy for state discrimination is always a von Neumann measurement, as has been recently proved by Eldar [54]. This means that the detection operators are mutually orthogonal projection operators in the Hilbert space of the quantum system, fulfilling the relation $\Pi_j \Pi_k = \delta_{jk} \Pi_j$, where $1 \leq j, k \leq N$. We note that the quantum states are called linearly independent when the combined set of all the eigenvectors of the density operators ρ_j $(j = 1, \ldots, N)$ forms a set of linearly independent state vectors. In this case the error-minimizing discrimination measurement is a projective measurement that can be realized by performing measurements on the original quantum system alone. Generally, however, this need not always be the case. In particular, when the number of different measurement outcomes, or the number of detection operators, respectively, exceeds the dimensionality of the physical state space of the quantum system, the detection operators cannot be represented by projection operators in that state space, as becomes immediately obvious from the constraint $\sum_j \Pi_j = I_{D_S}$. The state discrimination measurement then has to be described as a generalized measurement, based on positive-operator valued measures, and the detection operators Π_j are also called POVM-elements. According to Neumark's theorem [13], any generalized measurement can be realized with the help of a unitary transformation and a projective measurement in an extended Hilbert space.

In order to illustrate these general considerations we find it worthwhile to briefly consider a prominent example for a generalized minimum-error measurement, which consists in discriminating between the three states of a single qubit defined as

$$|\psi_1\rangle = -\frac{1}{2}\left(|0\rangle + \sqrt{3}\,|1\rangle\right),$$

$$|\psi_2\rangle = -\frac{1}{2}\left(|0\rangle - \sqrt{3}\,|1\rangle\right),$$
$$|\psi_3\rangle = |0\rangle, \tag{11.72}$$

where $|0\rangle$ and $|1\rangle$ are the orthonormal basis states of the qubit. The three given states form an overcomplete set of symmetric states that is known as the trine ensemble [55]. Provided that the occurrence of each of the three states is equally probable, the optimum detection operators for distinguishing among them with minimum error, or the optimum POVM-elements, respectively, are given by [55]

$$\Pi_j = A_j^\dagger A_j = \frac{2}{3}\,|\psi_j\rangle\langle\psi_j|, \tag{11.73}$$

for $j = 1, 2, 3$. The probability of correctly identifying any of the trine states in the error-minimizing measurement is $2/3$ and the probability of making an error is $1/3$. Since the states $|\psi_j\rangle$ are normalized, it is clear that the operators Π_j do not represent projection operators. For a physical implementation of the generalized measurement, means have to be found for extending the originally two-dimensional Hilbert space of the system. Then a unitary transformation on the extended system of basis states has to be performed in such a way that a final von Neumann measurement realizes the specific projective measurement that is necessary for applying Neumark's theorem.

As discussed before, there are two conceptually different ways of achieving an extension of the dimensionality of the Hilbert space. The first amounts to defining the original quantum system in such a way that auxiliary quantum states can be directly added. The extended Hilbert space is then the direct sum of the Hilbert space spanned by the states of the original system and of the Hilbert space spanned by the auxiliary states, being also called ancilla states. For the trine ensemble, it is possible to associate the three two-dimensional non-normalized detection states $A_j = \sqrt{2/3}\,|\psi_j\rangle$ with three orthonormal states in three dimensions, given by

$$|\psi_j\rangle = \sqrt{\frac{2}{3}}\,|\psi_j\rangle + \sqrt{\frac{1}{3}}\,|2\rangle, \tag{11.74}$$

where the auxiliary normalized state $|2\rangle$ is chosen to be orthogonal to the two basis states $|0\rangle$ and $|1\rangle$. By performing the von-Neumann measurement that consists of the three projections $|\psi_j\rangle\langle\psi_j|$ $(j = 1, 2, 3)$ in the enlarged, i. e. three-dimensional Hilbert space, the required generalized measurement is realized in the original two-dimensional Hilbert space of the qubit [56]. To implement this scheme, an original atomic qubit system can be defined to consist of only two electronic states of a multi-level atom. The unitary transformation necessary for the generalized measurement can then be accomplished with the help of a third level, by appropriately redistributing the population of the three atomic levels using sequences of Raman transitions induced by specially taylored classical pulses [56]. Finally the resulting level

population is detected for performing the projective measurement. Explicit theoretical proposals have been made for using this scheme in order to realize both optimum unambiguous discrimination between two nonorthogonal pure qubit states, and minimum-error discrimination for the trine ensemble [56]. Similarly, a single-photon qubit system could be represented by only two out of three possible input modes, or input ports, respectively, of an optical network built of beam splitters and phase shifters. The unitary transformation then would be implemented by this linear network, and the projective measurement could be performed by detecting the photon at one of the output ports. It is obvious that the direct-sum representation of extending the Hilbert space relies on assuming that the original qubit secretly consists of two components of a qutrit [57].

The second way of enlarging the Hilbert space relies on coupling the original system to an auxiliary system, or ancilla system, with the help of a physical interaction. The Hilbert space of the combined system is the tensor product of the Hilbert spaces of both subsystems. As has been stated by Jozsa, *et al.* [58], from a physical point of view the adjoining of an additional ancilla system is the only available means of extending a space while retaining the original system intact. The formalism of (nonorthogonal) POVMs is then a mathematical artifice that expresses the residual effect on the original system when a von Neumann measurement is performed on the combined system after the interaction [58]. Clarke, *et al.* demonstrated experimentally the minimum-error discrimination for the trine ensemble represented by polarization states of a single photon, using an interferometric set-up [55]. They also performed the optimal error minimizing measurements on the tetrad ensemble, which, as its name implies, consists of four states [55]. The points corresponding to these states on the Bloch sphere lie on the corners of a tetrahedron.

