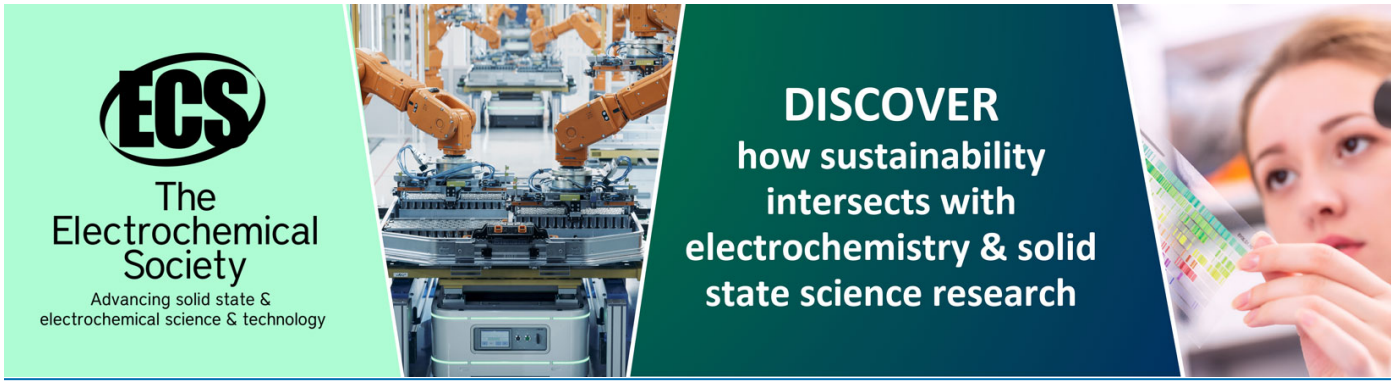
Quantum state discrimination and selected applications



Journal of Physics: Conference Series **84** (2007) 012001 doi:10.1088/1742-6596/84/1/012001

Quantum state discrimination and selected applications

János A Bergou

Department of Physics and Astronomy, Hunter College of The City University of New York, 695 Park Avenue, New York, NY 10021, USA

E-mail: jbergou@hunter.cuny.edu

Abstract. Determining the state of a quantum system is a central task in quantum information processing since it encompasses the read-out problem. Very often the optimized state discrimination strategy is a generalized measurement (Positive Operator Valued Measure, POVM). Therefore, we begin with a brief introduction to the theory of generalized measurements and illustrate the power of the concept on examples relevant to applications in quantum cryptography.

1. Introduction

In quantum information and quantum computing the carrier of information is a quantum system and information is encoded in its state [1]. The state, however, is not an observable in quantum mechanics [2] and, thus, a fundamental problem arises: after processing the information - i.e. after the desired transformation is performed on the input state by the quantum processor - the information has to be read out or, in other words, the state of the system has to be determined. When the set of possible target states is known and the states in the set are mutually orthogonal, this is a relatively simple task. One simply has to set up detectors along these orthogonal directions and a click in one of the detectors will unambiguously identify the state of the system. However, when the possible target states are not orthogonal they cannot be discriminated perfectly, and optimum discrimination with respect to some appropriately chosen criteria is far from being trivial even if the set of the possible nonorthogonal states is known. The problem of discriminating among nonorthogonal states is ubiquitous in quantum information and quantum computing, underlying many of the communication (cryptography) and computing (probabilistic algorithms) schemes that have been suggested so far. It is, in general, a measurement optimization problem (for recent reviews see [3, 4]). It is the purpose of this contribution to introduce the various theoretical and experimental tools that have been developed for discriminating among nonorthogonal quantum states. Interestingly, the field of discriminating among nonorthogonal quantum states was founded by the seminal works of Helstrom [5] and Holevo [6] long before the term Quantum Information Theory was even coined. Stimulated by the rapid developments in quantum information theory of the 90's, the question of how to discriminate between nonorthogonal quantum states *in an optimum way* has gained new momentum and state discrimination quickly became an integral part of quantum information theory. In particular, the suggestion to use nonorthogonal quantum states for communication in secure quantum cryptographic protocols, most notably in the quantum key distribution (QKD) scheme based on the two-state procedure as developed by Bennett [7] (henceforth referred to as the B92 protocol), has given major impetus to this field.

