Quantum state discrimination

Stephen M. Barnett[1,*] and Sarah Croke[2]

[1]Department of Physics, University of Strathclyde, Glasgow G4 0NG, UK

[2]Perimeter Institute for Theoretical Physics, Waterloo, Ontario, N2L 2Y5, Canada

*Corresponding author: steve@phys.strath.ac.uk

Received October 10, 2008; revised December 10, 2008; accepted December 11, 2008; posted December 11, 2008 (Doc. ID 102658); published February 11, 2009

It is a fundamental consequence of the superposition principle for quantum states that there must exist nonorthogonal states, that is, states that, although different, have a nonzero overlap. This finite overlap means that there is no way of determining with certainty in which of two such states a given physical system has been prepared. We review the various strategies that have been devised to discriminate optimally between nonorthogonal states and some of the optical experiments that have been performed to realize these. © 2009 Optical Society of America

OCIS codes: 270.5565, 270.5585.

1. Introduction. 240
2. Generalized Measurements. 241
3. State Discrimination—Theory. 243
   3.1. Minimum Error Discrimination. 243
      3.1a. Minimum Error Conditions. 244
      3.1b. Square-Root Measurement. 246
      3.1c. Other Results. 248
   3.2. Unambiguous Discrimination. 249
      3.2a. $N>2$ Pure States. 251
      3.2b. Mixed States. 251
   3.3. Maximum Confidence Measurements. 252
      3.3a. Other Similar Strategies. 256
      3.3b. Related Problems—Quantum State Filtration. 256
   3.4. Comments on the Relationships between Strategies. 257
   3.5. Mutual Information. 258
   3.6. No-Signaling Bounds on State Discrimination. 259
      3.6a. Unambiguous Discrimination. 260
      3.6b. Maximum Confidence Measurements. 261
4. State Discrimination—Experiments. 261
   4.1. Photon Polarization. 262
   4.2. Minimum Error Discrimination. 264
      4.2a. Two States. 264
      4.2b. Three or Four States. 264

1943-8206/09/020238-41/$15.00 © OSA

4.3. Unambiguous Discrimination. 268
4.4. Maximum Confidence Measurements. 269
4.5. Mutual Information. 270
5. Conclusion. 272
Acknowledgments. 272
References and Notes. 273

Quantum state discrimination

Stephen M. Barnett and Sarah Croke

1. Introduction

The state of a quantum system is a mysterious object and has been the subject of much attention since the earliest days of quantum theory. We know that it provides a way of calculating the observed statistical properties of any desired observable but that it is not, itself, observable. This means that we cannot determine by observation the state of any single physical system. If we have some prior information, however, then we may be able to use this to determine, at least to some extent, the state. Consider, for example, a single photon that we know has been prepared with either horizontal or vertical polarization. A suitably oriented polarizing beam splitter can be used to transmit the photon if it is vertically polarized and reflect it if its polarization is horizontal. Determining the path of the photon by absorbing it with a suitable detector then determines the state to have been one of horizontal or vertical polarization.

Suppose, however, that we are told that our photon was prepared with either horizontal or with left-circular polarization. These quantum states of polarization are not orthogonal in that states of circular polarization are superpositions of those of both vertical and horizontal polarization:

$$|L\rangle = \frac{1}{\sqrt{2}}(|H\rangle + i|V\rangle) \Rightarrow \langle H|L\rangle = \frac{1}{\sqrt{2}} \neq 0. \tag{1}$$

If we subject our photon to the polarization measurement outlined in the preceding paragraph, then a left-circularly polarized photon will appear to be horizontally polarized with probability $1/2$ and vertically polarized with the same probability.

The problem of discriminating between such states is fundamental to the quantum theory of communications [1–4] and underlies the secrecy of the now well-reviewed science of quantum cryptography [5–9]. Indeed, we can use the connection between quantum state discrimination and quantum communications to motivate the problem of state discrimination. We suppose that two parties, conventionally named Alice and Bob, wish to communicate by using a quantum channel. To do this Alice selects from a given set of states, $|\psi_i\rangle$ (or more generally mixed states with density operators $\hat{\rho}_i$) with a given set of probabilities p_i. The selected state is encoded in the preparation of a given physical system, such as photon polarization, and this is sent to Bob. Bob will know both the set of possible states and the associated preparation probabilities. His task is to determine, by means of a suitable measurement, the state selected by Alice and hence the intended message. This, then is the quantum state discrimination problem: how can we best discriminate among

a known set of possible states $|\psi_i\rangle$, each having been prepared with a known probability p_i.

The quantum state discrimination problem, as posed here, has been the subject of active theoretical investigation for a long time [1–3,10–14], but it is only comparatively recently that experiments have been performed, and most of these have been based on optics. There exist in the literature a number of reviews of and introductions to quantum state discrimination [4,15–21]. Our purpose in preparing this review is twofold: first, to bring the rapidly developing field up to date and, second, to introduce the idea of state discrimination to a wider audience in optics. It seems especially appropriate to do this as it is in simple optical experiments that the ideas are most transparent and where most of the important practical developments have been made.

2. Generalized Measurements

Most of us are introduced to the idea of measurements in quantum theory in a manner that is, essentially, that formulated by von Neumann [22]. Each observable property O is associated with a Hermitian operator $\hat{O}$ (or more precisely a self-adjoint one), the eigenvalues of which are the possible results of a measurement of O. If the eigenvalues and eigenvectors are o_m and $|o_m\rangle$, then we can write the operator $\hat{O}$ in the diagonal form

$$\hat{O} = \sum_m o_m |o_m\rangle\langle o_m|. \tag{2}$$

If the system to be measured has been prepared in the state $|\psi\rangle$, then the probability that a measurement of O will give the result o_m is

$$P(o_m) = |\langle o_m|\psi\rangle|^2. \tag{3}$$

Consider, for example, a measurement to determine whether the polarization of a single photon is horizontal or vertical. A suitable operator, corresponding to this measurement, would be

$$\hat{\mathrm{Pol}} = H|H\rangle\langle H| + V|V\rangle\langle V|. \tag{4}$$

The probability that a measurement of this property on a photon prepared in the circularly polarized state $|L\rangle$ will give the result H, corresponding to horizontal polarization, is

$$P(H) = |\langle H|L\rangle|^2 = \frac{1}{2}. \tag{5}$$

It is helpful, in what follows, to rewrite the above probabilities as the expectation value of an operator. In this way the probability that a measurement of optical polarization shows the photon to be horizontally polarized is

$$P(H) = \langle |H\rangle\langle H|\rangle = \langle \hat{P}_H\rangle, \tag{6}$$

where $\hat{P}_H = |H\rangle\langle H|$, the projector onto the state $|H\rangle$, is the required operator. More generally, for our operator $\hat{O}$, the probability that a measurement gives the value o_m is

$$P(o_m) = \langle |o_m\rangle\langle o_m|\rangle = \langle \hat{P}_m \rangle. \tag{7}$$

We note that the measured value, itself, makes no explicit appearance in this probability; it is not the eigenvalue but only the corresponding eigenvector that determines the form of the projector and hence the probability for the associated measurement outcome.

The projectors have four important mathematical properties, and it is helpful to list these:

- The projectors are Hermitian operators, $\hat{P}_m^\dagger = \hat{P}_m$. This property is associated with the fact that probabilities are, themselves, observable quantities.
- They are positive operators, which means that $\langle\psi|\hat{P}_m|\psi\rangle \geq 0$ for all possible states $|\psi\rangle$. This reflects the fact that the expectation value of the projector is a probability and must, therefore, be positive or zero.
- They are complete in that $\Sigma_m \hat{P}_m = \hat{\mathbb{1}}$, so that the sum of the probabilities for all possible measurement outcomes is unity.
- They are orthonormal in that $\hat{P}_m \hat{P}_n = 0$ unless $m = n$. This property is sometimes associated with the fact that measurement outcomes must be distinct (you can only get one of them). This view is, as we shall see, not correct. You can indeed only get one outcome, but this does not require the orthonormality property.

The theory of generalized measurements can be formulated simply by dropping the final requirement. To see how this works, we introduce a set of probability operators $\{\hat{\pi}_m\}$, each of which we wish to associate with a measurement outcome such that the probability that our measurement gives the result labeled m is

$$P(m) = \langle \hat{\pi}_m \rangle. \tag{8}$$

We insist on the first three of the properties of the projectors, as these are required if we are to maintain the probability interpretation (8), but we drop the final requirement so that our probability operators have the following properties:

- The probability operators are Hermitian: $\hat{\pi}_m^\dagger = \hat{\pi}_m$.
- They are positive operators: $\langle\psi|\hat{\pi}_m|\psi\rangle \geq 0$ for all possible states $|\psi\rangle$.
- They are complete: $\Sigma_m \hat{\pi}_m = \hat{\mathbb{1}}$.

Note that Hermiticity follows from positivity, since if all expectation values are nonnegative, $\langle\psi|\hat{\pi}_m|\psi\rangle \geq 0$ for all possible states $|\psi\rangle$, then in particular they are all real numbers and $\hat{\pi}_m$ must be Hermitian. The set of probability operators characterizing the possible outcomes of any generalized measurement is called a probability operator measure, usually abbreviated to POM [1,4]. You will often find this set referred to as a positive operator-valued measure or POVM [23,24]. If the latter name is used then the probability operators become elements of a positive operator-valued measure.

The differences between the projectors and more general probability operators are best appreciated by reference to some simple examples, and these will be given in the following sections. There are, however, some important and

perhaps even surprising points, and it is sensible to emphasize these here. First, the three properties described above have a remarkable generality in that (i) any measurement can be described by the appropriate set of probability operators and (ii) any set of operators that satisfy the three properties correspond to a possible measurement [4,23]. This means that we can seek the optimum measurement in any given situation mathematically, by searching among all sets of operators that satisfy the required properties. Having found this optimum measurement, we know that a physical realization of it will exist, and we can seek a way to implement it in the laboratory. The second point to emphasize is that the number of (orthogonal) projectors can only be less than or equal to the dimension of the state space. For optical polarization, for example, there are only two orthogonal polarizations, and the state space is therefore two dimensional. It follows that any von Neumann measurement of polarization can have only two outcomes. By dropping the requirement for orthogonality, we allow a generalized measurement to have any number of outcomes, so a generalized measurement of polarization can have three, four, or more different outcomes. Finally, a generalized measurement allows us to describe the simultaneous observation of incompatible observables, such as position and momentum or, in the context of quantum optics, orthogonal field quadratures [25,26]. Perhaps the first reported generalized optical measurement was of precisely this form [27,28].

3. State Discrimination—Theory

3.1. Minimum Error Discrimination

In quantum state discrimination we wish to design a measurement to distinguish optimally between a given set of states. As we have seen in Section 2, any physically realizable measurement can be described by a POM. Thus, by mathematically formulating a figure of merit describing the performance of a measurement, we can search for the set of probability operators describing the optimal measurement. There are several possible figures of merit, each one corresponding to a different strategy. Possibly the simplest criteria that may be applied is to minimize the probability of making an error in identifying the state. We begin with the special case where the state is known to be one of two possible pure states, $|\psi_0\rangle$, $|\psi_1\rangle$, with associated probabilities p_0, $p_1 = 1 - p_0$. If outcome 0 (associated with the probability operator $\hat{\pi}_0$) is taken to indicate that the state was $|\psi_0\rangle$, and outcome 1 (associated with $\hat{\pi}_1$) is taken to indicate that the state was $|\psi_1\rangle$, the probability of making an error in determining the state is given by

$$\begin{aligned} P_{\mathrm{err}} &= P(\psi_0)P(1|\psi_0) + P(\psi_1)P(0|\psi_1) \\ &= p_0\langle\psi_0|\hat{\pi}_1|\psi_0\rangle + p_1\langle\psi_1|\hat{\pi}_0|\psi_1\rangle \\ &= p_0 - \mathrm{Tr}((p_0|\psi_0\rangle\langle\psi_0| - p_1|\psi_1\rangle\langle\psi_1|)\hat{\pi}_0), \end{aligned} \tag{9}$$

where in the last line we have used the completeness condition $\hat{\pi}_0 + \hat{\pi}_1 = \hat{\mathbb{1}}$. This expression takes its minimum value when the second term reaches a maximum, which in turn is achieved if $\hat{\pi}_0$ is a projector onto the positive eigenket of the operator $p_0|\psi_0\rangle\langle\psi_0| - p_1|\psi_1\rangle\langle\psi_1|$. Note that two pure states define a two-dimensional space, and without loss of generality we can choose an orthogonal basis $\{|0\rangle, |1\rangle\}$ of this space such that the components of each state

in this basis are real. Thus we can express $|\psi_0\rangle$, $|\psi_1\rangle$ as follows:

$$|\psi_0\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle,$$

$$|\psi_1\rangle = \cos\theta|0\rangle - \sin\theta|1\rangle, \tag{10}$$

and the eigenvalues of $p_0|\psi_0\rangle\langle\psi_0| - p_1|\psi_1\rangle\langle\psi_1|$ can be calculated directly as

$$\lambda_\pm = \frac{1}{2}(p_0 - p_1 \pm \sqrt{1 - 4p_0p_1\cos^2 2\theta}). \tag{11}$$

The minimum probability of making an error is then given by the so-called *Helstrom bound* [1],

$$P_{\text{err}} = \frac{1}{2}(1 - \sqrt{1 - 4p_0p_1|\langle\psi_0|\psi_1\rangle|^2}), \tag{12}$$

and the optimal measurement is simply a von Neumann measurement. In particular, for $p_0 = p_1 = 1/2$, the optimal measurement is a projective measurement onto the states

$$|\varphi_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle),$$

$$|\varphi_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \tag{13}$$

These are symmetrically located about the signal states, as may be expected from the symmetry of the problem. As p_0 is increased, $|\varphi_0\rangle$ moves closer to $|\psi_0\rangle$, and the optimal measurement becomes biased toward making fewer errors when the more probable state is sent (see Fig. 1). Finally, as may be expected intuitively, if p_0 is much bigger than p_1, the optimal measurement is very close to simply asking, "Is the state $|\psi_0\rangle$ or not?"