Error-Minimizing Discrimination with a Fixed Number of Inconclusive Results

As an extension of the measurement strategy described so far, the error minimizing discrimination strategy has been also studied under the condition that inconclusive results are now allowed to occur, but with a fixed prescribed probability. This probability is assumed to be smaller than the minimum failure probability resulting from the measurement scheme for optimum unambiguous discrimination, where errors do not occur at all. The optimum measurement minimizing the error under this prescribed condition is then intermediate between generic minimum-error discrimination and optimum unambiguous discrimination. It was first investigated for pure states by Chefles and Barnett [59] and by Zhang *et* al. [60]. Later the method was generalized to the case of mixed states by Fiurášek and Ježek [25] and by Eldar [37]. The possible occurrence of inconclusive results has to be accounted for in the basic equations by introducing an additional detection operator Π_0

such that $\mathrm{Tr}(\rho\Pi_0)$ describes the probability to get an inconclusive outcome provided the system is in a state ρ. The given fixed value of this probability, or the given failure probability, respectively, modifies the constraint on the detection operators. By optimization, the minimum probability for the occurrence of errors in the conclusive outcomes can be determined.

11.3.4 Selected Problems of Minimum-Error Discrimination

Distinguishing N Symmetric Pure States

We conclude our review of minimum-error discrimination by a more detailed consideration of a few special problems. In particular, in this connection we also present the results of some of our own recent investigations. To begin with, we recall a pure-state discrimination problem that is exactly solvable and has found wide application in quantum communication. It consists in the so called *square-root measurement* that discriminates with minimum error between N equally probable symmetric states. Symmetric pure states are defined in such a way that each state results from its predecessor by applying a unitary operator V in a cyclic way [47],

$$|\psi_j\rangle = V|\psi_{j-1}\rangle = V^{j-1}|\psi_1\rangle, \qquad |\psi_1\rangle = V|\psi_N\rangle, \tag{11.75}$$

implying that $V^N = I$. For the case that the states occur with equal a priori probability, i. e. that $\eta_j = 1/N$ for each of the states, Ban *et* al. found that the optimum detection operators for minimum-error discrimination are given by [47]

$$\Pi_j = A_j^\dagger A_j = B^{-1/2}|\psi_j\rangle\langle\psi_j|B^{-1/2} \equiv |\mu_j\rangle\langle\mu_j|, \tag{11.76}$$

where

$$B = \sum_{j=1}^{N} |\psi_j\rangle\langle\psi_j|. \tag{11.77}$$

The states $|\mu_j\rangle = B^{-1/2}|\psi_j\rangle$ are in general non-normalized and are called detection states. It is obvious that the special structure of the detection operators, or of the detection states, respectively, suggests the name "square-root-measurement". The minimum error probability P_E for this measurement is [47]

$$P_E = 1 - \frac{1}{N}\sum_{j=1}^{N} |\langle\mu_j|\psi_j\rangle|^2, \tag{11.78}$$

in accordance with the fact that in the corresponding optimized measurement scheme the quantum system is inferred to have been prepared in the state

$|\psi_j\rangle$ provided that the state $|\mu_j\rangle$ is detected. When the detection states $|\mu_j\rangle$ are orthonormal, the detection operators are projection operators and the minimum-error measurement is a von-Neumann measurement, otherwise it is a generalized measurement. The latter always holds true when the number of states exceeds the dimensionality of the physical state space of the quantum system, as can be immediately seen from the fact that the detection operators have to sum up to the unit operator in that space. In this case the given states are linearly dependent and form an overcomplete set in the Hilbert space of the system.

Let us apply the general solution in order to investigate minimum-error discrimination for the set of the N symmetric states

$$|\psi_j\rangle = \sum_{k=1}^{D} c_k \, \mathrm{e}^{\mathrm{i}\frac{2\pi}{N}j(k-1)} \, |\gamma_k\rangle, \qquad (N \geq D), \tag{11.79}$$

where the coefficients c_k are arbitrary non-zero complex numbers with $\sum_k |c_k|^2 = 1$, and the states $|\gamma_k\rangle$ $(k = 1, \ldots, D)$ form a D-dimensional orthonormal basis. The given symmetric states are non-orthogonal except for the case that both the conditions $N = D$ and $|c_k|^2 = 1/N$ are fulfilled. For distinguishing them with minimum error, provided that they occur with equal a priori probability, we obtain the optimum detection states

$$|\mu_j\rangle = \frac{1}{\sqrt{N}} \sum_{k=1}^{D} \frac{c_k}{|c_k|} \, \mathrm{e}^{\mathrm{i}\frac{2\pi}{N}jk} \, |\gamma_k\rangle, \tag{11.80}$$

yielding the minimum error probability [61]

$$P_E = 1 - \frac{1}{N} \left(\sum_{k=1}^{D} |c_k| \right)^2 . \tag{11.81}$$

When $N > D$ the given symmetric states are linearly dependent and form an overcomplete set. In this case the detection states $|\mu_j\rangle$ are non-orthogonal and non-normalized, with $\langle \mu_j | \mu_j \rangle = D/N$. When $N = D$, however, the states $|\psi_j\rangle$ are linearly independent and therefore can be discriminated also unambiguously, as has been pointed out in the corresponding section of this review. Assuming equal a priori probabilities, the minimum failure probability, Q_F, for unambiguous discrimination of symmetric states has been derived to be $Q_F = 1 - N \, \min |c_k|^2$ [24]. Comparing this with the expression for P_E, the minimum error probability is found to be smaller than Q_F. It is worth mentioning that the minimum error probability, P_E, on the one hand, and the failure probability, Q_F, on the other hand, have been considered as distinguishability measures for ordering different ensembles of N equally probable symmetric pure states, and it has been found that these two measures impose different orderings [62].