In order to devise an optimum state-discriminating measurement, strategies have been developed with respect to various criteria. Often, these optimized strategies involve generalized measurements (Positive Operator Valued Measures, POVMs). Therefore, in Sec. 2 we begin with a brief overview of the quantum

theory of measurement with special emphasis on POVMs, including their experimental implementation. Then, as an application of the theory, in Sec. 3 we discuss the two most obvious criteria for optimizing a measurement scheme that is designed for discriminating between different states of a quantum system. The two methods are optimum unambiguous discrimination of the states, on the one hand, and state discrimination with minimum error, on the other hand. They will be outlined in detail in the subsections 3.1 and 3.2. As an application of these two main discrimination strategies, optimal figures of merit for the B92 protocol will be reviewed in subsection 3.3. Section 4 is devoted to the recently emerging subfield of discriminating among mixed states with an application which can be considered as a nontrivial generalization of the B92 protocol. We conclude with a brief outlook in Sec. 5.

2. Quantum measurements

Measurements are an integral part of quantum information processing. Reading out the quantum information at the end of the processing pipeline is equivalent to learning what final state the system is in at the output since information is encoded in the state. Since finding out the state of a system can be done only by performing measurements on it, we need a thorough understanding of the quantum theory (and practice) of measurements. To this end we will begin by a brief review of the postulates of standard quantum measurement theory, due essentially to von Neumann. Then, by analyzing the underlying assumptions we will show that some of the postulates can be replaced by more relaxed ones and this will lead us to the concept of generalized measurements (Positive Operator Valued Measures, POVMs) which are particularly useful in measurement optimization problems. Next, by invoking Neumark's theorem we will show how to actually implement POVMs experimentally. These general concepts we will be illustrated in the next section by explicitly working out optimized measurement schemes for various state discrimination strategies.

2.1. Standard quantum measurements

We begin with a brief summary of the postulates of standard quantum measurements. Standard, or projective, measurements were introduced by von Neumann [8] by analyzing a simple model for the coupling between the system to be measured and the meter and by appropriately generalizing the predictions of the model. Deatails of this analysis are omitted here, for the sake of brevity.

Without losing too much of the generality, we assume that the Hilbert space is finite and discrete, in order to keep the notation simple. Then the postulates read as

0. To every observable in quantum mechanics there corresponds a Hermitian operator X which has the spectral representation $X = \sum_j \lambda_j |j\rangle\langle j|$. From the hermiticity of X it follows that the eigenvalues λ_j are real. For simplicity we assume that the eigenvalues are nondegenerate and the corresponding eigenvectors, $\{|j\rangle\}$, form a complete orthonormal basis set.

1. The projectors $P_j = |j\rangle\langle j|$ span the entire Hilbert space, $\sum_j P_j = 1$.

2. From the orthogonality of the states we have $P_i P_j = P_i \delta_{ij}$. In particular, $P_i^2 = P_i$ from where it follows that the eigenvalues of any projector are 0 and 1.

3. A measurement of X yields one of the eigenvalues λ_j.

4. The state of the system after the measurement is $|\phi_j\rangle = \frac{P_j|\psi\rangle}{\sqrt{\langle\psi|P_j|\psi\rangle}}$ if the outcome is λ_j.

5. The probability that this particular outcome is found as the measurement result is $p_j = ||P_j\psi\rangle||^2 = \langle\psi|P_j^2|\psi\rangle = \langle\psi|P_j|\psi\rangle$ where we used the property 2.

6. If we perform the measurement but we do not record the results, the postmeasurement state can be described by the density operator $\rho = \sum_j p_j |\phi_j\rangle\langle\phi_j| = \sum_j P_j|\psi\rangle\langle\psi|P_j$.

These seven postulates adequately describe what happens to the system during the measurement if it was initially in a pure state. If the system is initially in the mixed state ρ the last three postulates are to be replaced by their immediate generalizations:

4a. The state of the system after the measurement is $\rho_j = \frac{P_j \rho P_j}{Tr(P_j \rho P_j)} = \frac{P_j \rho P_j}{Tr(P_j \rho)}$ if the outcome is λ_j.