3.1a. Minimum Error Conditions

The above analysis is easily extended to two mixed states $\hat{\rho}_0$, $\hat{\rho}_1$, in which case the optimal measurement becomes a projective measurement onto the subspaces corresponding to positive and negative eigenvalues of $p_0\hat{\rho}_0 - p_1\hat{\rho}_1$. In the general case of N possible states $\{\hat{\rho}_i\}$ with associated *a priori* probabilities $\{p_i\}$, the aim is to minimize the expression

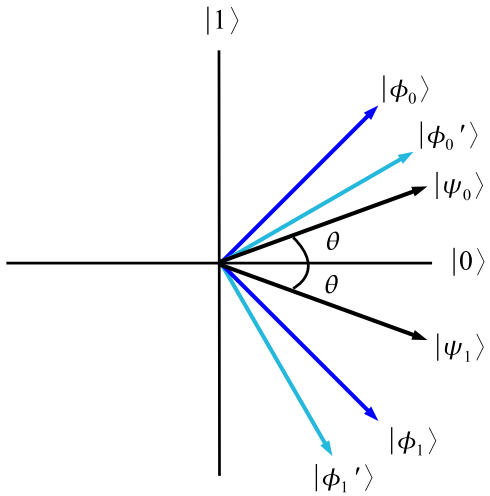
$$P_{\text{err}} = \sum_{i=0}^{N-1} p_i \sum_{j\neq i} \text{Tr}(\hat{\rho}_i\hat{\pi}_j) \tag{14}$$

or, equivalently, to maximize

$$P_{\text{corr}} = 1 - P_{\text{err}} = \sum_{i=0}^{N-1} p_i \text{Tr}(\hat{\rho}_i\hat{\pi}_i). \tag{15}$$

The optimal measurement is known only in certain special cases; however, necessary and sufficient conditions that must be satisfied by the optimal POM for the general case are known [1,12,13] and are given in Eqs. (16) and

Figure 1



The optimal minimum error measurement for discriminating between the pure states $|\psi_0\rangle$, $|\psi_1\rangle$ is a von Neumann measurement. For $p_0=p_1=1/2$ this is a projective measurement onto the states $|\varphi_0\rangle$, $|\varphi_1\rangle$, symmetrically located on either side of the signal states and shown in blue here. For $p_0>p_1$ the optimal measurement performs better when the state $|\psi_0\rangle$ sent, shown here in light blue (labeled $|\varphi_0'\rangle$, $|\varphi_1'\rangle$), is the case $p_0=3/4$.

(17). For simplicity, we prove only sufficiency of the conditions here, but we note that there is also a straightforward proof of their necessity [29].

Necessary and sufficient conditions that must be satisfied by the POM achieving minimum error in distinguishing between the states $\{\hat{\rho}_i\}$, occurring with probabilities $\{p_i\}$, are given by

$$\sum_i p_i\hat{\rho}_i\hat{\pi}_i - p_j\hat{\rho}_j \geq 0, \quad \forall\, j, \tag{16}$$

$$\hat{\pi}_i(p_i\hat{\rho}_i - p_j\hat{\rho}_j)\hat{\pi}_j = 0, \quad \forall\, i,j. \tag{17}$$

Note that these conditions are not independent; the second may be derived from the first, as shown in the text.

If $\{\hat{\pi}_i\}$ corresponds to an optimal measurement, then for all other POMs $\{\hat{\pi}_i'\}$ we require

$$\sum_i p_i \mathrm{Tr}(\hat{\rho}_i\hat{\pi}_i) \geq \sum_j p_j \mathrm{Tr}(\hat{\rho}_j\hat{\pi}_j'). \tag{18}$$

Inserting the identity $\Sigma_j\hat{\pi}_j'=\hat{\mathbb{1}}$ gives

$$\sum_j \mathrm{Tr}\left(\left(\sum_i p_i\hat{\rho}_i\hat{\pi}_i - p_j\hat{\rho}_j\right)\hat{\pi}'_j\right) \geq 0. \tag{19}$$

Note that $\hat{\pi}'_j \geq 0$; thus the above holds if Eq. (16) holds, which is therefore a sufficient condition.

For any POM satisfying this condition, it follows that the operator $\hat{\Gamma} = \Sigma_i p_i\hat{\rho}_i\hat{\pi}_i$ is positive, and therefore Hermitian. Thus we have

$$\sum_j \left(\sum_i p_i\hat{\pi}_i\hat{\rho}_i - p_j\hat{\rho}_j\right)\hat{\pi}_j = \sum_i \hat{\pi}_i\left(p_i\hat{\rho}_i - \sum_j p_j\hat{\rho}_j\hat{\pi}_j\right) = 0, \tag{20}$$

where we have used the fact that the probability operators form a resolution of the identity $\Sigma_i\hat{\pi}_i = \hat{\mathbb{1}}$. As both $\Sigma_i p_i\hat{\pi}_i\hat{\rho}_i - p_j\hat{\rho}_j$ and $\hat{\pi}_j$ are positive operators, each term in the sum over j must be identically zero. Using similar reasoning we can argue that each term in the sum over i must be identically zero. Thus, in terms of $\hat{\Gamma}$ we obtain

$$(\hat{\Gamma} - p_j\hat{\rho}_j)\hat{\pi}_j = \hat{\pi}_i(p_i\hat{\rho}_i - \hat{\Gamma}) = 0, \quad \forall\, i,j. \tag{21}$$

Eliminating $\hat{\Gamma}$ gives Eq. (17), which is therefore also a sufficient condition.

3.1b. Square-Root Measurement

For any given set of states we can construct an associated measurement, the square-root measurement [30–33], as follows:

$$\hat{\pi}_i = p_i\hat{\rho}^{-1/2}\hat{\rho}_i\hat{\rho}^{-1/2}, \tag{22}$$

where $\hat{\rho} = \Sigma_i p_i\hat{\rho}_i$. It is clear that the probability operators $\{\hat{\pi}_i\}$ are positive and sum to the identity, and thus form a complete measurement. For many of the cases in which the optimal minimum error measurement is known, it is the square-root measurement [34–39]. We will present here the example of N symmetric pure states occurring with equal *a priori* probabilities $p_i = 1/N$, considered by Ban *et al.* [34], and given by

$$|\psi_i\rangle = \hat{V}|\psi_{i-1}\rangle = \hat{V}^i|\psi_0\rangle, \quad i = 0, \ldots, N-1, \tag{23}$$

for some unitary operator $\hat{V}$ satisfying $\hat{V}^N = \hat{\mathbb{1}}$. For this set we have

$$\hat{\rho} = \frac{1}{N}\sum_{i=0}^{N-1} |\psi_i\rangle\langle\psi_i| = \frac{1}{N}\sum_{i=0}^{N-1} \hat{V}^i|\psi_0\rangle\langle\psi_0|\hat{V}^{\dagger i}, \tag{24}$$

and it is useful to note that

$$\begin{aligned}
\hat{V}\hat{\rho}\hat{V}^\dagger &= \frac{1}{N}\sum_{i=0}^{N-1} \hat{V}|\psi_i\rangle\langle\psi_i|\hat{V}^\dagger \\
&= \frac{1}{N}\sum_{i=0}^{N-1} \hat{V}^{i+1}|\psi_0\rangle\langle\psi_0|\hat{V}^{\dagger i+1} \\
&= \frac{1}{N}\left(\sum_{i=1}^{N-1} \hat{V}^i|\psi_0\rangle\langle\psi_0|\hat{V}^{\dagger i} + \hat{V}^N|\psi_0\rangle\langle\psi_0|\hat{V}^{\dagger N}\right) \\
&= \hat{\rho},
\end{aligned} \tag{25}$$

where in the last line we have used the property $\hat{V}^N = \hat{\mathbb{1}}$. Thus

$$\hat{V}\hat{\rho} = \hat{V}\hat{\rho}\hat{V}^\dagger\hat{V} = \hat{\rho}\hat{V}, \tag{26}$$

and $\hat{\rho}$ commutes with $\hat{V}$. The square-root measurement consists of the operators

$$\hat{\pi}_i = \frac{1}{N}\hat{\rho}^{-1/2}|\psi_i\rangle\langle\psi_i|\hat{\rho}^{-1/2} = \frac{1}{N}\hat{\rho}^{-1/2}\hat{V}^i|\psi_0\rangle\langle\psi_0|\hat{V}^{\dagger i}\hat{\rho}^{-1/2}, \tag{27}$$

and condition (17) is equivalent to the requirement

$$\langle\psi_i|\hat{\rho}^{-1/2}|\psi_i\rangle\langle\psi_i|\hat{\rho}^{-1/2}|\psi_j\rangle - \langle\psi_i|\hat{\rho}^{-1/2}|\psi_j\rangle\langle\psi_j|\hat{\rho}^{-1/2}|\psi_j\rangle = 0. \tag{28}$$

Noting that

$$\langle\psi_i|\hat{\rho}^{-1/2}|\psi_i\rangle = \langle\psi_0|\hat{V}^i\hat{\rho}^{-1/2}\hat{V}^{\dagger i}|\psi_0\rangle = \langle\psi_0|\hat{\rho}^{-1/2}|\psi_0\rangle, \quad \forall\, i, \tag{29}$$

we see that this requirement is satisfied. We now proceed to evaluate $\hat{\Gamma}$:

$$\begin{aligned}\hat{\Gamma} &= \frac{1}{N}\sum_{i=0}^{N-1}|\psi_i\rangle\langle\psi_i|\frac{1}{N}\hat{\rho}^{-1/2}|\psi_i\rangle\langle\psi_i|\hat{\rho}^{-1/2}\\ &= \frac{1}{N}\langle\psi_0|\hat{\rho}^{-1/2}|\psi_0\rangle\sum_{i=0}^{N-1}\frac{1}{N}|\psi_i\rangle\langle\psi_i|\hat{\rho}^{-1/2}\\ &= \frac{1}{N}\langle\psi_0|\hat{\rho}^{-1/2}|\psi_0\rangle\hat{\rho}^{1/2}.\end{aligned} \tag{30}$$

To satisfy condition (16) we require

$$\langle\varphi|\left(\hat{\Gamma} - \frac{1}{N}|\psi_i\rangle\langle\psi_i|\right)|\varphi\rangle \geq 0, \quad \forall\, i, |\varphi\rangle. \tag{31}$$

Writing $\hat{\Gamma} = (1/N)\langle\psi_i|\hat{\rho}^{-1/2}|\psi_i\rangle\hat{\rho}^{1/2}$, we can show that

$$\begin{aligned}\langle\varphi|\hat{\Gamma}|\varphi\rangle &= \frac{1}{N}\langle\psi_i|\hat{\rho}^{-1/2}|\psi_i\rangle\langle\varphi|\hat{\rho}^{1/2}|\varphi\rangle\\ &= \frac{1}{N}\langle\psi_i|\hat{\rho}^{-1/4}\hat{\rho}^{-1/4}|\psi_i\rangle\langle\varphi|\hat{\rho}^{1/4}\hat{\rho}^{1/4}|\varphi\rangle\\ &\geq \frac{1}{N}|\langle\psi_i|\hat{\rho}^{-1/4}\hat{\rho}^{1/4}|\varphi\rangle|^2\\ &= \frac{1}{N}|\langle\psi_i|\varphi\rangle|^2,\end{aligned} \tag{32}$$

where we have used the Cauchy–Schwarz inequality. Thus condition (16) holds, and the square-root measurement is optimal. Note that the case of two equiprobable pure states discussed above is an example of a symmetric set. In this case $\hat{V} = \hat{\sigma}_z$, and it may easily be verified that $\hat{\sigma}_z|\psi_0\rangle = |\psi_1\rangle$ and $\hat{\sigma}_z^2 = \hat{\mathbb{1}}$.

Another example of a symmetric set is the so-called trine ensemble [32,40], given by

$$|\psi_0\rangle = |0\rangle,$$

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

$$|\psi_2\rangle = -\frac{1}{2}|0\rangle - \frac{\sqrt{3}}{2}|1\rangle \tag{33}$$

and obtained from one another by rotation through an angle of $2\pi/3$. These states form a resolution of the identity, and the square-root measurement consists of equally weighted projectors onto the states themselves, $\hat{\pi}_i = \frac{2}{3}|\psi_i\rangle\langle\psi_i|$.

The above solution has been extended to multiply symmetric states [37] and mixed states [38,39]. The square-root measurement has also been generalized by Mochon [41], who considered measurements of the form

$$\hat{\pi}_i = \hat{\sigma}^{-1/2} p_i' \hat{\rho}_i \hat{\sigma}^{-1/2}, \tag{34}$$

where $\hat{\sigma} = \Sigma_i p_i' \hat{\rho}_i$, i.e., the square-root measurements corresponding to the same set of states but constructed using different *a priori* probabilities. For pure states, each such measurement is optimal for at least one discrimination problem with the same states, occurring with probabilities given analytically in [41].

3.1c. Other Results

Most of the known results for minimum error discrimination correspond to one of the two cases discussed above: that of just two states, or those for which the square-root measurement is optimal. Another example that is interesting to note is the no-measurement strategy [42]. Sometimes the optimal discrimination strategy is not to measure at all, but just to guess the state which is *a priori* most likely, a measurement that may be described by the POM $\{\hat{\pi}_i = \hat{\mathbb{1}}, \hat{\pi}_j = 0, \forall j \neq i\}$, where i is such that $p_i \geq p_j, \forall j$. Condition (17) holds trivially for this POM. Thus the no-measurement solution is optimal when condition (16) holds, which then reads as

$$p_i \hat{\rho}_i - p_j \hat{\rho}_j \geq 0, \quad \forall j. \tag{35}$$

Clearly this is never optimal if $\hat{\rho}_i$ is pure; a necessary (but not sufficient) condition is that the eigenvectors of $\hat{\rho}_i$ span the entire Hilbert space in which the states $\{\hat{\rho}_j\}$ lie. A practical example is discriminating signal states from random noise, described by the density operator $\hat{\rho}_i \propto \hat{\mathbb{1}}$. If the signal-to-noise ratio is small enough, the minimum error strategy is to always guess that the state received was random noise [42]. It is therefore useful to know the noise threshold at which this occurs, which may be deduced from condition (35).

Other examples for which explicit results are known include three mirror symmetric qubit states, for both pure [43] and mixed states [44], and the case of equiprobable pure states, a weighted sum of which equals the identity operator [13]. The form of the solution for any set of qubit states has also been explored in some detail by Hunter [45,46], including a complete

characterization of the solution for equiprobable pure qubit states. In the general case, for which explicit results are not known, it is possible to deduce both upper [47,48] and lower [49,50] bounds on the error probability. Alternatively, numerical algorithms exist that can find the optimal measurement for a specified set of states to within any desired accuracy [51,52].

3.2. Unambiguous Discrimination

In the minimum error measurement, each possible outcome is taken to indicate some corresponding state. It is perhaps surprising that it is sometimes advantageous to allow for measurement outcomes that do not lead us to identify any state. Suppose again that we wish to discriminate between the two pure states given by Eq. (10), occurring with *a priori* probabilities p_0, p_1. Consider the von Neumann measurement

$$\hat{\pi}_? = |\psi_1\rangle\langle\psi_1|,$$

$$\hat{\pi}_0 = (\sin\theta|0\rangle + \cos\theta|1\rangle)(\sin\theta\langle 0| + \cos\theta\langle 1|). \tag{36}$$

If outcome ?, associated with the probability operator $\hat{\pi}_?$, is realized, we cannot say for sure what state was prepared. However, note that $\langle\psi_1|\hat{\pi}_0|\psi_1\rangle = 0$, and thus when outcome 0, corresponding to POM element $\hat{\pi}_0$, is realized, we can say for certain that the state was $|\psi_0\rangle$. Thus, by allowing for measurement outcome ?, which does not lead us to identify any state, we can construct a measurement that sometimes allows us to determine unambiguously which state was prepared. This measurement, however, only ever identifies the state $|\psi_0\rangle$; ideally we would like to design a measurement that can identify either state unambiguously, at the cost of sometimes giving an inconclusive result. The generalized measurement formalism outlined above allows for exactly such a measurement, a possibility that was first pointed out by Ivanovic [53], Dieks [54], and Peres [55].