We used the preceding general results for studying minimum-error discrimination between m-photon polarization states, referring to a fixed number m of indistinguishable photons ($m = 1, 2, \dots$). These states are superposition states of the $m + 1$ orthonormal polarization basis states

$$|\gamma_k^{(m)}\rangle = \frac{1}{\sqrt{(m-k)!k!}}(a_1^\dagger)^{m-k}(a_2^\dagger)^k|0\rangle \equiv |m-k, k\rangle, \tag{11.82}$$

where $k = 0, 1, \dots m$. The basis states correspond to the $m + 1$ different possibilities of distributing m indistinguishable photons among two orthogonal polarization modes, characterized by the photon creation operators $a_1^\dagger$ and $a_2^\dagger$, respectively, and $|0\rangle$ is the vacuum state of the field. A specific set of symmetric m-photon polarization states is defined by

$$|\psi_j^{(m)}\rangle = \frac{1}{\sqrt{m!}}\left(\cos\theta\, a_1^\dagger + \sin\theta\, \mathrm{e}^{\mathrm{i}\frac{2\pi}{N}j}\, a_2^\dagger\right)^m |0\rangle, \tag{11.83}$$

where $j = 1, \dots N$. Interestingly, for $N = 4$ and $\theta = \pi/4$ these symmetric states are identical with the states that result when the standard protocol for quantum key distribution [65] is applied to m-photon pulses. Assuming again equal a priori probabilities, the minimum-error probability for discriminating the states reads for $m \leq N - 1$

$$P_E^{(m)}(\theta) = 1 - \frac{1}{N}\left[\sum_{k=0}^{m}\sqrt{\binom{m}{k}}\cos^{m-k}\theta\, \sin^k\theta\right]^2. \tag{11.84}$$

For the case of two-photon-polarization states, $m = 2$, we found [61] that for $\theta = \pi/4$ the minimum error probability is given by $P_E = 1-(3+2\sqrt{2})/(2N)$ which is smaller than the value that would result for the corresponding single-photon polarization states, being $P_E = 1-2/N$. In general, the states $|\psi_j^{(m)}\rangle$ can be considered to consist of m identical copies of indistinguishable qubits, or photons, respectively, being each in the state $\cos\theta|1,0\rangle + \sin\theta \mathrm{e}^{\mathrm{i}\frac{2\pi}{N}j}|0,1\rangle$. Therefore our results show that by performing a joint measurement on m copies, instead of a measurement on a single copy only, the probability of making a correct guess for the actual state can be enhanced. However, in order to physically realize a joint measurement of this kind, in many cases an m-photon interaction process would be necessary, while a measurement on a single copy can always be achieved much more simply with the help of linear optics. We also gave a recipe for a linear optical multiport performing the generalized measurement that discriminates with minimum error between N equiprobable symmetric single-photon polarization states and we discussed how the corresponding two-photon-polarization states could be discriminated, at least in principle, with the help of polarization-dependent two-photon absorption [61].

Subset Discrimination and Quantum Filtering in a Two-Dimensional Hilbert Space

While in the previous section we dealt with distinguishing between N individual pure states, we now turn to the error-minimizing discrimination between two subsets of a given set of N pure states. In our work [63] we studied this task for two sets of linearly dependent pure states that collectively span only a two-dimensional Hilbert space. Let us formulate the problem more precisely. We want to devise a measurement that allows us to decide, with the smallest possible error and without inconclusive answers, whether the actual state of the system belongs to the subset of states $\{|\psi_1\rangle, \ldots |\psi_M\rangle\}$, or to the complementary subset of the remaining states $\{|\psi_{M+1}\rangle, \ldots |\psi_N\rangle\}$ with $M < N$. To avoid confusion, in this section we denote the a priori probabilities of the individual pure states by $\eta_j^{\rm i}$, where $j = 1, \ldots, N$. Note that for $M = 1$ our subset-discrimination problem is also called minimum-error *quantum state filtering*, in correspondence to the problem of unambiguous quantum state filtering that has been treated in the respective section of this review. The detection operators Π_1 and Π_2, referring to the two possible measurement outcomes for minimum-error subset-discrimination, are defined in such a way that the quantity $\langle\psi_j|\Pi_1|\psi_j\rangle$ accounts for the probability to infer, from performing the measurement, the state of the system to belong to the first subset, if it has been prepared in the state $|\psi_j\rangle$. Obviously, this inference is correct if $j \leq M$. Similarly, the quantity $\langle\psi_j|\Pi_2|\psi_j\rangle$ is defined as the probability for inferring the state to belong to the second subset. The overall error probability reads

$$P_{\rm err}^{M(N)} = 1 - \left(\sum_{j=1}^{M} \eta_j^{\rm i} \langle\psi_j|\Pi_1|\psi_j\rangle + \sum_{j=M+1}^{N} \eta_j^{\rm i} \langle\psi_j|\Pi_2|\psi_j\rangle \right), \qquad (11.85)$$

where $\sum_{j=1}^{N} \eta_j^{\rm i} = 1$ and $\Pi_1 + \Pi_2 = I$. In general, with respect to the optimum measurement strategies, the problem of subset-discrimination is equivalent to the problem of distinguishing between the two mixed states

$$\rho_1 = \frac{1}{\eta_1} \sum_{j=1}^{M} \eta_j^{\rm i} |\psi_j\rangle\langle\psi_j| \qquad \text{with} \qquad \eta_1 = \sum_{j=1}^{M} \eta_j^{\rm i}, \qquad (11.86)$$

$$\rho_2 = \frac{1}{\eta_2} \sum_{j=M+1}^{N} \eta_j^{\rm i} |\psi_j\rangle\langle\psi_j| \qquad \text{with} \qquad \eta_2 = \sum_{j=M+1}^{N} \eta_j^{\rm i}, \qquad (11.87)$$

provided that the mixed states occur just with the a priori probabilities given by η_1 and η_2, respectively [63]. This equivalence also becomes immediately obvious from comparing (11.48) and (11.85). Thus the minimum error probability for subset-discrimination can be obtained by applying the Helstrom solution, (11.58), to the problem of discriminating between ρ_1 and ρ_2.