5a. The probability that this particular outcome is found as the measurement result is $p_j = Tr(P_j \rho P_j) = Tr(P_j^2 \rho) = Tr(P_j \rho)$ where, again, we used the property 2.

6a. If we perform the measurement but we do not record the results, the postmeasurement state can be described by the density operator $\tilde{\rho} = \sum_j p_j \rho_j = \sum_j P_j \rho P_j$.

Of course, $4a$–$6a$ reduce to 4–6 for the pure state density matrix $\rho = |\psi\rangle\langle\psi|$. Therefore, in what follows we use the density matrix to describe a general (pure or mixed) quantum state unless we want to emphasize that the state is pure.

Let us summarize the message of these postulates. They essentially tell us that the measurement process is random, we cannot predict its outcome. What we can predict is the spectrum of the possible outcomes and the probability that a particular outcome is found in an actual measurement. This leads us to the ensemble interpretation of quantum mechanics. The state $|\psi\rangle$ (or ρ for mixed states) describes not a single system but an ensemble of identically prepared systems. If we perform the same measurement on each member of the ensemble we can predict the possible measurement results and the probabilities with which they occur but we cannot predict the outcome of an individual measurement. Except, of course, when the probability of a certain outcome is 0 or 1. With the help of these postulates we can then calculate the moments of the probability distribution, $\{p_j\}$, generated by the measurement. The first moment is the average of a large number of identical measurements performed on the initial ensemble. It is called the expectation value of X and is denoted as $\langle X \rangle$,

$$\langle X \rangle = \sum_j \lambda_j p_j = \sum_j \lambda_j Tr(P_j \rho) = Tr(X\rho) \,, \tag{1}$$

where we used the spectral representation of X. The second moment, $Tr(X^2\rho)$, is related to the variance,

$$\langle (X - \langle X \rangle)^2 \rangle = Tr(X^2\rho) - \langle X \rangle^2 \,. \tag{2}$$

Higher moments can also be calculated in a straightforward manner but typically the first and second moments are the most important ones to consider.

2.2. Positive Operator Valued Measures (POVMs)

Now we are in the position to put the postulates of standard measurement theory under closer scrutiny. What the last three postulates provide us with is, in fact, an algorithm to generate probabilities. The generated probabilities are non-negative, $0 \leq p_j \leq 1$, and the probability distribution is normalized to unity, $\sum_j p_j = 1$ which is a consequence of the first two postulates. Furthermore, the number of possible outcomes is bounded by the number of terms in the orthogonal decomposition of the identity operator of the Hilbert space. Obviously, one cannot have more orthogonal projections than the dimensionality, N_A, of the Hilbert space of the system, so $j \leq N_A$. It would, however, be often desirable to have more outcomes than the dimensionality while keeping the positivity and normalization of the probabilities. We will first show that this is formally possible: if we relax the above rather restrictive postulates and replace them with more flexible ones we can still obtain a meaningful probability generating algorithm. Then we will show that there are physical processes that fit these more general postulates.

Let us begin with the formal considerations and take a closer look at Postulate $5a$ (or 5) which is the one that gives us the prescription for the generation of probabilities. We notice that in order to get a positive probability by this prescription it is sufficient if P_j^2 is a positive operator, we do not need to

require the positivity of an underlying P_j operator. So let us try the following. We introduce a positive operator, $\Pi_j \geq 0$, which is the generalization of P_j^2, and prescribe $p_j = Tr(\Pi_j \rho)$. Of course, we want to ensure that the probability distribution generated by this new prescription is still normalized. Inspecting the postulates we can easily figure out that normalization is a consequence of Postulate 1 and, therefore, require that $\sum_j \Pi_j = I$, that is the positive operators still represent a decomposition of the identity. We will call a decomposition of the identity in terms of positive operators, $\sum_j \Pi_j = I$, a POVM (Positive Operator Valued Measure) and $\Pi_j \geq 0$ the elements of the POVM. These generalizations will form the core of our new postulates $2'$ and $6'$.