Consider therefore the operators

$$\hat{\pi}_0 = a_0(\sin\theta|0\rangle + \cos\theta|1\rangle)(\sin\theta\langle 0| + \cos\theta\langle 1|),$$

$$\hat{\pi}_1 = a_1(\sin\theta|0\rangle - \cos\theta|1\rangle)(\sin\theta\langle 0| - \cos\theta\langle 1|), \tag{37}$$

chosen such that $\langle\psi_0|\hat{\pi}_1|\psi_0\rangle = \langle\psi_1|\hat{\pi}_0|\psi_1\rangle = 0$, and where $0 \le a_0, a_1 \le 1$. Thus when outcome 0 is realized, we can say for sure that the corresponding state was $|\psi_0\rangle$, while when outcome 1 occurs, we know the state was $|\psi_1\rangle$ with certainty. Note that these cannot form a complete measurement for any choice of a_0, a_1, unless $|\psi_0\rangle$, $|\psi_1\rangle$ are orthogonal, and thus an inconclusive outcome is needed, associated with the probability operator

$$\hat{\pi}_? = \hat{\mathbb{1}} - \hat{\pi}_0 - \hat{\pi}_1. \tag{38}$$

The probability of occurrence of the inconclusive result is given by

$$P(?) = p_0\langle\psi_0|\hat{\pi}_?|\psi_0\rangle + p_1\langle\psi_1|\hat{\pi}_?|\psi_1\rangle = 1 - \sin^2 2\theta(p_0 a_0 + p_1 a_1), \tag{39}$$

and the unambiguous discrimination strategy may be further optimized by minimizing this probability, subject to the constraints $a_0, a_1 \ge 0$, $\hat{\pi}_? \ge 0$. For equal *a priori* probabilities, $p_0 = p_1 = \frac{1}{2}$, the minimum value or Ivanovic–Dieks–

Peres (IDP) limit [53–55] is given by $P(?)=\cos 2\theta=|\langle\psi_0|\psi_1\rangle|$ and is achieved by the measurement

$$\hat{\pi}_0=\frac{1}{2\cos^2\theta}(\sin\theta|0\rangle+\cos\theta|1\rangle)(\sin\theta\langle 0|+\cos\theta\langle 1|),$$

$$\hat{\pi}_1=\frac{1}{2\cos^2\theta}(\sin\theta|0\rangle-\cos\theta|1\rangle)(\sin\theta\langle 0|-\cos\theta\langle 1|),$$

$$\hat{\pi}_?=(1-\tan^2\theta)|0\rangle\langle 0|. \tag{40}$$

The optimal $P(?)$ for arbitrary prior probabilities was first given by Jaeger and Shimony [56]. As p_0 is increased, the optimal measurement is given by Eqs. (37) and (38) with

$$a_0=\frac{1-\sqrt{p_1/p_0}\cos 2\theta}{\sin^2 2\theta},$$

$$a_1=\frac{1-\sqrt{p_0/p_1}\cos 2\theta}{\sin^2 2\theta}, \tag{41}$$

giving $P(?)=2\sqrt{p_0p_1}\cos 2\theta$. Thus the measurement becomes biased toward unambiguously identifying the state that is *a priori* more probable. Clearly, when $\sqrt{p_0/p_1}\cos 2\theta>1$, this no longer defines a physical measurement; the optimal measurement then is simply the von Neumann measurement given by Eq. (36). In this case $|\psi_1\rangle$ always gives the inconclusive result, and the probability of failure is $P(?)=p_0|\langle\psi_0|\psi_1\rangle|^2+p_1$. Thus for p_0 much bigger than p_1, the optimal strategy is the one that rules out the less probable state, in contrast to the minimum error measurement, which in this regime (approximately) identifies or rules out the **more** probable state.

A simple example from quantum optics might help to illustrate the main idea [57]. Let us suppose that we have an optical pulse known to have been prepared, with equal probability, in one of the two coherent states [58] $|\alpha\rangle$ or $|-\alpha\rangle$. If we interfere the pulse with a second pulse prepared in the state $|i\alpha\rangle$ by using a 50:50 symmetric beam splitter, then one of the output modes will be left in its vacuum state $|0\rangle$:

$$|\alpha\rangle|i\alpha\rangle\rightarrow|0\rangle|i\sqrt{2}\alpha\rangle,$$

$$|-\alpha\rangle|i\alpha\rangle\rightarrow|-\sqrt{2}\alpha\rangle|0\rangle. \tag{42}$$

The state can be identified simply by detecting the light in the associated output mode. The ambiguous outcome is a consequence of the fact that the coherent states have a nonzero overlap with the vacuum state, and the probability for this result is

$$P_?=|\langle i\sqrt{2}\alpha|0\rangle|^2=|\langle-\sqrt{2}\alpha|0\rangle|^2=|\langle\alpha|-\alpha\rangle|, \tag{43}$$

which is the Ivanovic–Dieks–Peres (IDP) limit.

3.2a. $N>2$ Pure States

In the general case of discriminating unambiguously between N pure states $\{|\psi_i\rangle\}$, $i=0,\ldots,N-1$, we wish to find probability operators $\{\hat{\pi}_i\}$ such that

$$\langle\psi_i|\hat{\pi}_j|\psi_i\rangle = P_i\delta_{ij}, \tag{44}$$

where $0\leq P_i\leq 1$. Thus outcome j is obtained only if the state is $|\psi_j\rangle$, in which case it occurs with probability P_j. We first note that this is only possible if the states $\{|\psi_i\rangle\}$ are linearly independent, as was shown by Chefles [59]. When this is the case, we can construct states $|\psi_j^\perp\rangle$ such that

$$\langle\psi_i|\psi_j^\perp\rangle = \langle\psi_j|\psi_j^\perp\rangle\delta_{ij}, \tag{45}$$

i.e., $|\psi_j^\perp\rangle$ is orthogonal to all allowed states except $|\psi_j\rangle$. The POM elements

$$\hat{\pi}_j = \frac{P_j}{|\langle\psi_j|\psi_j^\perp\rangle|^2}|\psi_j^\perp\rangle\langle\psi_j^\perp| \tag{46}$$

thus satisfy Eq. (44) and unambiguously discriminate between the linearly independent states $\{|\psi_i\rangle\}$. As before, an inconclusive outcome is necessary to form a complete measurement

$$\hat{\pi}_? = \hat{\mathbb{1}} - \sum_j \hat{\pi}_j. \tag{47}$$

The above defines the unambiguous discrimination strategy for N linearly independent states. The occurrence of outcome j indicates unambiguously that the state was $|\psi_j\rangle$. As in the two-state case, a further condition that may be applied is to minimize the probability of obtaining an inconclusive result. Analytical solutions for the minimum achievable $P(?)$ are not known in the general case, but the solution for three states is given by Peres and Terno [60], who also discuss how the method used can be extended to higher dimensions. For the special case in which the probability of unambiguously identifying a state $|\psi_j\rangle$ is the same for all j ($P_j=P$, $\forall j$), the minimum probability of obtaining an inconclusive result is known [59]. Further, the optimal strategy minimizing this probability is given for N linearly independent symmetric states in [61]. For the general case, upper [62] and lower bounds [63,64] have been given for the probability of successful unambiguous discrimination of N linearly independent states, and numerical optimization techniques have also been considered [64,65].

3.2b. Mixed States

It is only relatively recently that unambiguous discrimination has been extended to mixed states [66], where it may be applied to problems such as quantum state comparison [66,67], subset discrimination [68], and determining whether a given state is pure or mixed [69]. Consider the problem of discriminating between two mixed states $\hat{\rho}_0$, $\hat{\rho}_1$, which may be written in terms of their eigenvalues and eigenvectors as follows:

$$\hat{\rho}_0 = \sum_i \lambda_i^{(0)}|\lambda_i^{(0)}\rangle\langle\lambda_i^{(0)}|, \quad \hat{\rho}_1 = \sum_i \lambda_i^{(1)}|\lambda_i^{(1)}\rangle\langle\lambda_i^{(1)}|. \tag{48}$$

where $0<\lambda_i^{(j)}\leq 1$. Define the projectors

$$\hat{\Lambda}_{\ker}^{(0)} = \hat{\mathbb{1}} - \sum_i |\lambda_i^{(0)}\rangle\langle\lambda_i^{(0)}|,$$

$$\hat{\Lambda}_{\ker}^{(1)} = \hat{\mathbb{1}} - \sum_i |\lambda_i^{(1)}\rangle\langle\lambda_i^{(1)}|, \tag{49}$$

such that $\hat{\Lambda}_{\ker}^{(0)}\hat{\rho}_0 = \hat{\Lambda}_{\ker}^{(1)}\hat{\rho}_1 = 0$. These are the projectors onto the kernels of $\hat{\rho}_0$ and $\hat{\rho}_1$, respectively [70]. If we now define $\hat{\pi}_1$ to lie in the kernel of $\hat{\rho}_0$, then $\hat{\pi}_1 = \hat{\Lambda}_{\ker}^{(0)}\hat{\pi}_1\hat{\Lambda}_{\ker}^{(0)}$, and clearly

$$\mathrm{Tr}(\hat{\rho}_0\hat{\pi}_1) = \mathrm{Tr}(\hat{\rho}_0\hat{\Lambda}_{\ker}^{(0)}\hat{\pi}_1\hat{\Lambda}_{\ker}^{(0)}) = 0. \tag{50}$$

Thus if there exists a positive operator $\hat{\pi}_1$ in the kernel of $\hat{\rho}_0$ for which $\mathrm{Tr}(\hat{\rho}_1\hat{\pi}_1) \neq 0$, then $\hat{\rho}_1$ may be unambiguously discriminated from $\hat{\rho}_0$. Similarly, $\hat{\pi}_0$ should lie in the kernel of $\hat{\rho}_1$. Thus a necessary and sufficient condition for unambiguous discrimination between two mixed states is that they have nonidentical kernels, and thus nonidentical supports [66]. Unless the states are orthogonal, an inconclusive outcome will be needed, as before, $\hat{\pi}_? = \hat{\mathbb{1}} - \hat{\pi}_0 - \hat{\pi}_1$. The problem of finding the strategy that minimizes the probability of occurrence of the inconclusive result is again a difficult one, and one that has received much attention in the past few years. The solutions for some special cases are known; some examples are when both states have one-dimensional kernels [66], unambiguous discrimination between a pure and a mixed state, first in two dimensions [71] and later extended to N dimensions [72]; other examples may be found in [73–75]. Reduction theorems given in [76] show that it is always possible to reduce the general problem to one of discriminating two states each of rank r, which together span a $2r$-dimensional space. Thus the simplest case that is not reducible to pure state discrimination is the problem of two rank-2 density operators in a four-dimensional space, which was recently analyzed in detail by Kleinmann *et al.* [77]. Upper and lower bounds for the general case are given in [66,78,79], a further reduction theorem in [74], and numerical algorithms are discussed in [80].

3.3. Maximum Confidence Measurements

As pointed out in the previous section, unambiguous discrimination is possible only when the allowed states are all linearly independent. If this is not the case, there will always be errors associated with identifying some states, even if an inconclusive outcome is allowed. Nevertheless, for more general sets of states we can construct an analogous measurement, one that allows us to be as confident as possible that when the outcome of measurement leads us to identify a given state $\hat{\rho}_i$, that was indeed the state prepared [81]. Just as with unambiguous discrimination, this measurement is concerned with optimizing the information given about the state by particular measurement outcomes, specifically the posterior probabilities

$$P(\hat{\rho}_i|i) = \frac{P(\hat{\rho}_i)P(i|\hat{\rho}_i)}{P(i)}. \tag{51}$$

Physically, in many runs of an experiment, this probability refers to the proportion of occurrences of outcome i that were due to state $\hat{\rho}_i$. In a single-shot measurement, this therefore corresponds to the probability that it was state

$\hat{\rho}_i$ that gave rise to outcome i. Thus, we can think of this quantity as our confidence in taking outcome i to indicate state $\hat{\rho}_i$. In terms of the probability operator $\hat{\pi}_i$ associated with outcome i, we can write

$$P(\hat{\rho}_i|i) = \frac{p_i \mathrm{Tr}(\hat{\rho}_i \hat{\pi}_i)}{\mathrm{Tr}(\hat{\rho}\hat{\pi}_i)}, \tag{52}$$

where $\hat{\rho}=\Sigma_j p_j \hat{\rho}_j$ is the *a priori* density operator. We note that $\hat{\pi}_i$ appears in both the numerator and the denominator of this expression and thus can be determined only up to a multiplicative constant. It is always possible, therefore, to choose the overall normalization such that

$$\sum_i \hat{\pi}_i \leq \hat{\mathbb{1}}, \tag{53}$$

and a physically realizable measurement may be constructed by adding an inconclusive result. Thus the only constraint we need worry about is that $\hat{\pi}_i \geq 0$. Optimization of this figure of merit is greatly facilitated by the use of the ansatz

$$\hat{\pi}_i = \hat{\rho}^{-1/2} \hat{Q}_i \hat{\rho}^{-1/2}, \tag{54}$$

where, by construction, $\hat{\pi}_i \geq 0$ if $\hat{Q}_i \geq 0$. With this, Eq. (52) becomes

$$P(\hat{\rho}_i|i) = \mathrm{Tr}\left(\hat{\rho}^{-1/2} p_i \hat{\rho}_i \hat{\rho}^{-1/2} \frac{\hat{Q}_i}{\mathrm{Tr}(\hat{Q}_i)} \right), \tag{55}$$

where we have used the cyclical property of the trace. Note that $\hat{Q}_i/\mathrm{Tr}(\hat{Q}_i)$ is a positive, trace one operator, and so has the mathematical properties of a density operator. The density operator that has largest overlap with any operator $\hat{A}$ is simply a projector onto the largest eigenvector of $\hat{A}$ (or any density operator in the eigensubspace corresponding to the largest eigenvalue if this is degenerate). For pure states the optimal probability operators are therefore given by

$$\hat{\pi}_i \propto \hat{\rho}^{-1} \hat{\rho}_i \hat{\rho}^{-1} \tag{56}$$

while for mixed states they may be written as

$$\hat{\pi}_i \propto \hat{\rho}^{-1/2} \hat{\sigma}_i \hat{\rho}^{-1/2}, \tag{57}$$

where $\hat{\sigma}_i$ is any density operator lying in the eigensubspace of $\hat{\rho}^{-1/2} p_i \hat{\rho}_i \hat{\rho}^{-1/2}$ corresponding to its largest eigenvalue. Finally, the limit is given by

$$[P(\hat{\rho}_i|i)]_{\max} = \gamma_{\max}(\hat{\rho}^{-1/2} p_i \hat{\rho}_i \hat{\rho}^{-1/2}), \tag{58}$$

where $\gamma_{\max}(\hat{A})$ denotes the largest eigenvalue of $\hat{A}$.