In our work we adopted another approach that does not entail any increase in the overall calculation effort, but has the advantage of yielding direct information about the method that realizes the optimum measurement. We first proved that from the restriction to a two-dimensional Hilbert space it follows that the optimum detection operator Π_1 (or Π_2, respectively) can be expressed as the projector onto a particular optimum pure state. By solving the extremal problem that minimizes the expression (11.85) we then determined this optimum pure state, as well as the resulting minimum value $P_E^{M(N)}$. Thus we found that the minimum probability of making an error in distinguishing to which of the two subsets $\{|\psi_1\rangle, \ldots |\psi_M\rangle\}$ or $\{|\psi_{M+1}\rangle, \ldots |\psi_N\rangle\}$ a given quantum state belongs is given by [63]

$$P_E^{M(N)} = \frac{1}{2} - \sqrt{R^2 + |S|^2}, \tag{11.88}$$

where R and S can be expressed as

$$R = \sum_{j=1}^{M} \eta_j^{\mathrm{i}} \left(|\langle\psi_1|\psi_j\rangle|^2 - \frac{1}{2} \right) - \sum_{j=M+1}^{N} \eta_j^{\mathrm{i}} \left(|\langle\psi_1|\psi_j\rangle|^2 - \frac{1}{2} \right), \tag{11.89}$$

$$S = \sum_{j=1}^{M} \eta_j^{\mathrm{i}} \frac{\langle\psi_2|\psi_j\rangle\langle\psi_j|\psi_1\rangle - \langle\psi_2|\psi_1\rangle\,|\langle\psi_1|\psi_j\rangle|^2}{\sqrt{1 - |\langle\psi_1|\psi_2\rangle|^2}} - \sum_{j=M+1}^{N} \eta_j^{\mathrm{i}} \frac{\langle\psi_2|\psi_j\rangle\langle\psi_j|\psi_1\rangle - \langle\psi_2|\psi_1\rangle\,|\langle\psi_1|\psi_j\rangle|^2}{\sqrt{1 - |\langle\psi_1|\psi_2\rangle|^2}}. \tag{11.90}$$

In particular, we applied this result to the case of quantum state filtering for three arbitrary, but linearly dependent states, spanning a two-dimensional Hilbert space. Let us introduce the matrix C by the definition $C_{ij} = \langle\psi_i|\psi_j\rangle$ for $i, j = 1, 2, 3$. Linear dependence then implies $\det(C) = 0$, or

$$|\langle\psi_1|\psi_2\rangle|^2 + |\langle\psi_1|\psi_3\rangle|^2 + |\langle\psi_2|\psi_3\rangle|^2 = 1 + 2\,\mathrm{Re}\langle\psi_1|\psi_2\rangle\langle\psi_1|\psi_3\rangle\langle\psi_1|\psi_3\rangle\ . \tag{11.91}$$

Therefore the minimum error probability obtained in [63] can be written as

$$P_E^{1(3)} = \frac{1}{2} - \sqrt{\frac{1}{4} - \eta_2^{\mathrm{i}}\eta_3^{\mathrm{i}}\ (1 - |\langle\psi_2|\psi_3\rangle|^2) - \eta_1^{\mathrm{i}} \sum_{j=2}^{3} \eta_j^{\mathrm{i}}\,|\langle\psi_1|\psi_j\rangle|^2}. \tag{11.92}$$

For $\eta_3^{\mathrm{i}} = 0$ the expression reduces to the Helstrom formula (11.59) for discriminating two pure states.

As an example, we considered three equally probable symmetric states,

$$|\psi_k\rangle = \cos\theta\,|\gamma_1\rangle + \mathrm{e}^{\mathrm{i}\frac{2\pi}{3}(k-1)}\ \sin\theta\,|\gamma_2\rangle, \qquad k = 1, 2, 3\ , \tag{11.93}$$

where $|\gamma_1\rangle$ and $|\gamma_2\rangle$ denote two orthonormal basis states. For these states also the minimum error probability P_E for distinguishing them individually can be analytically expressed, cf. (11.81). We found that the ratio $P_E^{1(3)}(\theta)/P_E$ varies between 0.56 (for $\theta \approx \pi/12$) and 0.5 (for $\theta = 0$ or $\pi/4$). Hence the minimum-error probability for distinguishing one state from the set of the two others is only about half as large as the minimum-error probability for distinguishing all three states separately. For the special example of the equally probable trine states, given by (11.72), the minimum error probability for quantum state filtering is obtained to be $P_E^{1(3)} = 1/6$. The error-minimizing filtering strategy consists in performing a projection onto the state $|\psi_1\rangle$ and guessing the system to be in this state when the detector clicks, and to be in one of the other states when a projection on the orthogonal state is successful. On the other hand, in the case of the trine states, optimum unambiguous quantum state filtering would yield the minimum failure probability $Q_F^{1(3)} = 1/3$, as can be verified from the formula given in the corresponding chapter of the review. In this case the measurement corresponds to projecting onto the direction orthogonal to $|\psi_1\rangle$, which unambiguously identifies the set of the other states. Obviously the minimized error probability and the optimized failure probability for unambiguous discrimination differ just by the factor one half, in agreement with the limit set by (11.66).

Distinguishing a Pure State from a Uniformly Mixed State in Arbitrary Dimensions

The solution of the minimum-error discrimination problem for two arbitrary quantum states, either pure or mixed, is well known and results in the compact Helstrom formula (11.58) for the minimum error probability, P_E. However, the explicit analytical evaluation of P_E poses severe difficulties when the dimensionality D of the relevant Hilbert space is larger than two. This is due to the fact that applying the Helstrom formula amounts to calculating the eigenvalues of a D-dimensional matrix. In the following we consider a simple yet non-trivial example of an error-minimizing state discrimination problem in an arbitrary dimensional Hilbert space that can be solved analytically, and, in addition, might be related to potential applications. Our problem consists in deciding with minimum error whether a quantum system is prepared either in a given pure state or in a given uniformly mixed state [42], i. e. we have to discriminate between the two quantum states described by

$$\rho_1 = |\psi\rangle\langle\psi|, \qquad \rho_2 = \frac{1}{d}\sum_{j=1}^{d} |u_j\rangle\langle u_j|. \tag{11.94}$$

Here the states $|u_j\rangle$ are supposed to be mutually orthonormal, i. e. $\langle u_i|u_j\rangle = \delta_{ij}$. With D_S denoting the dimensionality of the physical state space of the quantum system, the relation $d \leq D_S$ has to be fulfilled. We note that in

the special case $d = D_S$ the state ρ_2 is the maximally mixed state that describes a completely random state of the quantum system, containing no information at all. Discriminating between the density operators $|\psi\rangle\langle\psi|$ and ρ_2 then amounts to deciding whether the state $|\psi\rangle$ has been reliably prepared, or whether the preparation has totally failed [44].