As observed in the previous paragraph, for a POVM to exist we do not have to require orthogonality and positivity of the underlying P_j operators. Therefore, the underlying operators that, via Postulates 4 (or $4a$) and 6 (or $6a$), determine the postmeasurement state can be just about any operators, even non-hermitian ones. For projective measurements orthogonality was essentially a consequence of Postulate 2, which was our most constraining postulate because it restricted the number of terms in the decomposition of the identity to at most the dimensionality of the system. Let us now see how far we can get by abandoning it.

If we abandon Postulate 2 then the operators that generate the probability distribution are no longer the same as the ones that generate the postmeasurement states and we have a considerable amount of freedom in choosing them. Let us denote the operators that generate the postmeasurement state by A_j, they are the generalizations of the orthogonal projectors, P_j. In other words, we define the non-normalized postmeasurement state by $A_j|\psi\rangle$ and the corresponding normalized state after the measurement by $|\phi\rangle = A_j|\psi\rangle / \sqrt{\langle\psi|A_j^\dagger A_j|\psi\rangle}$. This expression will form the essence of our new Postulate $4'$. It immediately tells us that Π_j has the structure $\Pi_j = A_j^\dagger A_j$ which by construction is a positive operator. Let us now use our freedom in designing the postmeasurement state. First note that, since the POVM elements are positive operators, $\Pi_j^{1/2}$ exists. Obviously, this is a possible choice for A_j. So is

$$A_j = U_j \Pi_j^{1/2} \,, \tag{3}$$

where U_j is an arbitrary unitary operator. This is the most general form of the detection operators, satisfying $A_j^\dagger A_j = \Pi_j$ and the above expression corresponds to their polar decomposition. What we see is that the POVM elements determine the absolute value operator through $|A_j| = \Pi_j^{1/2}$ but leave its unitary part open. The A_j operators represent a generalization of the projectors P_j whereas Π_j is a generalization of P_j^2. The set $\{A_j\}$ is called the set of detection operators and these operators figure prominently in our new postulates $2', 4'$ and $6'$ replacing the corresponding ones of the standard measurements.

With this we completed our goal that we set out to do at the beginning of this section, namely the generalization of all of the postulates of the standard measurement theory to more flexible ones while keeping the spirit of the old ones. What we see is that Postulate 0. is no longer necessary and rest of the new postulates read as

$1'$. We consider the decomposition of the identity, $\sum_j \Pi_j = 1$, in terms of positive operators, $\Pi_j \geq 0$. Such a decomposition is called a POVM (Positive Operator Valued Measure) and the Π_j the elements of the POVM.

$2'$. The elements of the POVM, Π_j, can be expressed in terms of the detection operators A_j as $\Pi_j = A_j^\dagger A_j$ where, in general, the detection operators are non-hermitian ones, restricted only by the requirement $\sum_j A_j^\dagger A_j = I$. Then, by construction, the POVM elements are positive operators.

$3'$. A detection yields one of the alternatives corresponding to an element of the POVM.

$4'$. The state of the system after the measurement is $|\phi\rangle = A_j|\psi\rangle/\sqrt{\langle\psi|A_j^\dagger A_j|\psi\rangle}$ if it was initially in the pure state $|\psi\rangle$, and $\rho_j = \frac{A_j\rho A_j^\dagger}{Tr(A_j\rho A_j^\dagger)} = \frac{A_j\rho A_j^\dagger}{Tr(A_j^\dagger A_j\rho)}$ if it was initially in the mixed state ρ.

$5'$. The probability that this particular alternative is found as the measurement result is $p_j = Tr(A_j\rho A_j^\dagger) = Tr(A_j^\dagger A_j\rho) = Tr(\Pi_j\rho)$ where we used the cyclic property of the trace operation.

$6'$. If we perform the measurement but we do not record the results, the postmeasurement state is described by the density operator $\tilde{\rho} = \sum_j p_j\rho_j = \sum_j A_j\rho A_j^\dagger$.