The simplest nontrivial example of a set of linearly dependent states is that of three states in two dimensions. To illustrate this strategy we consider the problem of discriminating between the states

$$|\psi_0\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle,$$

$$|\psi_1\rangle = \cos\theta|0\rangle + e^{2\pi i/3}\sin\theta|1\rangle,$$

$$|\psi_2\rangle = \cos\theta|0\rangle + e^{-2\pi i/3}\sin\theta|1\rangle, \tag{59}$$

where $0 \leq \theta \leq \pi/4$, occurring with equal *a priori* probabilities $p_i = 1/3$, $i = 0, 1, 2$. These states are symmetrically located at the same latitude of the Bloch sphere, as shown in Fig. 2. The *a priori* density operator for this set is

$$\hat{\rho} = \cos^2\theta|0\rangle\langle 0| + \sin^2\theta|1\rangle\langle 1|, \tag{60}$$

and the maximum confidence POM elements may be readily calculated by using Eq. (56). These have the form $\hat{\pi}_i = \alpha_i|\varphi_i\rangle\langle\varphi_i|$, where we have some freedom in choosing the constants α_i, $i = 0, 1, 2$, and

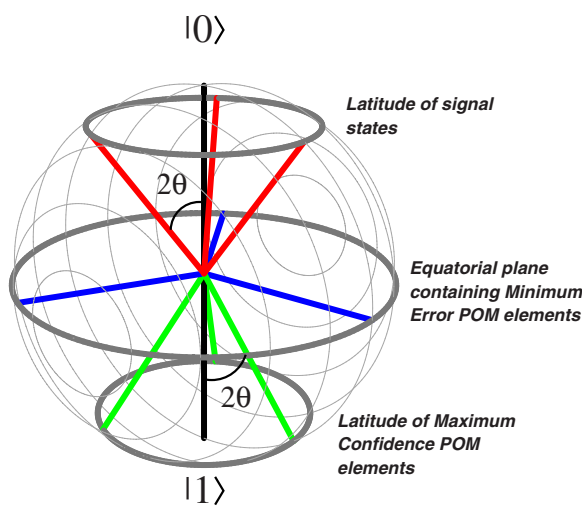
$$|\varphi_0\rangle = \sin\theta|0\rangle + \cos\theta|1\rangle,$$

$$|\varphi_1\rangle = \sin\theta|0\rangle + e^{2\pi i/3}\cos\theta|1\rangle,$$

$$|\varphi_2\rangle = \sin\theta|0\rangle + e^{-2\pi i/3}\cos\theta|1\rangle. \tag{61}$$

These states correspond to reflections of the input states in the equatorial plane of the Bloch sphere, and are also shown in Fig. 2. It is not possible in general to choose $\alpha_0, \alpha_1, \alpha_2$ such that $\{\hat{\pi}_i\}$ form a complete measurement, and thus an additional operator, $\hat{\pi}_? = \hat{\mathbb{1}} - \Sigma_i \hat{\pi}_i$, associated with an inconclusive result, is needed. We may choose to complete the measurement by minimizing the probability of an inconclusive result:

Figure 2



Bloch sphere representation of states. The states used in the example, along with the states onto which the optimal maximum confidence and minimum error POM elements project, are shown [81]. © 2006 by the American Physical Society.

$$P(?) = \mathrm{Tr}(\hat{\rho}\hat{\pi}_?) = 1 - 2(\alpha_0 + \alpha_1 + \alpha_2)\cos^2\theta \sin^2\theta. \tag{62}$$

As $P(?)$ is a monotonically decreasing function of α_i, the optimal values of these parameters lie on the boundary of the allowed domain, defined by the constraint $\hat{\pi}_? \geq 0$. It is straightforward to show that this leads us to choose $\alpha_0 = \alpha_1 = \alpha_2 = (3\cos^2\theta)^{-1}$, giving

$$\hat{\pi}_? = (1 - \tan^2\theta)|0\rangle\langle 0|. \tag{63}$$

It is useful to compare this measurement with the minimum error (ME) measurement, which for this set is given by the square-root measurement discussed earlier:

$$\hat{\pi}_i^{\mathrm{ME}} = \frac{1}{3}\hat{\rho}^{-1/2}|\psi_i\rangle\langle\psi_i|\hat{\rho}^{-1/2} = \frac{2}{3}|\varphi_i^{\mathrm{ME}}\rangle\langle\varphi_i^{\mathrm{ME}}|, \tag{64}$$

where

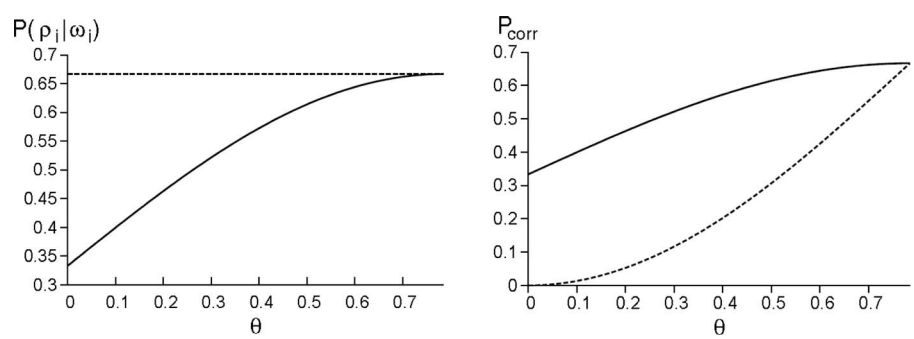
$$|\varphi_0^{\mathrm{ME}}\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle),$$

$$|\varphi_1^{\mathrm{ME}}\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i/3}|1\rangle),$$

$$|\varphi_2^{\mathrm{ME}}\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{-2\pi i/3}|1\rangle). \tag{65}$$

In the Bloch sphere representation, these states correspond to the projection of the input states onto the equatorial plane, as can be seen in Fig. 2. The minimum error and maximum confidence figures of merit are shown for each measurement in Fig. 3. For the minimum error measurement, each outcome

Figure 3



Graphs showing the maximum confidence (left) and minimum error (right) figures of merit, for various values of the parameter θ for the example discussed in the text. In each case the values achieved by the optimal maximum confidence measurement are indicated by a dashed curve, and those corresponding to the optimal minimum error measurement are indicated by a solid curve.

leads us to identify some state, and the average probability of making an error is minimized. However, the confidence in identifying a state may be increased by allowing for an inconclusive result, as may be seen from the plots. When a noninconclusive result is obtained in the maximum confidence measurement, the probability that the state prepared really was the one identified is $\frac{2}{3}$, compared with $\frac{1}{3}(1+\sin 2\theta)$ for the minimum error measurement.

3.3a. Other Similar Strategies

A related strategy may be constructed by applying a worst-case optimality criterion to the conditional probability considered here, $P(\hat{\rho}_i|i)$ [82]. This approach does not allow for inconclusive results, but searches for the measurement for which the smallest value of $P(\hat{\rho}_i|i)$ is maximized. This more complicated problem is difficult to solve analytically, but may be cast as a quasi-convex optimization, for which efficient numerical techniques are available. An alternative strategy allows inconclusive results to occur with a certain fixed probability, P_I, and maximizes the probability of correctly identifying the state when a noninconclusive outcome is obtained. For linearly independent pure states this approach interpolates between minimum error and unambiguous discrimination [83,84]. The performance of projective (von Neumann) measurements versus generalized measurements for strategies with both errors and inconclusive results is analyzed in [85]. Here the rate of inconclusive results is minimized for a bounded-error rate, and it is shown that as small, but experimentally realistic errors are allowed, the advantage of generalized measurements over von Neumann measurements is reduced for some sets of states. For arbitrary mixed states an approach that allows for both errors and inconclusive results is also possible [86] and may be interpreted as interpolating between a minimum error measurement and a maximum confidence strategy. It is clear that the probability of obtaining a correct result, renormalized over only the results that are not inconclusive, denoted P_{RC}, can never be larger than the largest value of $P(\hat{\rho}_i|i)_{\max}$ for a given set, regardless of how much we increase P_I. This upper bound is achieved by a maximum confidence strategy which only ever identifies the state(s) $\hat{\rho}_i$ such that [from Eq. (58)]

$$\gamma_{\max}(\hat{\rho}^{-1/2}p_i\hat{\rho}_i\hat{\rho}^{-1/2}) \geq \gamma_{\max}(\hat{\rho}^{-1/2}p_j\hat{\rho}_j\hat{\rho}^{-1/2})\ \forall\ \hat{\rho}_j, \tag{66}$$

while all other results are interpreted as inconclusive. Although it is difficult to find the optimal measurement for general P_I, it is indeed found that P_{RC} is saturated at some value of P_I, and the maximum P_{RC} achievable corresponds to the strategy outlined here [86].

3.3b. Related Problems—Quantum State Filtration

Quantum state filtration refers to the problem of whether the state of a system is a given state $|\psi_i\rangle$ or simply in any one of the other states in a given set $\{|\psi_j\rangle\}, j \neq i$. This problem is less demanding than complete discrimination among all possible states, and in the minimum error approach the probability of error may be smaller in the state filtration case [87]. For the maximum confidence measurement, however, the optimality of the probability operator $\hat{\pi}_i$ in Eq. (57) is independent of the number and interpretation of other possible outcomes. Thus the confidence in identifying a given state from a set cannot be increased by considering this simpler problem. This figure of merit is

dependent only on the geometry of the set, and in this sense can be thought of as a measure of how distinguishable $\hat{\rho}_i$ is in the given set.

3.4. Comments on the Relationships between Strategies

The maximum confidence strategy was introduced as an analogy to unambiguous discrimination for linearly dependent states [81]. In fact, unambiguous discrimination is a special case of maximum confidence discrimination. The maximum confidence measurement maximizes the conditional probability $P(\hat{\rho}_i|i)$. If this figure of merit is equal to unity for some state $\hat{\rho}_i$, the optimal measurement is such that, when outcome i is obtained, we can be absolutely certain that $\hat{\rho}_i$ was in fact the state received, corresponding to unambiguous discrimination. We can use the maximum confidence formalism to investigate when unambiguous discrimination is possible. Equation (52) may be written as

$$P(\hat{\rho}_i|i) = \frac{p_i \mathrm{Tr}(\hat{\rho}_i \hat{\pi}_i)}{p_i \mathrm{Tr}(\hat{\rho}_i \hat{\pi}_i) + \Sigma_{j\neq i} p_j \mathrm{Tr}(\hat{\rho}_j \hat{\pi}_i)}. \tag{67}$$

Clearly the limit of unity may be achieved if there exists any projector $\hat{\Lambda}_i$ for which $\hat{\Lambda}_i \Sigma_{j\neq i} p_j \hat{\rho}_j \hat{\Lambda}_i = 0$ while $\hat{\Lambda}_i \hat{\rho}_i \hat{\Lambda}_i$ is nonzero. $\hat{\pi}_i$ is then any operator lying in the subspace with projector $\hat{\Lambda}_i$. This reproduces the known results that unambiguous discrimination is possible between pure states if the states are linearly independent [59] and between mixed states if they have distinct supports [66]. More precisely, a measurement is possible that will sometimes allow us to identify $\hat{\rho}_i$ unambiguously if $\hat{\rho}_i$ has support in the kernel of $\Sigma_{j\neq i} p_j \hat{\rho}_j$. This condition is less restrictive than the previous, which does not hold in the case where it is possible to unambiguously discriminate some but not all states in a set. Unambiguous discrimination is still possible in this case, but some states are never identified. For example, it was pointed out by Sun *et al.* [71] that it is possible to apply unambiguous discrimination to the problem of determining whether a system is in a given state $|\psi_0\rangle$ or in either of two other possible states, $|\psi_1\rangle$, $|\psi_2\rangle$, even if the states span only two dimensions and therefore are linearly dependent. This may be more easily understood as unambiguous discrimination between a mixed state and a pure state in two dimensions [66]. Let

$$\rho_0 = |\psi_0\rangle\langle\psi_0|,$$

$$\rho_1 = \frac{p_1}{p_1+p_2}|\psi_1\rangle\langle\psi_1| + \frac{p_2}{p_1+p_2}|\psi_2\rangle\langle\psi_2| = q|0\rangle\langle 0| + (1-q)|1\rangle\langle 1|, \tag{68}$$

where $|0\rangle$, $|1\rangle$ are the eigenkets of ρ_1, $0<q<1$, and without loss of generality we can write $|\psi_0\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle$. It is clear that the von Neumann measurement

$$\hat{\pi}_0 = |\psi_0\rangle\langle\psi_0|,$$

$$\hat{\pi}_1 = \hat{\mathbb{1}} - |\psi_0\rangle\langle\psi_0| \tag{69}$$

can unambiguously discriminate the two possibilities—if outcome 1 is obtained, we can say for sure that the state was ρ_1, while the result 0 is interpreted as inconclusive. However, this measurement never tells us if the state was $|\psi_0\rangle$. In this case it may be useful to consider unambiguous discrimination within the framework of maximum confidence measurements. It is then possible to construct a measurement that sometimes identifies $\hat{\rho}_1$ with certainty, but also sometimes identifies $\hat{\rho}_0$ as confidently as possible. In general an inconclusive result will also be necessary.

Now suppose that, instead of maximizing the conditional probability in Eq. (52) independently for each state in the set, we choose to maximize a weighted average of these probabilities. We would then obtain as our figure of merit

$$P(\hat{\rho}_i|i)_{\text{avg}} = \sum_i P(i)P(\hat{\rho}_i|i) = \sum_i P(\hat{\rho}_i)P(i|\hat{\rho}_i), \tag{70}$$

which is precisely the figure of merit maximized by the minimum error measurement. Thus these two strategies can be thought of as applying a different optimality condition to the same quantity. The minimum error measurement also has the additional constraint that the operators $\{\hat{\pi}_i\}$ must form a complete measurement, as it is never optimal to allow inconclusive results to occur. This constraint makes finding the optimal measurement a difficult problem, although the conditions that the optimal measurement must satisfy are known, as we have shown. By contrast, the maximum confidence strategy allows a closed form solution for an arbitrary set of states. In the special case where the maximum confidence figure of merit is the same for all states $\hat{\rho}_i$ and no inconclusive result is needed, the two strategies coincide. More generally, it is clear by examination of Eq. (70) that an upper bound for the minimum error figure of merit is given by the largest value of $P(\hat{\rho}_i|i)_{\max}$ for a given set [i.e., the largest value of Eq. (58)].