To simplify the representation, in the following we restrict ourselves to the special case that the a priori probabilities of the two states are $\eta_1 = 1/(d+1)$ for the pure state, and $\eta_2 = d/(d+1)$ for the mixed state, respectively [42]. This means that in the corresponding quantum state filtering scenario all possible pure states would have equal a priori probabilities. In order to calculate the minimum error probability, using the Helstrom formula (11.58), it is necessary to determine the eigenvalues λ of the operator

$$\Lambda = \frac{1}{d+1}\left(\sum_{j=1}^{d} |u_j\rangle\langle u_j| \; - \; |\psi\rangle\langle\psi|\right). \tag{11.95}$$

We found that the eigenvalues are given by [42]

$$\begin{aligned} \lambda_1 &= -\frac{1}{d+1}\sqrt{1-\|\psi^{\|}\|^2}, \\ \lambda_2 &= -\lambda_1, \qquad \lambda_k = \frac{1}{d+1} \qquad (k = 3, \ldots d+1), \end{aligned} \tag{11.96}$$

where we introduced the notation $|\psi^{\|}\rangle$ for the component of $|\psi\rangle$ that lies in the subspace spanned by the states $|u_1\rangle, \ldots, |u_d\rangle$,

$$\|\psi^{\|}\|^2 = \langle\psi^{\|}|\psi^{\|}\rangle = \sum_{j=1}^{d} |\langle u_j|\psi\rangle|^2. \tag{11.97}$$

When the quantum states to be discriminated are linearly independent, i. e. when $\|\psi^{\|}\| \neq 1$, there exists exactly one eigenvalue that is negative, given by λ_1. Therefore according to (11.55) the detection operators for performing the minimum-error measurement are given by $\Pi_1 = |\phi_1\rangle\langle\phi_1|$ and $\Pi_2 = I_{D_S} - \Pi_1$, where $|\phi_1\rangle$ is the eigenstate belonging to the negative eigenvalue, λ_1. On the other hand, when ρ_1 and ρ_2 are linearly dependent, i. e. when $\|\psi^{\|}\| = 1$, a negative eigenvalue does not exist. In this case the optimum measurement strategy is described by the detection operator $\Pi_2 = I_{D_S}$ which means that the resulting minimum error probability, $P_E = 1/(d+1)$, is achievable by guessing the system always to be in the state ρ_2, without performing any measurement at all. The minimum error probability resulting from the above eigenvalues reads [42]

$$P_E = \frac{1}{d+1}\left(1 - \sqrt{1-\|\psi^{\|}\|^2}\right). \tag{11.98}$$

We still mention that the previous considerations can be easily extended to the case that the pure state and the uniformly mixed state given in (11.94) occur with arbitrary a priori-probabilities η_1 and $\eta_2 = 1 - \eta_1$, respectively. The minimum error probability for distinguishing between them is then given by [64]

$$P_E = \frac{1}{2}\left[\eta_1 + \frac{\eta_2}{d} - \sqrt{\left(\eta_1 + \frac{\eta_2}{d}\right)^2 - 4\eta_1\frac{\eta_2}{d}\|\psi^{\|}\|^2}\right]. \tag{11.99}$$

Let us now compare the minimum probability of errors, P_E, with the smallest possible failure probability, Q_F, that can be obtained in a strategy optimized for unambiguously discriminating between the pure state and the uniformly mixed state. The solution of the latter problem coincides with the solution to the problem of optimum unambiguous quantum state filtering. Assuming again that $\eta_1 = 1/(d+1)$, the minimum failure probability is $Q_F = 2\|\psi^{\|}\|/(d+1)$ [42]. Supposing nonorthogonality of the two states, characterized by $0 < \|\psi^{\|}\| \leq 1$, we observe that $P_E/Q_F \leq 1/2$, in accordance with (11.66). When the two states are linearly dependent, i. e. when $\|\psi^{\|}\| = 1$, it follows that $P_E/Q_F = 1/2$. On the other hand, for nearly orthogonal states, where $\|\psi^{\|}\| \ll 1$, we find that $P_E/Q_F \approx \|\psi^{\|}\|/4$. Obviously in this case the minimum error probability is drastically smaller than the optimum failure probability for unambiguous discrimination.

As an application of the minimum-error strategy described above, we discussed the problem of discriminating between a pure and a mixed two-qubit state [42]. An arbitrary bipartite qubit state, shared among two parties A (Alice) and B (Bob), can be expressed with the help of the four orthonormal basis states $|00\rangle, |01\rangle, |10\rangle$ and $|11\rangle$, where $|mn\rangle$ stands for $|m\rangle_A \otimes |n\rangle_B$, with $|0\rangle$ and $|1\rangle$ denoting any two orthonormal basis states of a single qubit. As an interesting special case we considered the problem that Alice and Bob want to decide whether the two-qubit system is either in a given pure state $|\psi\rangle$, occuring with the a priori probability $\eta_1 = 1/4$, or in a uniform mixture of the three symmetric states

$$|u_1\rangle = |00\rangle, \qquad |u_2\rangle = |11\rangle, \qquad |u_3\rangle = \frac{1}{2}(|01\rangle + |10\rangle). \tag{11.100}$$

We found that in this case the minimum error probability is given by

$$P_E = \frac{1}{4}\left(1 - \frac{1}{\sqrt{2}}\,|\langle 01|\psi\rangle - \langle 10|\psi\rangle|\right). \tag{11.101}$$

The same result would hold true if $|u_1\rangle$ and $|u_2\rangle$ were replaced by the two symmetric Bell states $(|00\rangle \pm |11\rangle)/\sqrt{2}$. Minimum-error discrimination is achieved by performing a projection measurement onto the eigenstate $|\phi_1\rangle$ that belongs to the negative eigenvalue λ_1 of the operator Λ, cf. (11.95). In general, this eigenstate will be a superposition of the four two-qubit basis states. The optimum measurement strategy therefore requires a correlated measurement that

has to be carried out collectively on the two qubits. On the other hand, if in our specific example the discrimination would have to performed by local measurements on the qubits only, without any communication between Alice and Bob, the smallest achievable error probability would be always 1/4, independent of the choice of the state $|\psi\rangle$ [42]. We note that the problem we considered is of particular interest in the context of quantum state comparison [30], where one wants to determine whether the states of quantum systems are identical or not. It has been shown [34] that for comparing two unknown single-particle states it is crucial to discriminate the anti-symmetric state of the combined two-particle system from the uniform mixture of the mutually orthogonal symmetric states. In this context it is interesting that recently an application of particle statistics to the problem of minimum-error discrimination has been presented [66]. Discrimination between bipartite states will be discussed in more detail in a separate chapter.