Very often we are not concerned with the state of the system after such operation is performed but only with the resulting probability distribution. For this, it is sufficient to consider Postulates $1'$ and $5'$ defining the probability of finding alternative j as the detection result. Note, that at no step did we require the orthogonality of the Π_j's. Since orthogonality is no longer a requirement, the number of terms in this decomposition of the identity is not bounded by N_A. In fact, the number of terms can be arbitrary. Obviously, what we arrived at is a generalization of the von Neumann projective measurement. It is a surprising generalization as it tells us that just about any operation that satisfies Postulates $1'$ and $2'$ is a legitimate operation that generates a valid probability distribution. It is a rather natural generalization of the standard quantum measurement, as it provides us with a well-defined algorithm that generates a well-behaved probability distribution. So this procedure can be regarded as a *generalized measurement* and, indeed, for most purposes it is a sufficient generalization of the standard quantum measurement.

Of course, up to this point all this is just a formal mathematical generalization of the standard quantum measurement. The important question is: How can we implement such a thing physically? In the next section we set out to answer this question and then we will study examples of POVMs.

2.3. Neumark's theorem and the implementation of a POVM via generalized measurements

First, let us take a look at what happens if we couple our system to another system called ancilla, let them evolve, and then measure the ancilla. The Hilbert space of the system is $\mathcal{H}_A$ and the Hilbert space of the ancilla is $\mathcal{H}_B$. We want to gain information about the state of the system that we now denote as $|\psi_A\rangle$. Let $\{|m_B\rangle\}$ be an orthonormal basis for $\mathcal{H}_B$, and U_{AB} a unitary operator acting on $\mathcal{H}_A \otimes \mathcal{H}_B$. The probability p_m of measuring $|m_B\rangle$ is then given by

$$p_m = \| \ (I_A \otimes |m_B\rangle\langle m_B|)U_{AB}(|\psi_A\rangle \otimes |\phi_B\rangle) \ \|^2 \ . \tag{4}$$

Define

$$A_m|\psi_A\rangle \equiv \langle m_B|U_{AB}(|\psi_A\rangle \otimes |\phi_B\rangle) \ . \tag{5}$$

Then A_m is a linear operator on $\mathcal{H}_A$ that depends on $|m_B\rangle$, $|\phi_B\rangle$ and U_{AB}. With the help of this definition we can write the measurement probability as

$$p_m = \| \ A_m|\psi_A\rangle \otimes |m_B\rangle \ \|^2 = \langle\psi_A|A_m^\dagger A_m|\psi_A\rangle \ . \tag{6}$$

Note that

$$\begin{aligned} \sum_m \langle\psi_A|A_m^\dagger A_m|\psi_A\rangle &= \sum_m (\langle\psi_A| \otimes \langle\phi_B|)U_{AB}^\dagger|m_B\rangle\langle m_B|U_{AB}(|\psi_A\rangle \otimes |\phi_B\rangle) \\ &= 1 \ . \end{aligned} \tag{7}$$

Since this is true for any $|\psi_A\rangle$, we must have that

$$\sum_m A_m^\dagger A_m = I_A \ , \tag{8}$$

where I_A is the identity in $\mathcal{H}_A$.

The non-normalized state of the total 'system plus ancilla' after the measurement is $A_m|\psi_A\rangle \otimes |m_B\rangle$ so the (normalized) postmeasurement state of the system alone is

$$|\phi_A\rangle = \frac{1}{\sqrt{\langle\psi_A|A_m^\dagger A_m|\psi_A\rangle}} A_m|\psi_A\rangle \; . \tag{9}$$

The set $\{A_m^\dagger A_m\}$ thus gives a decomposition of the identity in terms of positive operators. Therefore, we can identify them with a POVM. In fact, what we see here is the first half of Neumark's theorem: If we couple our system to an ancilla, let them evolve so that they become entangled, and perform a measurement on the ancilla, which collapses the ancilla to one of the basis vectors of the ancilla space, then this procedure will also transform the system because the ancilla degrees of freedom are now entangled to the system. The transformation of the state of the system is, however, neither unitary nor a projection. It can adequately be described as a POVM so the above procedure corresponds to a POVM in the system Hilbert space. Thus, we have just found a procedure that, when we look at the system only, looks like a POVM. We now know that there are physical processes that can adequately be described as POVMs.