3.5. Mutual Information

In communications theory the performance of a communications channel is quantified not by an error probability but rather by the information conveyed. We can give a precise meaning to this by invoking Shannon's noisy channel coding theorem [88,89], which states that the maximum communications rate, or channel capacity, is obtained by maximizing the mutual information between the transmitter and receiver. If the transmitted message, A, is one of the set $\{a_i\}$ and the reception event, B, is one of the set $\{b_j\}$, then the mutual information is defined to be

$$H(A{:}B) = \sum_{ij} P(a_i,b_j)\log\left(\frac{P(a_i,b_j)}{P(a_i)P(b_j)}\right), \tag{71}$$

where the logarithm is usually taken to be base 2 so that the information is expressed in bits. For a quantum channel, the state $\hat{\rho}_i$ is selected with probability p_i, and the measurement result b_j is associated with the probability operator $\hat{\pi}_j$. It follows that the mutual information is

$$H(A{:}B) = \sum_{ij} p_i \mathrm{Tr}(\hat{\rho}_i \hat{\pi}_j) \log\left(\frac{\mathrm{Tr}(\hat{\rho}_i \hat{\pi}_j)}{\mathrm{Tr}(\hat{\rho}\hat{\pi}_j)}\right), \tag{72}$$

where $\hat{\rho}=\Sigma_i p_i \hat{\rho}_i$. The maximum value of the mutual information is found by varying both the preparation probabilities, p_i, and the measurement strategies. This is a very difficult optimization problem, and there are very few exact solutions known [90,91]. A scarcely simpler problem is to fix the preparation probabilities and then seek the maximum value to give what is referred to as the accessible information [92].

For two pure states, it is known that the mutual information is maximized if the states are prepared with equal probability and if the minimum error measurement is employed [91]. For three or more states, the accessible information is known if the states are equally likely to be selected and possess a degree of symmetry. In particular, for the so-called trine ensemble of three equally probable states (33), the accessible information is obtained with a generalized measurement with probability operators

$$\hat{\pi}_0 = \frac{2}{3}|1\rangle\langle 1|,$$

$$\hat{\pi}_1 = \frac{2}{3}\left(\frac{1}{2}|1\rangle + \frac{\sqrt{3}}{2}|0\rangle\right)\left(\frac{1}{2}\langle 1| + \frac{\sqrt{3}}{2}\langle 0|\right),$$

$$\hat{\pi}_2 = \frac{2}{3}\left(\frac{1}{2}|1\rangle - \frac{\sqrt{3}}{2}|0\rangle\right)\left(\frac{1}{2}\langle 1| - \frac{\sqrt{3}}{2}\langle 0|\right). \tag{73}$$

Note that the accessible information is obtained not by maximizing the probability for determining the state but rather for **eliminating** one of the states so that

$$\langle \psi_i|\hat{\pi}_j|\psi_i\rangle = \frac{1}{2}(1-\delta_{ij}). \tag{74}$$

A similar strategy works well for four equiprobable states arranged so as to form a regular tetrahedron on the Bloch or Poincaré sphere [90]. For more states, optimal strategies have been demonstrated with fewer measurement outcomes than states [92].

3.6. No-Signaling Bounds on State Discrimination

Until now we have discussed the limits on quantum state discrimination by mathematically formulating figures of merit that may then be evaluated and compared for any allowed measurement by virtue of the generalized measurement formalism. It is interesting to note, however, that it is possible to place tight bounds on state discrimination without any reference to generalized measurements by appealing to the no-signaling principle, the condition that information may not propagate faster than the speed of light.

Although entanglement appears to allow spacelike separated quantum systems to influence one another instantaneously, it may be shown that quantum

mechanical correlations do not allow signaling [93–96]. Further, owing to the implications of this in reconciling quantum mechanics with special relativity, it has been suggested that the no-signaling principle be given the status of a physical law, which may be used to limit quantum mechanics and possible extensions of it [97,98]. In practice, bounds on the fidelity of quantum cloning machines [99,100], the success probability of unambiguous discrimination [101,102], and the maximum confidence figure of merit [103] have been derived by using no-signaling arguments. In particular, the no-signaling principle may be used to put a tight bound on unambiguous discrimination of two pure states [101] and to derive the maximum confidence strategy [103]. We will discuss these two cases here.

3.6a. Unambiguous Discrimination

Consider the entangled state

$$|\Psi\rangle = \sqrt{p_0}|\psi_0\rangle_L|0\rangle_R + \sqrt{1-p_0}|\psi_1\rangle_L|1\rangle_R, \tag{75}$$

where $|\psi_0\rangle_L$, $|\psi_1\rangle_L$ are nonorthogonal states of the left system [given by Eq. (10)], and $|0\rangle_R$, $|1\rangle_R$ form an orthonormal basis for the right system. The reduced density operator of the right system may be obtained by taking the partial trace over the left system and is given by

$$\hat{\rho}_R = \mathrm{Tr}_L(|\Psi\rangle\langle\Psi|) = \begin{pmatrix} p_0 & \sqrt{p_0(1-p_0)}\cos 2\theta \\ \sqrt{p_0(1-p_0)}\cos 2\theta & 1-p_0 \end{pmatrix}. \tag{76}$$

According to the no-signaling principle, no operation performed on the left system may be detected by measurement of the right system alone, as this could be used to signal faster than light. Thus, after any physically allowed transformation of the left system, the reduced density operator of the right system must remain the same. Consider now a measurement that discriminates unambiguously between the states $|\psi_0\rangle_L$, $|\psi_1\rangle_L$ of the left system. If outcome 0 is realized, which occurs with some probability q_0, the right system is projected into the state $|0\rangle_R$, due to the initial entanglement between the systems. Similarly outcome 1 projects the right system into state $|1\rangle_R$, with probability q_1. There is also the inconclusive result, which transforms the right system to some as yet unknown state

$$\hat{\rho}_? = \begin{pmatrix} \rho_?^{00} & \rho_?^{01} \\ \rho_?^{10} & \rho_?^{11} \end{pmatrix} \tag{77}$$

with probability $q_?$. No signaling implies

$$\hat{\rho}_R = \begin{pmatrix} q_0 & 0 \\ 0 & q_1 \end{pmatrix} + q_? \begin{pmatrix} \rho_?^{00} & \rho_?^{01} \\ \rho_?^{10} & \rho_?^{11} \end{pmatrix}. \tag{78}$$

The task is then to minimize $q_?$ subject to the above condition and the conditions $\hat{\rho}_? \geq 0$, q_0, q_1, $q_? \geq 0$. This optimization is straightforward [101] and, remarkably, gives precisely the Jaeger and Shimony result [56] discussed in Subsection 3.2. Thus the no-signaling condition may be used to place a tight bound on the success probability of unambiguous discrimination, without any reference to generalized measurements.

3.6b. Maximum Confidence Measurements

The confidence in identifying a given state $|\psi_j\rangle$ as a result of a state discrimination measurement on the ensemble $\{|\psi_i\rangle, p_i\}$ is simply the probability that it was state $|\psi_j\rangle$ that gave rise to the measurement outcome observed. Consider now the entangled state

$$|\Psi\rangle = \sum_{i=0}^{N-1} \sqrt{p_i}|\psi_i\rangle_L|i\rangle_R, \tag{79}$$

where $\{|\psi_i\rangle_L\}$ are nonorthogonal states of the left system which together span a $D \leq N$-dimensional space, and $\{|i\rangle_R\}$ forms an orthonormal basis for the right system. Now for any measurement performed on the left system of the entangled pair, the probability that it was state $|\psi_j\rangle_L$ which gave rise to the observed outcome is equivalent to the probability that the right system is now found in state $|j\rangle_R$. Thus, if measurement outcome j causes the right system to transform to $\hat{\rho}_{R|j}$, we can write

$$P(\psi_j|j) = {}_R\langle j|\hat{\rho}_{R|j}|j\rangle_R. \tag{80}$$

It may be shown (by reference to the Schmidt decomposition of $|\Psi\rangle$ [104]) that although the right system lies in an N-dimensional Hilbert space, it is confined to a D-dimensional subspace (with the projector denoted $\hat{P}_D$ below) because of the entanglement with the left system. The key point, then, is to notice that any operation performed on the left system cannot take the right system out of this subspace, since this could be detected with some probability by a measurement on the right system alone, and thus could be used to signal. Thus ${}_R\langle j|\hat{\rho}_{R|j}|j\rangle_R$ is restricted by the requirement that $\hat{\rho}_{R|j}$ lies in this subspace, and is clearly bounded by the magnitude of the projection of $|j\rangle_R$ onto this space:

$$P(\psi_j|j) = {}_R\langle j|\hat{\rho}_{R|j}|j\rangle_R \leq {}_R\langle j|\hat{P}_D|j\rangle_R. \tag{81}$$

Further, this bound is achievable and is equivalent to that obtained previously [Eq. (58)] [103]. Similar arguments may be applied to the mixed state case, and the maximum confidence strategy is derived in a natural way from no-signaling considerations. Finally, we note that in the case where the states $\{|\psi_i\rangle_L\}$ are linearly independent, $D=N$, and the right system occupies the entire N-dimensional Hilbert space. In this case the limit is unity, corresponding to unambiguous discrimination.

4. State Discrimination—Experiments

The theory of generalized measurements has a mathematically appealing generality in that it depends only on the overlaps of the possible states to be discriminated and on the probabilities that each was the state prepared. The nature of the physical states, be they nuclear spins, optical coherent states, or electronic energy levels in an atom, is unimportant. In performing experimental demonstrations, however, the choice of physical system is of primary importance. We require a physical system in which superpositions are relatively stable, easy to prepare and to manipulate, and also, of course, to measure. For all of these reasons, the system of choice has usually been photon polarization and forms the basis of our review.

4.1. Photon Polarization

At least within paraxial optics [105], the electric and magnetic fields are very nearly perpendicular to the direction of propagation of the light. It is conventional to define the polarization by the orientation of the electric field in this transverse plane [106]. Two orthogonal polarizations then correspond to fields in which the electric fields are oriented at 90° to each other. The polarization of a single photon is an excellent two-state quantum system, or qubit [4,104], as we can identify the states of horizontal and vertical polarization with the logical $|0\rangle$ and $|1\rangle$ states of a qubit:

$$|0\rangle = |H\rangle, \quad |1\rangle = |V\rangle. \tag{82}$$

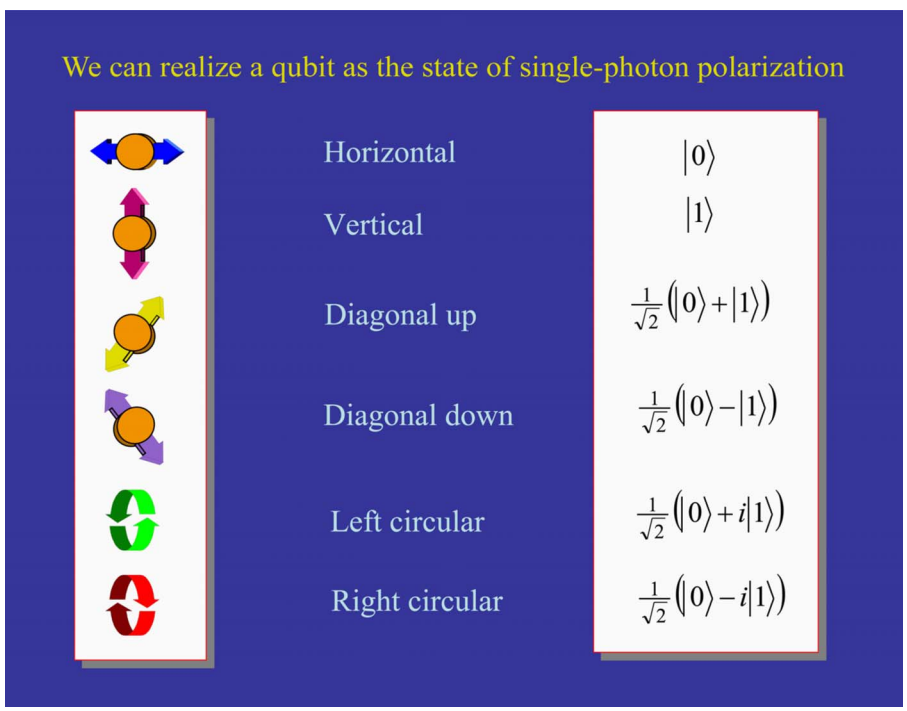
Other polarizations are superpositions of these states. In particular, as illustrated in Fig. 4, linear polarization at ±45° to the horizontal and circular polarizations are the superpositions

$$|+45°\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad |-45°\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle),$$

$$|L\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle), \quad |R\rangle = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle). \tag{83}$$

The set of all possible pure states of polarization can be represented on the surface of a sphere, the Poincaré sphere [107,108], which is a representation equivalent to the Bloch sphere used for qubits in quantum information

Figure 4



Polarization of light as a two-level system, or qubit.

theory [4,104]. States of optical polarization can be changed coherently by delaying one polarization compared with the orthogonal polarization, usually by a quarter or half a wavelength, by using birefringent wave plates. A combination of three suitably oriented half- and quarter-wave plates can perform any desired transformation, corresponding to a rotation on the Poincaré sphere through any desired angle about any desired axis. In this way we can realize any desired single-qubit unitary transformation.

It is important, in order to realize generalized measurements, to be able to superpose fields and also to be able to spatially separate different polarizations. These tasks can be performed using beam splitters and polarizing beam splitters. For fully overlapping modes with the same frequency, we can write the output annihilation operators in terms of those for the input modes. For a symmetric polarization-independent beam splitter we find [58]

$$\hat{a}_3^{H,V} = r\hat{a}_1^{H,V} + t\hat{a}_2^{H,V},$$

$$\hat{a}_4^{H,V} = t\hat{a}_1^{H,V} + r\hat{a}_2^{H,V}, \tag{84}$$

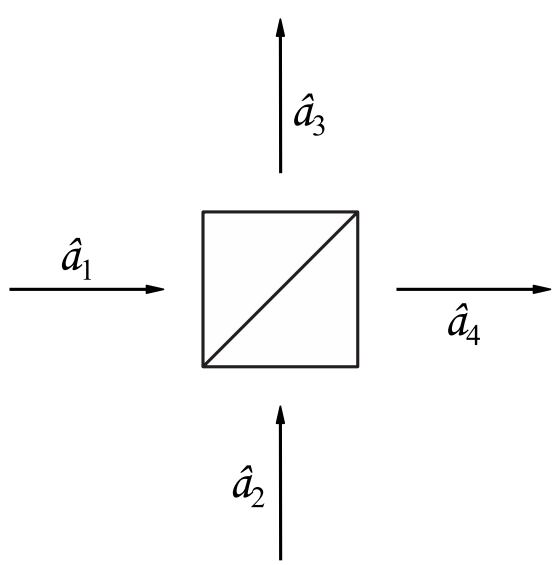
where the input and output modes are labeled as in Fig. 5.