11.4 Discriminating Multiparticle States

So far we have been considering the situation in which the person discriminating between states is in possession of the entire system that is guaranteed to be in one of the allowed states. However, if our system consists of subsystems, these can be distributed among different parties, and then these parties have to determine which state they are sharing by measuring their subsystems and communicating among themselves. This adds another layer of complexity to the problem.

The simplest example is to suppose that we have a two-qubit state, and we give one of the qubits to Alice and the other to Bob. Alice and Bob know that the state is either $|\psi_0\rangle$ or $|\psi_1\rangle$, and by performing local operations and communicating classically (this is abbreviated as LOCC), they want to determine which state they have. We shall consider both the case of minimum-error and unambiguous discrimination. The object is to develop a procedure that Alice and Bob can use to discriminate between the states.

It is possible to immediately obtain some bounds on how successful these procedures can be. If both states are equally likely, and both qubits are measured together, then we know that the states can be successfully unambiguously discriminated with a probability of $P_{IDP} = 1 - |\langle\psi_0|\psi_1\rangle|$ [7]- [9]. This clearly represents an upper bound on what can be accomplished using LOCC. In the case of minimum-error discrimination, the best probability of correctly identifying the state that is obtainable if both states are measured together is

$$P = \frac{1}{2} + \frac{1}{2}\mathrm{Tr}|(\eta_0|\psi_0\rangle\langle\psi_0| - \eta_1|\psi_1\rangle\langle\psi_1|)|, \tag{11.102}$$

where η_j is the a priori probability for $|\psi_j\rangle$, for $j = 1, 2$. This is again an upper bound to what can be achieved using LOCC. A natural question is

whether these bounds are, in fact, achievable. It was recently shown that they are.

In order to see how, let us start with an extreme case; we shall assume that $|\psi_0\rangle$ and $|\psi_1\rangle$ are orthogonal. Walgate, *et al.* proved that in this case the states can be distinguished perfectly using only LOCC [67]. They did, in fact, much more that this, they showed that two orthogonal states, of any dimension, shared by any number of parties can be perfectly distinguished by LOCC. Their proof rests on the fact that any two bipartite states can be expressed in the form

$$|\psi_0\rangle = \sum_{j=0}^{n} |j\rangle_A |\xi_j\rangle_B$$
$$|\psi_1\rangle = \sum_{j=0}^{n} |j\rangle_A |\xi_j^{\perp}\rangle_B, \tag{11.103}$$

where $\{|j\rangle_A | j = 1, \ldots n\}$ is an orthonormal basis for Alice's space and the states $|\xi_j\rangle_B$ and $|\xi_j^{\perp}\rangle_B$, which are not normalized, are orthogonal in Bob's. Alice measures her state in the $|j\rangle_A$ basis and communicates her result to Bob. If her result was $|j_0\rangle$, then Bob measures his particle in order to determine whether it is in the state $|\xi_{j_0}\rangle_B$ or $|\xi_{j_0}^{\perp}\rangle_B$, which he can do perfectly since the states are orthogonal. Note that the measurement that Bob makes depends on the result of Alice's measurement. If the states are split among more than two parties, this procedure can be applied several times. For example, if there are three parties, Alice, Bob and Charlie, then we initially group Bob and Charlie together so that the state can be considered bipartite. Alice performs her measurement and tells Bob and Charlie the result. They now share one of two known, orthogonal states, and they can apply the above procedure again to find out which. The answer will tell them what the original state was.

The case when $|\psi_0\rangle$ and $|\psi_1\rangle$ are not orthogonal (and of arbitrary dimension) was investigated by Virmani, *et al.* [68], and they were able to apply the above decomposition to the problem of minimum-error discrimination. They found a strategy, for arbitrary a priori probabilities, using only LOCC that achieves the optimal success probability given in (11.102). In addition, they found strong numerical evidence that, when the two states are equally likely, unambiguous discrimination is possible with a probability of P_{IDP} using LOCC, and they found a class of states for which they could prove that this was true. A proof that this is true for all bipartite states was provided by Chen and Yang [69].

The procedure that makes LOCC unambiguous discrimination with a success probability of P_{IDP} possible is closely related to the one for discriminating orthogonal states. Alice makes a projective measurement on her particle that gives her no information about whether the state is $|\psi_0\rangle$ or $|\psi_1\rangle$, and she then communicates her result to Bob. Based on what Alice has told him, Bob chooses a measurement to make on his particle. In particular, he applies the

procedure for the optimal unambiguous discrimination of single qubit states to his particle. However, in this procedure one must know the two states that one is discriminating between, and it is this information that is provided by the result of Alice's measurement.

Together with Mimih we took a somewhat different approach to unambiguous discrimination of two-qubit states [70]. The motivation was to study bipartite state discrimination schemes that could be used in quantum communication protocols, quantum secret sharing, in particular [71]- [74]. A quantum cryptography protocol based on two-state unambiguous discrimination already existed [38], and this suggested that the discrimination of bipartite states might find application as well. In secret sharing, Alice and Bob are both sent information that allows them to decode a message if they act together, but neither party can decode it by themselves. The schemes discussed above are not well suited for this type of application, because the information gained by the two parties is not the same. In particular, after the measurements have been made (note that Alice must communicate her result to Bob so that he can make his), Alice knows nothing and Bob knows which state has been sent. We were interested in schemes that are more symmetric.