Next we address the question, given the set of operators $\{A_m\}$ acting on $\mathcal{H}_A$ such that $\sum_m A_m^\dagger A_m = I$, can this be interpreted as resulting from a measurement on a larger space? That is, can we find $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, $|\phi_B\rangle$, $\{|m_B\rangle\} \in \mathcal{H}_B$ and U_{AB} acting on $\mathcal{H}$ such that

$$A_m|\psi_A\rangle = \langle m_B|U_{AB}(|\psi_A\rangle \otimes |\phi_B\rangle) \tag{10}$$

holds?

The answer to this question is yes as we will now prove it constructively. Let us choose $\mathcal{H}_B$ to have dimension M and let $\{|m_B\rangle\}$ be an orthonormal basis for $\mathcal{H}_B$, and choose $|\phi_B\rangle$ to be an arbitrary state of $\mathcal{H}_B$. Let us further define a transformation U_{AB} via

$$U_{AB}(|\psi_A\rangle \otimes |\phi_B\rangle) = \sum_m A_m|\psi_A\rangle \otimes |m_B\rangle \; , \tag{11}$$

which implies the Eq. (10). U_{AB} is inner product preserving,

$$(\sum_{m'} \langle\psi'_A| \otimes \langle m'_B A_{m'}^\dagger)(\sum_m A_m|\psi_A\rangle \otimes |m_B\rangle) = \sum_m \langle\psi'_A|A_m^\dagger A_m|\psi_A\rangle = \langle\psi'_A|\psi_A\rangle \; , \tag{12}$$

so it is unitary on the one-dimensional subspace spanned by $|\phi_B\rangle$ and it can be extended to a full unitary operator on $\mathcal{H}_A \otimes \mathcal{H}_B$ because, e.g., on the subspace that is orthogonal to $|\phi_B\rangle$ it can be the identity.

This completes the proof of Neumark's theorem which asserts that there is a one-to-one correspondence between a POVM and the above procedure which sometimes is itself called a generalized measurement. Hence, a generalized measurement can be regarded as the physical implementation of a given POVM. As a final remark to this Section we note that there are two standard methods to extend a Hilbert space. One is the tensor product extension, the other is the direct sum extension. Both methods are employed in practical schemes and they lead to rather different implementations of the same POVM. We will not pursue this issue further here but details can be found in Ref [4], for example.

3. State discrimination strategies and their application to quantum cryptography

As examples of a measurement optimization task we will consider two schemes for the optimal discrimination of quantum states. The first is unambiguous discrimination and the second is discrimination with minimum error. We will see that the optimum measurement for the first strategy is a POVM while the optimum measurement for the second is a standard von Neumann measurement.

The two main discrimination strategies evolved rather differently from the very beginning. Unambiguous discrimination started with pure states and only very recently was it extended to discriminating among mixed quantum states. Minimum-error discrimination addressed the problem of discriminating between two mixed quantum states from the very beginning and the result for two pure states follows as a special case. The two strategies are, in a sense, complementary to each other. Unambiguous discrimination is relatively straightforward to generalize for more than two states, at least in principle, but it is difficult to treat mixed states. The error-minimizing approach, initially developed for two mixed states, is hard to generalize for more than two states. To close this chapter we will analyze the B92 protocol for quantum key distribution (QKD). QKD is the crucial ingredient of most quantum cryptographic protocols and in the B92 proposal all of the concepts of this chapter come together in a particularly clean and instructive form.

3.1. Unambiguous discrimination of two pure states

Unambiguous discrimination is concerned with the following problem. An ensemble of quantum systems is prepared so that each individual system is prepared in one of two known states, $|\psi_1\rangle$ or $|\psi_2\rangle$ with probability η_1 or η_2 (such that $\eta_1 + \eta_2 = 1$), respectively. The preparation probabilities are called a priori probabilities or, simply, priors. The states are, in general, not orthogonal, $\langle\psi_1|\psi_2\rangle \neq 0$ but linearly independent. The preparer, Alice, then draws a system at random from this ensemble and hands it over to an observer, called Bob, whose task is to determine which one of the two states he is given. The observer also knows how the ensemble was prepared, i.e. has full knowledge of the two possible states and their priors but does not know the actual state that was drawn. All he can do is to perform a single measurement or perhaps a POVM on the individual system he receives.