Enforcing the canonical commutation relations for the output modes constrains the reflection and transmission coefficients:

$$|t|^2 + |r|^2 = 1, \quad rt^* + tr^* = 0. \tag{85}$$

A polarizing beam splitter is designed to transmit horizontally polarized light and to reflect vertically polarized light. This means that input and output annihilation operators are related by

Figure 5



A beam splitter can be used to superpose or separate field modes. The input and output modes are labeled with the associated annihilation operators.

$$\hat{a}_3^H = \hat{a}_2^H, \quad \hat{a}_3^V = \hat{a}_1^V,$$

$$\hat{a}_4^H = \hat{a}_1^H, \quad \hat{a}_4^V = \hat{a}_2^V. \tag{86}$$

In correlating photon polarization and direction, a polarizing beam splitter can be used to prepare (filter) light with a desired polarization or, in conjunction with photodetectors placed in each output beam, to measure the polarization. They also allow us to perform different transformations on two orthogonal polarizations, and this is crucial in enabling us to perform generalized measurements.

We should make one important point before describing any of the experiments that have been performed, and this is that they have not been done with single-photon sources. All of them rely on linear optical elements and processes, and for these the single-photon probability amplitudes and the associated probabilities behave in the same way as the amplitudes and intensities of classical optics. Some of the experiments have been performed at light levels in the quantum regime, however, and this suggests strongly that the devices will work in the same way given single-photon sources and detectors.

4.2. Minimum Error Discrimination

4.2a. Two States

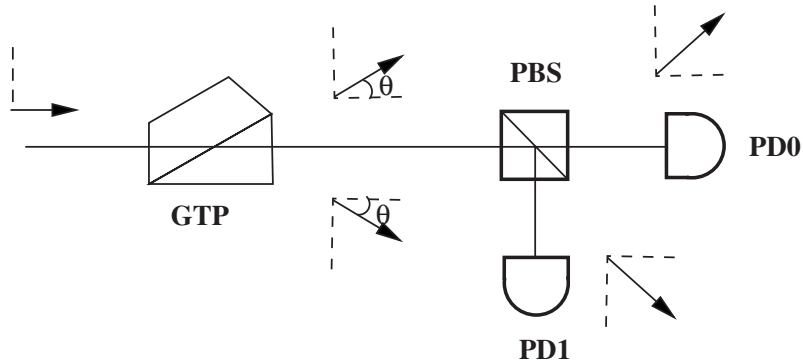
The simplest minimum error problem is, as we have seen, that for two pure states, Eqs. (10). For the photon polarizations described above these correspond to two states of linear polarization, oriented at $+\theta$ and $-\theta$ to the horizontal, so that the angle between them is 2θ, for a range of values of θ between 0 and $\pi/4$. If the two states are prepared with equal prior probability, then, as we have seen, the minimum error measurement corresponds to a familiar von Neumann, or projective, measurement with two projectors associated with the orthogonal states, Eqs. (13). For optical polarization, this corresponds to measuring the polarization at 45° to the horizontal. Thus the minimum error strategy in this case is a simple polarization measurement. The experiment to test this [109] was performed by using light pulses with on average 0.1 photons per pulse prepared in the desired polarization state by use of a Glan–Thompson polarizer oriented so as to produce polarized light at the angle $+\theta$ or $-\theta$ to the horizontal. These were then measured by using a polarizing beam splitter oriented so as to transmit light polarized at +45° to the horizontal and to reflect the orthogonal polarization. The experimental apparatus is shown in Fig. 6. Results, shown in Fig. 7, were found to be in excellent agreement with the Helstrom value (12) for equal prior probabilities:

$$P_{\text{err}} = \frac{1}{2}(1 - \sin 2\theta). \tag{87}$$

4.2b. Three or Four States

Finding a minimum error strategy for discriminating between more than two states is, in general a difficult problem, although very general statements about the solution can be made for qubits [45]. For the trine ensemble of three equiprobable linear polarization states

Figure 6



Schematic of the Barnett–Riis experiment achieving the Helstrom bound for state discrimination between two pure states. GTP, Glan–Thompson polarizer; PBS, polarizing beam splitter; PD0, PD1, photodetectors.

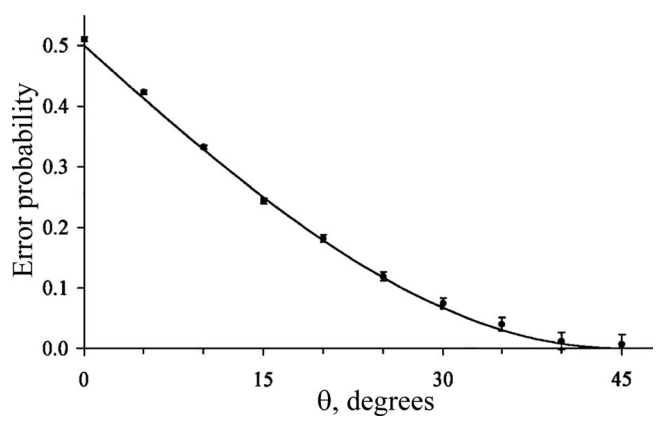
$$|\psi_1^3\rangle = -\frac{1}{2}|H\rangle - \frac{\sqrt{3}}{2}|V\rangle,$$

$$|\psi_2^3\rangle = -\frac{1}{2}|H\rangle + \frac{\sqrt{3}}{2}|V\rangle,$$

$$|\psi_3^3\rangle = |H\rangle \tag{88}$$

and the tetrad ensemble of four equiprobable states

Figure 7



Results from the Barnett–Riis experiment demonstrating minimum error state discrimination at the Helstrom bound. Reproduced with permission from [109], http://www.informaworld.com.

$$|\psi_1^4\rangle = \frac{1}{\sqrt{3}}(-|H\rangle + \sqrt{2}e^{-2\pi i/3}|V\rangle),$$

$$|\psi_2^4\rangle = \frac{1}{\sqrt{3}}(-|H\rangle + \sqrt{2}e^{2\pi i/3}|V\rangle),$$

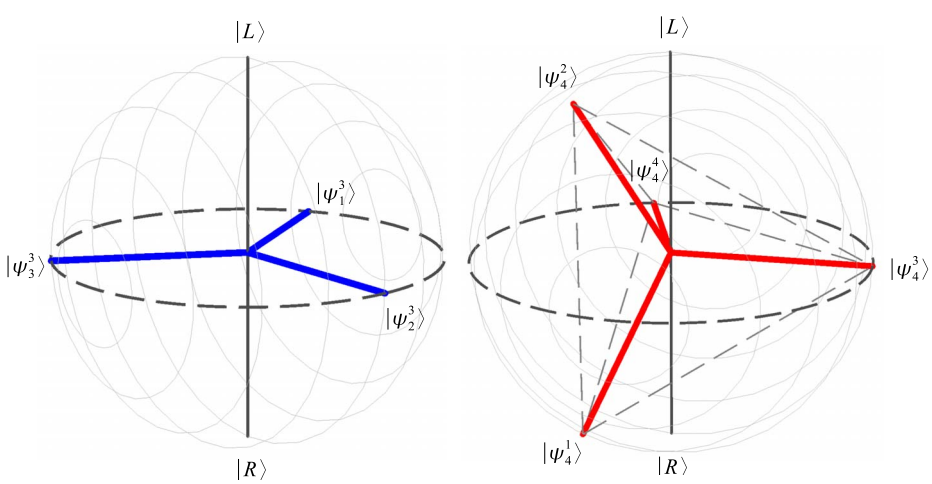
$$|\psi_3^4\rangle = \frac{1}{\sqrt{3}}(-|H\rangle + \sqrt{2}|V\rangle),$$

$$|\psi_4^4\rangle = |H\rangle \tag{89}$$

the square-root measurement is readily shown to give the minimum probability for error. The trine states are states of linear polarization separated by 60°, and the tetrad states are two states of linear polarization and two of elliptical polarization. In each case they form a set of maximally separated points on the surface of the Poincaré sphere, as shown in Fig. 8.

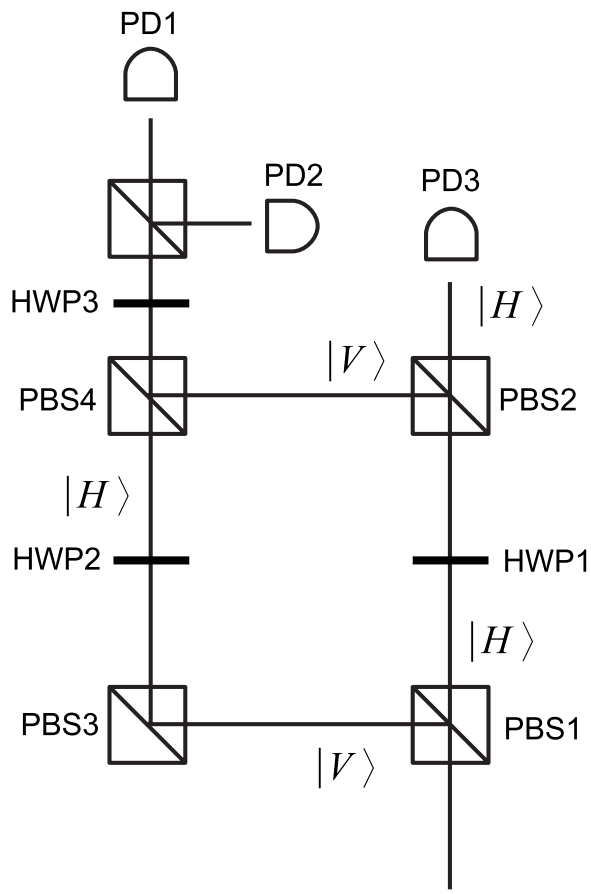
To measure more than two orthogonal states of polarization we need to introduce an additional degree of freedom, and a suitable one is provided by the path of the light beam. We shall illustrate this idea only for the trine ensemble, the experimental setup for which is shown in Fig. 9. Details for the tetrad ensemble can be found in [110]. The input polarizing beam splitter separates, coherently, the polarization components by transmitting the horizontal component and reflecting the vertical component. This allows us to manipulate these components independently. A half-wave plate placed in the path of the horizontally polarized beam rotates the polarization so that only the requisite fraction of it is transmitted at the next polarizing beam splitter. The vertically polarized beam is transformed into a horizontally polarized beam so that it can be recombined coherently with what is left of the originally horizontally polarized beam. Thus the polarization of this combined beam is analyzed by using a final polarizing beam splitter. The photon ends up in one of the three

Figure 8



Representation of the trine (left) and tetrad (right) states on the Poincaré sphere.

Figure 9



Schematic of the Clarke *et al.* experimental realization of minimum error discrimination between the trine states. PBS1–PBS4, polarizing beam splitters; HWP1–HWP3, half-wave plates; PD1–PD3, photodetectors. For details see [110].

photodetectors, and we can think of each of the trine polarization states being transformed into a superposition of exit paths from the interfermometer [110]:

$$|\psi_1^3\rangle \to -\frac{1}{\sqrt{6}}|P3\rangle - \frac{\sqrt{2}}{\sqrt{3}}|P1\rangle - \frac{1}{\sqrt{6}}|P2\rangle,$$

$$|\psi_2^3\rangle \to -\frac{1}{\sqrt{6}}|P3\rangle + \frac{1}{\sqrt{6}}|P1\rangle + \frac{\sqrt{2}}{\sqrt{3}}|P2\rangle,$$

$$|\psi_3^3\rangle \to \frac{\sqrt{2}}{\sqrt{3}}|P3\rangle + \frac{1}{\sqrt{6}}|P1\rangle - \frac{1}{\sqrt{6}}|P2\rangle, \tag{90}$$

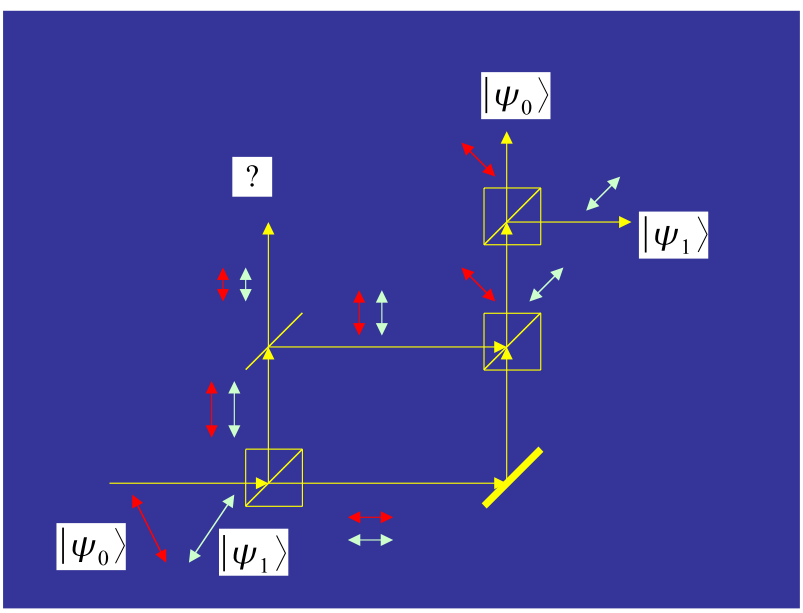
where a photon in path Pi will be detected in photodetector i. This measurement device is optimal, as it correctly identifies the initial polarization state with probability $2/3$.

4.3. Unambiguous Discrimination

Unambiguous discrimination between nonorthogonal polarization states, like the minimum error measurements described above, requires an extension of the two-dimensional state space, and an interferometer is the ideal device for implementing this. The idea is depicted in Fig. 10. We have two possible linear polarization states, each of which has a larger vertical component of polarization than horizontal. The double-headed arrows are intended to represent the magnitudes of the probability amplitudes at various places. The input polarizing beam splitter reflects the vertical component and transmits the horizontal component. The mirror in the upper arm of the interferometer transmits just enough for the reflected field to have the same amplitude as that in the lower arm. If the photon escapes from the interferometer at this point, then the measurement is inconclusive. If it does not, however, then the amplitudes for the vertical and horizontal fields are equal in magnitude and become orthogonal when recombined at the output polarizing beam splitter. At this stage they can be discriminated with certainty by using a final, suitably oriented, polarizing beam splitter.