Our approach was to examine situations in which the classical communication between the parties was limited. One possibility is to allow no classical communication. In that case each party has three possible measurement results, 0 corresponding to $|\psi_0\rangle$, 1 corresponding to $|\psi_1\rangle$, and f for failure to distinguish. If $|\psi_0\rangle$ is sent, then Alice and Bob both measure 0 or both measure f, so that they both know, without communicating, that $|\psi_0\rangle$ was sent or that the measurement failed. If $|\psi_1\rangle$ is sent, then they both measure either 1 or f. In the case of qubits, we found that the best that can be done is to identify one of the states and fail the rest of the time, i.e. we never get a positive identification for the second state. The situation improves if we go to qutrits. In that case there are examples of states that can be distinguished with the success probability equal to P_{IDP}. One is given by the two states (the states $|0\rangle$, $|1\rangle$, and $|2\rangle$ are an orthonormal basis for the qutrits)

$$|\psi_0\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |22\rangle)$$
$$|\psi_1\rangle = \frac{1}{\sqrt{2}}(|11\rangle + |22\rangle). \tag{11.104}$$

If Alice and Bob each measure 0, they know they were sent $|\psi_0\rangle$, if they each measure 1, they know they were sent $|\psi_1\rangle$, and if they both measure 2 they failed.

We also considered the situation in which and Alice and Bob make measurements, and later pool their results to determine which state was sent. However, conditional measurements were banned, i.e. situations in which the measurement made by one party depends on the measurement result of the other were not allowed. Conditional measurements seem to have the problem that, as has been noted, the information the parties receive is not the same,

and, that if there is a delay between the time the qubits are received and the time they are measured, then the qubits must be protected against decoherence until the measurement is made. In the procedures we studied, we found that it was optimal for each party to make projective measurements. If the two states shared the same Schmidt basis, then these procedures could successfully distinguish them with a probability of P_{IDP}, but if they did not, the probability of success was, in general, smaller than P_{IDP}.

The discrimination of more than two states shared by two parties has begun to be investigated, and presently results only exist for the case in which all of the states are orthogonal to each other. Ghosh, *et al.* have shown that it is not possible, in general, to deterministically distinguish either three or four orthogonal two-qubit states using only local operations and classical communication [75]. A general condition on when orthogonal, bipartite $2 \times d$ states (one of the particles is a qubit and the other a qudit), can be distinguished by LOCC was found recently by Walgate and Hardy [76]. In the case of 2×2 states, they found that for three of them to be perfectly distinguishable by LOCC, at least two of them must be product states, and for four, all of them must be product states. As far as we are aware, the problem of distinguishing three or more nonorthogonal states shared between two or more parties using LOCC has not yet been studied although, very recently, Chefles [77] derived a necessary and sufficient condition for a finite set of states in a finite dimensional multiparticle quantum system for LOCC unambiguous discrimination. This suggests that there is much still to be learned about distinguishing multipartite states using local operations and classical communication and, in general, separable quantum operations have to be considered.

11.5 Outlook

Quantum state discrimination is a very rapidly evolving field just like many other areas of quantum information and quantum computing. We have reviewed here the two most important - and simplest - state discrimination strategies, unambiguous discrimination and minimum-error discrimination. The minimum-error strategy for the discrimination of two mixed states has been one of the first problems that was solved exactly. With the recent progress toward the unambiguous discrimination of mixed states we expect that in the near future the problem of unambiguous discrimination of two mixed states will be solved completely. Then attention will quite naturally turn toward the discrimination of more than two states where so far only special cases were solved completely and partial results were obtained in the general case. A rapidly emerging field with a lot of room for quick progress is the discrimination of multiparticle states using LOCC only. Applications in the area of quantum cryptography and probabilistic quantum algorithms will surely follow but in a field with such a rapidly changing landscape it would not be responsible to predict more than the immediately foreseeable future.

References

1. M. A. Nielsen and I. L. Chuang: *Quantum Computation and Information* (Cambridge University Press, 2000)
2. A. Peres: *Quantum Theory: Concepts and Methods* (Kluwer, Dordrecht, 1995)
3. C. W. Helstrom: *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976)
4. A. Chefles: Contemp. Phys. **41**, 401 (2000); arXiv:quant-ph/0010114 (2000).
5. M. Sasaki, S. M. Barnett, R. Jozsa, M. Osaki and O. Hirota: Phys. Rev. A **59**, 3325 (1999)
6. S. M. Barnett, C. R. Gilson, and M. Sasaki, J. Phys. A: Math. Gen. **34**, 6755 (2001); arXiv:quant-ph/0107024 (2001)
7. I. D. Ivanovic: Phys. Lett. A **123**, 257 (1987)
8. D. Dieks: Phys. Lett. A **126**, 303 (1988)
9. A. Peres: Phys. Lett. A **128**, 19 (1988)
10. G. Jaeger and A. Shimony: Phys. Lett. A **197**, 83 (1995)
11. M. Ban: Phys. Lett. A **213**, 235 (1996)
12. K. Kraus: *States, Effects and Operations: Fundamental Notions of Quantum Theory* (Springer, Berlin, 1983)
13. M. A. Neumark: Izv. Akad. SSSR Ser. Mat. **4**, (1940)
14. J. Bergou, M. Hillery, and Y. Sun: J. Mod. Opt. **47**, 487 (2000)
15. B. Huttner, A. Muller, J. D. Gautier, H. Zbinden, and N. Gisin: Phys. Rev A **54**, 3783 (1996)
16. R. B. M. Clarke, A. Chefles, S. M. Barnett, and E. Riis: Phys. Rev. A **63**, 040305(R) (2001)
17. A. Chefles: Phys. Lett. A **239**, 339 (1998)
18. S. Zhang Y. Feng, X. Sun, and M. Ying: Phys. Rev. A **64**, 062103 (2001)
19. L. M. Duan and G. C. Guo: Phys. Rev. Lett. **80**, 4999 (1998)
20. X. Sun, S. Zhang, Y. Feng, and M. Ying: Phys. Rev. A **65**, 044306 (2002)
21. A. Peres and D. R. Terno: J. Phys. A **31**, 7105 (1998)
22. Y. Sun, M. Hillery, and J. A. Bergou: Phys. Rev. A **64**, 022311 (2001)
23. M. Mohseni, A. Steinberg, and J. A. Bergou: arXiv:quant-ph/0401002
24. A. Chefles and S. Barnett: Phys. Lett. A **250**, 223 (1998)
25. J. Fiurášek and M. Ježek: Phys. Rev. A 67, 012321 (2003); arXiv:quant-ph/0208126 (2002)
26. Y. Sun, J. A. Bergou, and M. Hillery: Phys. Rev. A **66** 032315 (2002)
27. J. A. Bergou, U. Herzog, and M. Hillery: Phys. Rev. Lett. **90**, 257901 (2003)
28. S. Zhang and M. Ying: Phys. Rev. A **65**, 062322 (2002)
29. M. Takeoka, M. Ban, and M. Sasaki: Phys. Rev. A **68**, 012307 (2003)
30. S. M. Barnett, A. Chefles, and I. Jex: Phys. Lett. A **307**, 189 (2003)
31. M. Ježek: Phys. Lett. A **299**, 441 (2002).
32. Y. Sun and J. Bergou: in preparation
33. J. Bergou, U. Herzog, and M. Hillery: in preparation
34. I. Jex, E. Andersson, and A. Chefles: J. Mod. Opt. **51**, 504 (2004); arXiv:quant-ph/0305120 (2003)
35. T. Rudolph, R. W. Spekkens, and P. S. Turner: Phys. Rev. A **68**, 010301(R) (2003)
36. Ph. Raynal, N. Lütkenhaus, and S. van Enk: Phys. Rev. A **68**, 022308 (2003)
37. Y. Eldar: Phys. Rev. A **67**, 042309 (2003)