In the unambiguous discrimination strategy the observer is not allowed to make an error, i.e. he is not permitted to conclude that he was given one state when actually he was given the other. First we show that this can not be done with 100% probability of success. To this end, let us assume the contrary and assume we have two detection operators, Π_1 and Π_2, that together span the Hilbert space of the two states,

$$\Pi_1 + \Pi_2 = I \, . \tag{13}$$

For unambiguous detection we also require that

$$\begin{aligned} \Pi_1|\psi_2\rangle &= 0 \, , \\ \Pi_2|\psi_1\rangle &= 0 \, , \end{aligned} \tag{14}$$

so that the first detector never clicks for the second state and vice versa, and we can identify the detector clicks with one of the states unambiguously. The probability of successfully identifying the first state if it is given is $p_1 = \langle\psi_1|\Pi_1|\psi_1\rangle$ and the probability of successfully identifying the second state if it is given is $p_2 = \langle\psi_2|\Pi_2|\psi_2\rangle$. Multiplying (13) with $\langle\psi_1|$ from the left and $|\psi_1\rangle$ from the right and taking into account (14), gives $p_1 = 1$ and, similarly, we obtain $p_2 = 1$, and it appears as though we could have perfect unambiguous discrimination. However, multiplying (13) with $\langle\psi_1|$ from the left and $|\psi_2\rangle$ from the right and taking into account (14) again, gives $0 = \langle\psi_1|\psi_2\rangle$ which can be satisfied for orthogonal states only. In fact, we have just proved that perfect discrimination of nonorthogonal quantum states is not possible.

Equation (13) allows two alternatives only, it assumes that we can have two operators that unambiguously identify the two states all the time. Since this is impossible, we are forced to modify this equation and have to allow for one other alternative. We introduce a third POVM element, Π_0, such that Eq. (14) is still satisfied but (13) is modified to

$$\Pi_1 + \Pi_2 + \Pi_0 = I \, . \tag{15}$$

The first and second POVM elements will continue to unambiguously identify the first and second state, respectively. However, Π_0 can click for both states and, thus, this POVM element corresponds to an

inconclusive detection result. It should be emphasized that this outcome is not an error, we will never identify the first state with the second and vice versa, we simply will not make any conclusion in this case. We can now introduce success and failure probabilities in such a way that $\langle\psi_1|\Pi_1|\psi_1\rangle = p_1$ is the probability of successfully identifying $|\psi_1\rangle$, and $\langle\psi_1|\Pi_0|\psi_1\rangle = q_1$ is the probability of failing to identify $|\psi_1\rangle$, (and similarly for $|\psi_2\rangle$). For unambiguous discrimination we have $\langle\psi_2|\Pi_1|\psi_2\rangle = \langle\psi_1|\Pi_2|\psi_1\rangle = 0$ from (14). Using this, we obtain from Eq. (15) $p_1 + q_1 = p_2 + q_2 = 1$. This means that if we allow inconclusive detection results to occur with a certain probability then in the remaining cases the observer can conclusively determine the state of the individual system.

It is rather easy to see that a simple von Neumann measurement can accomplish this task. Let us denote the Hilbert space of the two given states by $\mathcal{H}$ and introduce the projector P_1 for $|\psi_1\rangle$ and $\bar{P}_1$ for the orthogonal subspace, such that $P_1 + \bar{P}_1 = I$, the identity in $\mathcal{H}$. Then we know for sure that $|\psi_2\rangle$ was prepared if in the measurement of $\{P_1, \bar{P}_1\}$ a click in the $\bar{P}_1$ detector occurs. A similar conclusion for $|\psi_1\rangle$ can be reached with the roles of $|\psi_1\rangle$ and $|\psi_2\rangle$ reversed. Of course, when a click along P_1 (or P_2) occurs then we learn nothing about which state was prepared, this outcome thus corresponding to the inconclusive result. In the von Neumann set-ups one of the alternatives is missing. We either identify one state or we get an inconclusive result but we miss the other state completely. This scenario is actually allowed by (15).