The first demonstration of unambiguous discrimination between nonorthogonal polarization states used a specially selected length of polarization maintaining fiber [111]. This has the effect of maintaining, with low losses, the horizontal component of polarization but attenuating the orthogonal vertical component. If the length of the fiber is chosen appropriately, then any light exiting the fiber will be in one of two orthogonal polarizations and so can be discriminated with certainty. An interferometric experiment has the advantage that it allows us to measure the ambiguous results, also. The experimental setup [112] is very similar to that for the minimum error discrimination of the three trine states, but with the three measured outputs

Figure 10



Schematic of the Clarke *et al.* experimental realization of unambiguous discrimination between two nonorthogonal polarization states.

now corresponding to the unambiguous identification of the states $|\psi_0\rangle$, $|\psi_1\rangle$ and to the ambiguous result. The results of this experiment are shown in Fig. 11.

We have presented here only the simplest experiments, but more complicated problems have also been addressed. In particular, unambiguous discrimination has been demonstrated for three possible states and also between nonorthogonal pure and mixed states [113]. The generalized measurements described here have all been implemented by using light, but the principles are independent of the system used. It should be noted, particularly in the context of quantum information, that nonorthogonal states encoded in the energy levels of atoms or ions can similarly be subjected to generalized measurements with unoccupied levels used to assist in the process [114].

4.4. Maximum Confidence Measurements

Maximum confidence discrimination between three symmetric states in two dimensions (the simplest possible case) has also been demonstrated experimentally by using the polarization of light as a qubit [115,116]. In the experimental realization, the states given in Eq. (59) were encoded in the left–right circular polarization basis, and the setup distinguished between the elliptical polarizations

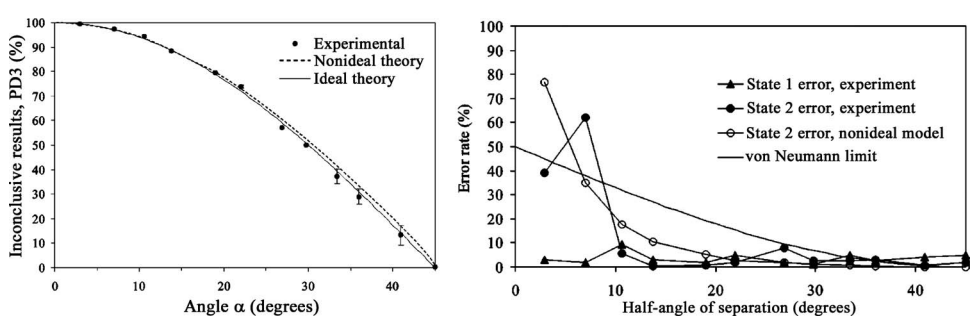
$$|\psi_0\rangle = \cos\theta|R\rangle + \sin\theta|L\rangle,$$

$$|\psi_1\rangle = \cos\theta|R\rangle + e^{2\pi i/3}\sin\theta|L\rangle,$$

$$|\psi_2\rangle = \cos\theta|R\rangle + e^{-2\pi i/3}\sin\theta|L\rangle. \tag{91}$$

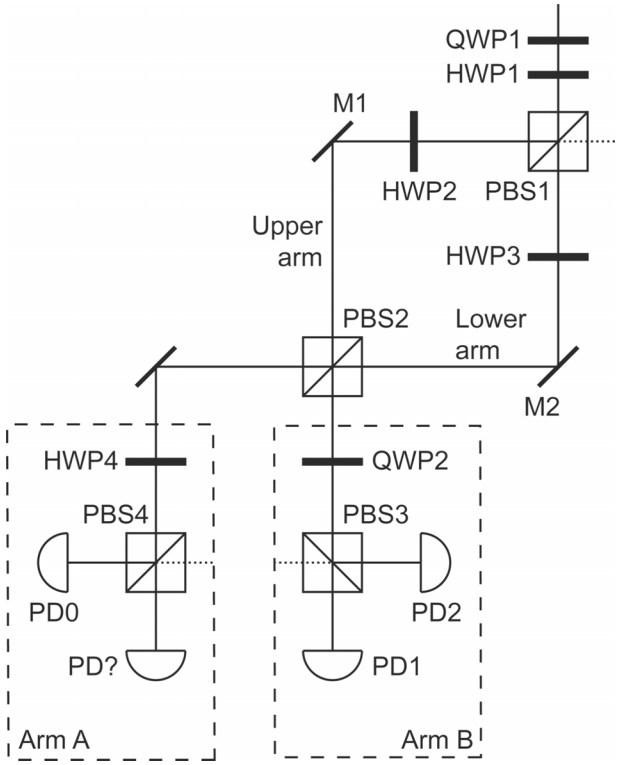
The maximum confidence measurement for these states, as we have seen, has four outcomes, one corresponding to each possible state and one inconclusive result. The apparatus used is shown in Fig. 12 and again features an interferometer to provide the extension to the state space necessary to realize all four outcomes. In this setup, the outcomes 0 and ? are grouped together

Figure 11



Results of the Clarke *et al.* experimental realization of unambiguous discrimination between two nonorthogonal polarization states. The rate of inconclusive results is shown on the left, and the error rate for each initial state is given on the right. A model taking into account the nonideal characteristics of the beam splitters was used to generate the nonideal theory plots in each case. For full details see [112]. © 2001 by the Americal Physical Society.

Figure 12



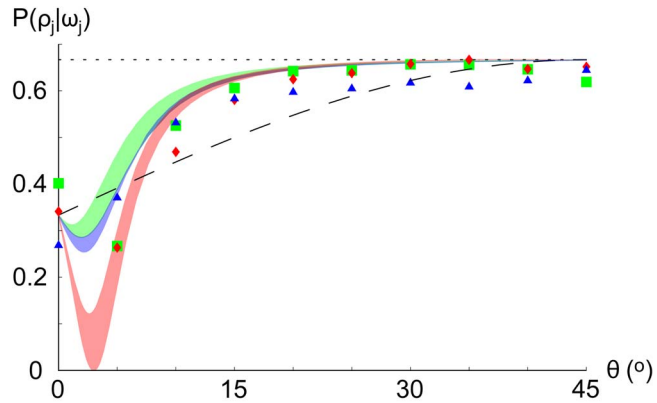
Schematic of the experimental apparatus used to demonstrate maximum confidence discrimination between three elliptical polarization states. PBS1–PBS4, polarizing beam splitters; QWP1–QWP4, quarter-wave plates; HWP1–HWP4, half-wave plates; PD0–PD2, PD?, photodetectors.

in one output arm of the interferometer, while the other arm corresponds to outcomes 1 and 2. Thus two detectors placed in output arms A and B of the apparatus would realize the two outcome generalized measurement described by the POM $\{\hat{\pi}_? + \hat{\pi}_0, \hat{\pi}_1 + \hat{\pi}_2\}$. In fact this setup is completely general and, by appropriate choice of orientations of the wave plates QWP1 and HWP1–3, may be used to implement any such two-element measurement. Further, any N outcome measurement may be, in principle, performed by using a number of such modules in series [116,117]. Thus, after PBS2, two orthogonal modes in arm A correspond to outcomes 0 and ?, while two orthogonal modes in arm B correspond to results 1 and 2. Finally HWP4, QWP2, and PBS3–4 are used to separate these modes, which are then detected at the photodetectors in the output arms. The results of this experiment demonstrated an improvement over the minimum error measurement in the confidence figure of merit for linearly dependent states and are shown in Fig. 13.

4.5. Mutual Information

The strategies for maximizing the mutual information for two pure states require us to perform a minimum error measurement [91]. With more states we require, in general, a generalized measurement [90,92]. For the trine and tetrad states we obtain the accessible information by eliminating, with certainty,

Figure 13



Results of the maximum confidence discrimination experiment, showing the confidence figure of merit for measurement outcomes 0 (red), 1 (green), and 2 (blue). Lines indicate the theoretical value of the figure of merit for the maximum confidence (dotted) and minimum error (dashed) measurement strategies. Shaded areas indicate the range of values consistent with a nonideal model, taking into account errors introduced at the polarizing beam splitters; for details see [116]. Figure reproduced from [115], © American Physical Society.

one of the possible states. This can be realized experimentally by using the same device as that devised for the minimum error measurement, simply by interchanging everywhere the horizontal and vertical components of polarization. In other words, the device for maximizing the mutual information for the trine or tetrad states is the same as that for minimizing the error in discriminating between a set of states **orthogonal** to the given trine or tetrad. For more than four states of linear polarization, we can maximize the mutual information by performing a measurement with just three possible outcomes [92].

The experiment to realize the minimum error discrimination between two nonorthogonal polarization states [109] also provided the maximum mutual information. For the pure states of Eqs. (10) with $\theta = 15°$, corresponding to linear polarizations at an angle of 30°, the mutual information derived from the measurements was [18]

$$H^{2\ \text{states}}(A{:}B) = 0.196 \pm 0.007 \text{ bits}, \tag{92}$$

which compares well with the theoretical value of 0.189 bits. For the trine and the tetrad [110] we found

$$H^{\text{trine}}(A{:}B) = 0.491^{+0.011}_{-0.027} \text{ bits}, \quad H^{\text{tetrad}}(A{:}B) = 0.363^{+0.09}_{-0.024} \text{ bits}, \tag{93}$$

which should be compared with the theoretical values of 0.585 bits and 0.415 bits, respectively. It is important to note that these experimental values are good enough to demonstrate the necessity of performing a generalized measurement, as the theoretical maximum mutual information for the trine and tetrad states using conventional projective measurements are 0.459 bits and 0.311 bits, respectively. A subsequent, more careful experiment produced

a substantially higher value for the mutual information obtained by using the trine ensemble of 0.556 bits and also realized the optimal measurements for sets of five and seven states of linear polarization [118].

5. Conclusion

Quantum theory allows us to prepare, at least in principle, even the simplest system in an uncountable infinity of different ways. The polarization for a single photon, for example, can be prepared in a state that corresponds to any point on the surface of the Poincaré sphere. It is a fundamental consequence of the superposition principle, however, that no measurement can discriminate with certainty between two nonorthogonal quantum states. The challenge for quantum state discrimination is to perform this task as well as is possible.

It is evident that selecting the best possible measurement in any given situation usually requires us to perform a generalized measurement. These are general in the sense that they represent, not just projective measurements of the kind envisaged by von Neumann [22], but rather the most general measurements possible within the confines of quantum theory. The POM formalism is, as we have seen, a remarkable tool in the search for optimal measurements. That this is the case is a consequence of the facts that (i) any set of probability operators satisfying the required properties listed in Section 2 correspond to a possible quantum measurement and (ii) all possible measurements can be described by an appropriate set of probability operators. This means that we can separate the mathematical task of finding the theoretically optimum measurement from the practical one of designing a measurement to implement it.

We have seen that optimal measurements have been found to minimize the error in identifying the state, to discriminate between states unambiguously, and to determine the state with the maximum level of confidence. These similar sounding goals are all subtly different and correspond, for the most part, to quite distinct measurements. We have also discussed yet another task relevant to quantum communications, that of maximizing the information transferred. The problem of state discrimination acquired much of its significance from considering the problem of quantum communication and in particular from quantum cryptography [4–9]. All existing implementations of these are based on optics, and it is perhaps not surprising, therefore, that it is in optics that the experimental advances in quantum state discrimination have been made. We have described, in particular, how quantum-limited measurements have been devised on optical polarization to realize the optimal measurements for detection with minimum error and unambiguous discrimination as well as detection with maximum confidence and maximum mutual information. As quantum information technology develops, the ability to optimize performance by performing the best possible measurements can only become more important.

Acknowledgments

This work was supported, in part, by the UK Engineering and Physical Sciences Research Council (EPSRC), the Royal Society and the Wolfson

Foundation (SMB), by the Synergy fund of the Universities of Glasgow and Strathclyde, and by Perimeter Institute for Theoretical Physics (SC). Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research & Innovation.

References and Notes

1. C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, 1976).
2. A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, 1982).
3. A. S. Holevo, *Statistical Structure of Quantum Theory* (Springer-Verlag, 2000).
4. S. M. Barnett, *Quantum Information* (Oxford U. Press, to be published).
5. S. J. D. Phoenix and P. D. Townsend, "How to beat the code breakers using quantum mechanics," Contemp. Phys. **36**, 165–195 (1995).
6. H.-K. Lo, S. Popescu, and T. Spiller, eds., *Introduction to Quantum Computation and Quantum Information* (World Scientific, 1998).
7. D. Bouwmeester, A. Ekert, and A. Zeilinger, eds., *The Physics of Quantum Information* (Springer-Verlag, 2000).
8. N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, "Quantum cryptography," Rev. Mod. Phys. **74**, 145–195 (2002).
9. S. Loepp and W. K. Wootters, *Protecting Information: from Classical Error Correction to Quantum Cryptography* (Cambridge U. Press, 2006).
10. C. W. Helstrom, "Detection theory and quantum mechanics," Inf. Control. **10**, 254–291 (1967).
11. C. W. Helstrom, "Detection theory and quantum mechanics II," Inf. Control. **13**, 156–171 (1968).
12. A. S. Holevo, "Statistical decision theory for quantum systems," J. Multivariate Anal. **3**, 337–394 (1973).
13. H. P. Yuen, R. S. Kennedy, and M. Lax, "Optimum testing of multiple hypotheses in quantum detection theory," IEEE Trans. Inf. Theory **IT-21**, 125–134 (1975).
14. E. B. Davies, "Information and quantum measurement," IEEE Trans. Inf. Theory **IT-24**, 596–599 (1978).
15. S. M. Barnett, "Quantum information via novel measurements," Philos. Trans. R. Soc. London, Ser. A **355**, 2279–2290 (1997).
16. A. Chefles, "Quantum state discrimination," Contemp. Phys. **41**, 401–424 (2000).
17. S. M. Barnett, "Quantum limited state discrimination," Fortschr. Phys. **49**, 909–913 (2001).
18. S. M. Barnett, "Optical demonstrations of statistical decision theory for quantum systems," Quantum Inf. Comput. **4**, 450–459 (2004).
19. J. A. Bergou, U. Herzog, and M. Hillery, "Discrimination of quantum states," Lect. Notes Phys. **649**, 417–465 (2004).
20. A. Chefles, "Quantum states: discrimination and classical information transmission. A review of experimental progress," Lect. Notes Phys. **649**, 467–511 (Springer, 2004).
21. J. A. Bergou, "Quantum state discrimination and selected applications," J. Phys.: Conf. Ser. **84**, 012001 (2007).