38. C. H. Bennett: Phys. Rev. Lett. **68**, 3121 (1992)
39. D. Deutsch and R. Jozsa: Proc. R. Soc. London A **439**, 553 (1992)
40. J. A. Bergou, U. Herzog, and M. Hillery: in preparation
41. C. A. Fuchs: PhD thesis, Univ. of New Mexico (1995); arXiv:quant-ph/9601020 (1996)
42. U. Herzog: J. Opt. B **6**, S24 (2004); arXiv:quant-ph/0307038 (2003)
43. S. M. Barnett and E. Rijs: J. Mod. Opt. **44**, 1061 (1997)
44. K. Hunter: Phys. Rev. A **68**, 012306 (2003); arXiv:quant-ph/0211148 (2002)
45. A. S. Holevo: J. Multivar. Anal. **3**, 337 (1973)
46. H. P. Yuen, R. S. Kennedy, and M. Lax: IEEE Trans. Inform. Theory **IT-21**, 125 (1975)
47. M. Ban, K. Kurokawa, R. Momose, and O. Hirota: Int. J. Theor. Phys. **55**, 22 (1997)
48. Y. C. Eldar, A. Megretski, and G. C. Verghese: IEEE Trans. Inform. Theory **IT-49**, 1007 (2003); arXiv:quant-ph/0211111 (2002)
49. C.-L. Chou and L. Y. Hsu: Phys. Rev. A **68**, 042305 (2003); arXiv:quant-ph/03044117 (2003)
50. R. S. Kennedy, M. I. T. Res. Lab: Electron. Quart. Progr. Rep. **110**, 142 (1973)
51. S. M. Barnett: Phys. Rev. A **64**, 030303(R) (2001)
52. E. Andersson, S. M. Barnett, C. B. Gilson, and K. Hunter: Phys. Rev. A **65**, 052308 (2002); arXiv:quant-ph/0201074 (2002)
53. M. Ježek, J. Řeháček, and J. Fiurášek: Phys. Rev. A **65**, 060301 (2002); arXiv:quant-ph/0201109 (2002)
54. Y. Eldar: Phys. Rev. A **68**, 052303 (2003); arXiv:quant-ph/0304077 (2003)
55. R. B. M. Clarke, V. M. Kendon, A. Chefles, S. M. Barnett, E. Riis, and M. Sasaki: Phys. Rev. A **64**, 012303 (2001)
56. S. Franke-Arnold, E. Andersson, S. M. Barnett, and S. Stenholm: Phys. Rev. A **63**, 052301 (2001)
57. J. Preskill. In *Lecture Notes for Physics 229: Quantum Information and Computation* (Cambridge University Press, 1998)
58. R. Jozsa, M. Koashi, N. Linden, S. Popescu, S. Presnell, D. Shepherd, and A. Winter: arXiv:quant-ph/0303167 (2003)
59. A. Chefles and S. M. Barnett: J. Mod. Opt. **45**, 1295 (1998)
60. C.-W. Zhang, C.-F. Li, and G.-C. Guo: Phys. Lett. A, **261**, 25 (1999)
61. U. Herzog: Fortschr. Phys. **49**, 981 (2001); arXiv:quant-ph/0105139 (2001)
62. A. Chefles: Phys. Rev. A **66**, 042325 (2002)
63. U. Herzog and J. A. Bergou: Phys. Rev. A **65**, 050305(R) (2002)
64. U. Herzog and J. A. Bergou: arXiv:quant-ph/0403124
65. C. H. Bennett: Phys. Rev. Lett. **68**, 3121 (1992)
66. S. Bose, A. Ekert, Y. Omar, N. Paunković and V. Vedral: Phys. Rev. A **68**, 052309 (2003); arXiv:quant-ph/0309090 (2003)
67. J. Walgate, A. Short, L. Hardy, and V. Vedral: Phys. Rev. Lett. **85**, 4972 (2000)
68. S. Virmani, M. F. Sacchi, M. B. Plenio, and D. Markham: Phys. Lett. A **288**, 62 (2001)
69. Yi-Xin Chen and Dong Yang: Phys. Rev. A **65**, 022320 (2002)
70. M. Hillery and J. Mimih: Phys. Rev. A **67**, 042304 (2003)
71. M. Hillery, V. Bužek, and A. Berthiaume: Phys. Rev. A **59**, 1829 (1999)
72. A. Karlsson, M. Koashi, and N. Imoto: Phys. Rev. A **59**, 162 (1999)
73. W. Tittel, H. Zbinden, and N. Gisin: Phys. Rev. A **63**, 042301 (2001)

74. R. Cleve, D. Gottesman, and H. -K. Lo: Phys. Rev. Lett. **83**, 648 (1999)
75. S. Ghosh, G. Kar, A. Roy, D. Sarkar, A. Sen(De), and U. Sen: Phys. Rev. A **65**, 062307 (2002)
76. J. Walgate and L. Hardy: Phys. Rev. Lett. **89**, 147901 (2002)
77. A. Chefles: arXiv:quant-ph/0302066 (2003)