22. J. Von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton U. Press, 1983).
23. A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, 1993).
24. P. Busch, M. Grabowski, and P. Lahti, *Operational Quantum Physics* (Springer, 1995).
25. C. Y. She and H. Heffner, "Simultaneous measurement of noncommuting observables," Phys. Rev. **152**, 1103–1110 (1966).
26. S. Stenholm, "Simultaneous measurement of conjugate variables," Ann. Phys. (N.Y.) **218**, 233–254 (1992).
27. N. G. Walker and J. E. Carroll, "Simultaneous phase and amplitude measurements on optical signals using a multiport junction," Electron. Lett. **20**, 981–983 (1984).
28. N. G. Walker, "Quantum theory of multiport optical homodyning," J. Mod. Opt. **34**, 15–60 (1987).
29. S. M. Barnett and S. Croke, "On the conditions for discrimination between quantum states with minimum error," J. Phys. A **42**, 062001 (2009).
30. A. S. Kholevo, "On asymptotically optimal hypothesis testing in quantum statistics," Theor. Probab. Appl. **23**, 411–415 (1979).
31. L. P. Hughston, R. Jozsa, and W. K. Wootters, "A complete classification of quantum ensembles having a given density matrix," Phys. Lett. A **183**, 14–18 (1993).
32. P. Hausladen and W. K. Wootters, "A 'pretty good' measurement for distinguishing quantum states," J. Mod. Opt. **41**, 2385–2390 (1994).
33. P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W. K. Wootters, "Classical information capacity of a quantum channel," Phys. Rev. A **54**, 1869–1876 (1996).
34. M. Ban, K. Kurokawa, R. Momose, and O. Hirota, "Optimum measurements for discrimination among symmetric states and parameter estimation," IEEE Transl. J. Magn. Jpn. **36**, 1269–1288 (1997).
35. M. Sasaki, K. Kato, M. Izutsu, and O. Hirota, "Quantum channels showing superadditivity in classical capacity," Phys. Rev. A **58**, 146–158 (1998).
36. Y. C. Eldar and G. D. Forney, Jr., "On quantum detection and the square-root measurement," IEEE Trans. Inf. Theory **47**, 858–872 (2001).
37. S. M. Barnett, "Minimum-error discrimination between multiply symmetric states," Phys. Rev. A **64**, 030303 (2001).
38. C.-L. Chou and L. Y. Hsu, "Minimum-error discrimination between symmetric mixed quantum states," Phys. Rev. A **68**, 042305 (2003).
39. Y. C. Eldar, A. Megretski, and G. C. Verghese, "Optimal detection of symmetric mixed quantum states," IEEE Trans. Inf. Theory **50**, 1198–1207 (2004).
40. A. Peres, "Neumark's theorem and quantum inseparability," Found. Phys. **20**, 1441–1453 (1990).
41. C. Mochon, "Family of generalized 'pretty good' measurements and the minimal-error pure-state discrimination problems for which they are optimal," Phys. Rev. A **73**, 032328 (2006).
42. K. Hunter, "Measurement does not always aid state discrimination," Phys. Rev. A **68**, 012306 (2003).
43. E. Andersson, S. M. Barnett, C. R. Gilson, and K. Hunter, "Minimum-error discrimination between three mirror-symmetric states," Phys. Rev. A **65**, 052308 (2002).

44. C.-L. Chou, "Minimum-error discrimination among mirror-symmetric mixed quantum states," Phys. Rev. A **70**, 062316 (2004).
45. K. Hunter, *Optimal Generalised Measurement Strategies*, Ph.D. thesis (University of Strathclyde, 2004).
46. K. Hunter, "Results in optimal discrimination," in *Proceedings of The Seventh International Conference on Quantum Communication, Measurement and Computing (QCMC04)*, Vol. 734 of American Institute of Physics Conference Series (AIP, 2004), pp. 83–86.
47. H. Barnum and E. Knill, "Reversing quantum dynamics with near-optimal quantum and classical fidelity," J. Math. Phys. **43**, 2097–2106 (2002).
48. A. Montanaro, "On the distinguishability of random quantum states," Commun. Math. Phys. **273**, 619–636 (2007).
49. A. Montanaro, "A lower bound on the probability of error in quantum state discrimination," in *IEEE Information Theory Workshop, 2008. ITW '08* (IEEE, 2008), pp. 378–380.
50. D. Qiu, "Minimum-error discrimination between mixed quantum states," Phys. Rev. A **77**, 012328 (2008).
51. M. Ježek, J. Řeháček, and J. Fiurášek, "Finding optimal strategies for minimum-error quantum-state discrimination," Phys. Rev. A **65**, 060301 (2002).
52. Y. C. Eldar, A. Mergretski, and G. C. Verghese, "Designing optimal quantum detectors via semidefinite programming," IEEE Trans. Inf. Theory **49**, 1007–1012 (2003).
53. I. D. Ivanovic, "How to differentiate between non-orthogonal states," Phys. Lett. A **123**, 257–259 (1987).
54. D. Dieks, "Overlap and distinguishability of quantum states," Phys. Lett. A **126**, 303–306 (1988).
55. A. Peres, "How to differentiate between non-orthogonal states," Phys. Lett. A **128**, 19–19 (1988).
56. G. Jaeger and A. Shimony, "Optimal distinction between two non-orthogonal quantum states," Phys. Lett. A **197**, 83–87 (1995).
57. B. Huttner, N. Imoto, N. Gisin, and T. Mor, "Quantum cryptography with coherent states," Phys. Rev. A **51**, 1863–1869 (1995).
58. R. Loudon, *The Quantum Theory of Light* (Oxford U. Press, 2000).
59. A. Chefles, "Unambiguous discrimination between linearly independent quantum states," Phys. Lett. A **239**, 339–347 (1998).
60. A. Peres and D. R. Terno, "Optimal distinction between non-orthogonal quantum states," J. Phys. A **31**, 7105–7111 (1998).
61. A. Chefles and S. M. Barnett, "Optimum unambiguous discrimination between linearly independent symmetric states," Phys. Lett. A **250**, 223–229 (1998).
62. S. Zhang, Y. Feng, X. Sun, and M. Ying, "Upper bound for the success probability of unambiguous discrimination among quantum states," Phys. Rev. A **64**, 062103 (2001).
63. L.-M. Duan and G.-C. Guo, "Probabilistic cloning and identification of linearly independent quantum states," Phys. Rev. Lett. **80**, 4999–5002 (1998).
64. X. Sun, S. Zhang, Y. Feng, and M. Ying, "Mathematical nature of and a family of lower bounds for the success probability of unambiguous discrimination," Phys. Rev. A **65**, 044306 (2002).
65. Y. C. Eldar, "A semidefinite programming approach to optimal

unambiguous discrimination of quantum states," IEEE Trans. Inf. Theory **49**, 446–456 (2003).
66. T. Rudolph, R. W. Spekkens, and P. S. Turner, "Unambiguous discrimination of mixed states," Phys. Rev. A **68**, 010301 (2003).
67. S. M. Barnett, A. Chefles, and I. Jex, "Comparison of two unknown pure quantum states," Phys. Lett. A **307**, 189–195 (2003).
68. S. Zhang and M. Ying, "Set discrimination of quantum states," Phys. Rev. A **65**, 062322 (2002).
69. C. Zhang, G. Wang, and M. Ying, "Discrimination between pure states and mixed states," Phys. Rev. A **75**, 062306 (2007).
70. The support of a mixed state $\hat{\rho}$ is the subspace spanned by its eigenvectors with nonzero eigenvalues. The kernel of a mixed state is the subspace orthogonal to its support.
71. Y. Sun, J. A. Bergou, and M. Hillery, "Optimum unambiguous discrimination between subsets of nonorthogonal quantum states," Phys. Rev. A **66**, 032315 (2002).
72. J. A. Bergou, U. Herzog, and M. Hillery, "Quantum filtering and discrimination between sets of Boolean functions," Phys. Rev. A **90**, 257901 (2003).
73. U. Herzog and J. A. Bergou, "Optimum unambiguous discrimination of two mixed quantum states," Phys. Rev. A **71**, 050301 (2005).
74. P. Raynal and N. Lütkenhaus, "Optimal unambiguous state discrimination of two density matrices: a second class of exact solutions," Phys. Rev. A **76**, 052322 (2007).
75. U. Herzog, "Optimum unambiguous discrimination of two mixed states and application to a class of similar states," Phys. Rev. A **75**, 052309 (2007).
76. P. Raynal, N. Lütkenhaus, and S. J. van Enk, "Reduction theorems for optimal unambiguous state discrimination of density matrices," Phys. Rev. A **68**, 022308 (2003).
77. M. Kleinmann, H. Kampermann, and D. Bruss, "Structural approach to unambiguous discrimination of two mixed states," arXiv.org, arXiv:0803.1083v1 (March 7, 2008).
78. Y. Feng, R. Duan, and M. Ying, "Unambiguous discrimination between mixed quantum states," Phys. Rev. A **70**, 012308 (2004).
79. P. Raynal and N. Lütkenhaus, "Optimal unambiguous state discrimination of two density matrices: lower bound and class of exact solutions," Phys. Rev. A **72**, 022342 (2005).
80. Y. C. Eldar, M. Stojnic, and B. Hassibi, "Optimal quantum detectors for unambiguous detection of mixed states," Phys. Rev. A **69**, 062318 (2004).
81. S. Croke, E. Andersson, S. M. Barnett, C. R. Gilson, and J. Jeffers, "Maximum confidence quantum measurements," Phys. Rev. Lett. **96**, 070401 (2006).
82. R. L. Kosut, I. Walmsley, Y. Eldar, and H. Rabitz, "Quantum state detector design: optimal worst-case *a posteriori* performance," arXiv.org, arXiv:quant-ph/0403150v1 (March 21, 2004).
83. A. Chefles and S. M. Barnett, "Strategies for discriminating between non-orthogonal quantum states," J. Mod. Opt. **45**, 1295–1302 (1998).
84. C.-W. Zhang, C.-F. Li, and G.-C. Guo, "General strategies for discrimination of quantum states," Phys. Lett. A **261**, 25–29 (1999).
85. M. A. P. Touzel, R. B. A. Adamson, and A. M. Steinberg, "Optimal bounded-error strategies for projective measurements in non-orthogonal state discrimination," Phys. Rev. A **76**, 062314 (2007).

86. J. Fiurášek and M. Ježek, "Optimal discrimination of mixed quantum states involving inconclusive results," Phys. Rev. A **67**, 012321 (2003).
87. U. Herzog and J. A. Bergou, "Minimum-error discrimination between subsets of linearly dependent quantum states," Phys. Rev. A **65**, 050305 (2002).
88. C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (U. Illinois Press, 1963).
89. T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, 1991).
90. E. B. Davies, "Information and quantum measurement," IEEE Trans. Inf. Theory **IT-24**, 596–599 (1978).
91. L. B. Levitin, "Optimal quantum measurements for two pure and mixed states," in *Quantum Communications and Measurement*, V. P. Belavkin, O. Hirota, and R. L. Hudson, eds. (Plenum, 1995), pp. 439–448.
92. M. Sasaki, S. M. Barnett, R. Jozsa, M. Osaki, and O. Hirota, "Accessible information and optimal strategies for real symmetrical quantum sources," Phys. Rev. A **59**, 3325–3335 (1999).
93. G. C. Ghirardi, A. Rimini, and T. Weber, "A general argument against superluminal transmission through the quantum mechanical measurement process," Lett. Nuovo Cimento **27**, 293–298 (1980).
94. P. J. Bussey, ""Super-luminal communication in Einstein–Podolsky–Rosen experiments," Phys. Lett. A **90**, 9–12 (1982).
95. T. F. Jordan, "Quantum correlations do not transmit signals," Phys. Lett. A **94**, 264–264 (1983).
96. P. J. Bussey, "Communication and non-communication in Einstein–Rosen experiments," Phys. Lett. A **123**, 1–3 (1987).
97. S. Popescu and D. Rohrlich, "Quantum nonlocality as an axiom," Found. Phys. **24**, 379–385 (1994).
98. L. Masanes, A. Acin, and N. Gisin, "General properties of nonsignaling theories," Phys. Rev. A **73**, 012112 (2006).
99. N. Gisin, "Quantum cloning without signaling," Phys. Lett. A **242**, 1–2 (1998).
100. S. Ghosh, G. Kar, and A. Roy, "Optimal cloning and no signaling," Phys. Lett. A **261**, 17–19 (1999).
101. S. M. Barnett and E. Andersson, "Bound on measurement based on the no-signaling condition," Phys. Rev. A **65**, 044307 (2002).
102. Y. Feng, S. Zhang, R. Duan, and M. Ying, "Lower bound on inconclusive probability of unambiguous discrimination," Phys. Rev. A **66**, 062313 (2002).
103. S. Croke, E. Andersson, and S. M. Barnett, "No-signaling bound on quantum state discrimination," Phys. Rev. A **77**, 012113 (2008).
104. M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge U. Press, 2000).
105. A. E. Siegman, *Lasers* (University Science Books, 1986).
106. G. R. Fowles, *Introduction to Modern Optics* (Holt, Rinehart and Winston, 1975).
107. M. Born and E. Wolf, *Principles of Optics* (Pergamon, 1980).
108. E. Wolf, *Introduction to the Theory of Coherence and Polarization of Light* (Cambridge U. Press, 2007).
109. S. M. Barnett and E. Riis, "Experimental demonstration of polarization discrimination at the Helstrom bound," J. Mod. Opt. **44**, 1061–1064 (1997).

110. R. B. M. Clarke, V. M. Kendon, A. Chefles, S. M. Barnett, and E. Riis, "Experimental realization of optimal detection strategies for overcomplete states," Phys. Rev. A **64**, 012303 (2001).
111. B. Huttner, A. Muller, G. Gautier, H. Zbinden, and N. Gisin, "Unambiguous quantum measurement of nonorthogonal states," Phys. Rev. A **54**, 3783–3789 (1996).
112. R. B. M. Clarke, A. Chefles, S. M. Barnett, and E. Riis, "Experimental demonstration of optimal unambiguous state discrimination," Phys. Rev. A **63**, 040305.
113. M. Mohseni, A. M. Steinberg, and J. A. Bergou, "Optical realization of optimal unambiguous discrimination for pure and mixed quantum states," Phys. Rev. Lett. **93**, 200403 (2004).
114. S. Franke-Arnold, E. Andersson, S. M. Barnett, and S. Stenholm, "Generalized measurements of atomic qubits," Phys. Rev. A **63**, 052301 (2001).
115. P. J. Mosley, S. Croke, I. A. Walmsley, and S. M. Barnett, "Experimental realization of maximum confidence quantum state discrimination for the extraction of quantum information," Phys. Rev. Lett. **97**, 193601 (2006).
116. S. Croke, P. J. Mosley, S. M. Barnett, and I. A. Walmsley, "Maximum confidence measurements and their optical implementation," Eur. Phys. J. D **41**, 589–598 (2007).
117. S. E. Ahnert and M. C. Payne, "General implementation of all possible positive-operator-value measurements of single-photon polarization states," Phys. Rev. A **71**, 012330 (2005).
118. J. Mizuno, M. Fujiwara, M. Akiba, T. Kawanishi, S. M. Barnett, and M. Sasaki, "Optimum detection for extracting maximum information from symmetric qubit sets," Phys. Rev. A **65**, 012315 (2001).