We now turn our attention to the determination of the optimum measurement strategy for unambiguous discrimination. It is the strategy, or measurement set-up, for which the average failure probability is minimum (or, equivalently, the average success probability is maximum). We want to determine the operators in (15) explicitly. If we introduce $|\psi_j^\perp\rangle$ as the vector orthogonal to $|\psi_j\rangle$ ($j = 1, 2$) then the condition of unambiguous detection, Eq. (14), mandates the choices

$$\Pi_1 = c_1|\psi_2^\perp\rangle\langle\psi_2^\perp| \,, \tag{16}$$

and

$$\Pi_2 = c_2|\psi_1^\perp\rangle\langle\psi_1^\perp| \,. \tag{17}$$

Here c_1 and c_2 are positive coefficients to be determined from the condition of optimum.

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

$$\begin{aligned}\Pi_1 &= \frac{p_1}{\sin^2\Theta}|\psi_2^\perp\rangle\langle\psi_2^\perp| \,, \\ \Pi_2 &= \frac{p_2}{\sin^2\Theta}|\psi_1^\perp\rangle\langle\psi_1^\perp| \,. \end{aligned} \tag{18}$$

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

$$\Pi_0 = I - \Pi_1 - \Pi_2 \,. \tag{19}$$

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

$$q_1 q_2 \geq |\langle\psi_1|\psi_2\rangle|^2 \,, \tag{20}$$

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

Eq. (20) represents the constraint imposed by the positivity requirement on the optimum detection operators. The task we set out to solve can now be formulated as follows. Let

$$Q = \eta_1 q_1 + \eta_2 q_2 \tag{21}$$

denote the average failure probability for unambiguous discrimination. We want to minimize this failure probability subject to the constraint, Eq. (20). Due to the relation, $P = \eta_1 p_1 + \eta_2 p_2 = 1 - Q$, the minimum of Q also gives us the maximum probability of success. Clearly, for optimum the product $q_1 q_2$ should be at its minimum allowed by (20), and we can then express q_2 with the help of q_1 as $q_2 = \cos^2\Theta/q_1$. Inserting this expression in (21) yields

$$Q = \eta_1 q_1 + \eta_2 \frac{\cos^2\Theta}{q_1} \,, \tag{22}$$

where q_1 can now be regarded as the independent parameter of the problem. Optimization of Q with respect to q_1 gives $q_1^{POVM} = \sqrt{\eta_2/\eta_1}\cos\Theta$ and $q_2^{POVM} = \sqrt{\eta_1/\eta_2}\cos\Theta$. Finally, substituting these optimal values into Eq. (21) gives the optimum failure probability,

$$Q^{POVM} = 2\sqrt{\eta_1\eta_2}\cos\Theta \,. \tag{23}$$

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

$$Q_1 = \eta_1 + \eta_2 |\langle\psi_1|\psi_2\rangle|^2 \,, \tag{24}$$

since $|\psi_1\rangle$ gives a click with probability 1 in this direction but it is only prepared with probability η_1 and $|\psi_2\rangle$ gives a click with probability $|\langle\psi_1|\psi_2\rangle|^2$ but it is only prepared with probability η_2.

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

$$Q_2 = \eta_1 |\langle\psi_1|\psi_2\rangle|^2 + \eta_2 \,. \tag{25}$$

What we can observe is that Q_1 and Q_2 are given as the arithmetic mean of two terms and Q^{POVM} is the geometric mean of the same two terms for either case. So, one would be tempted to say that the POVM performs better always. This, however, is not quite the case, it does so only when it exists. The obvious condition for the POVM solution to exist is that both $q_1^{POVM} \leq 1$ and $q_2^{POVM} \leq 1$. Using $\eta_2 = 1 - \eta_1$, a little algebra tells us that the POVM exists in the range $\cos^2\Theta/(1+\cos^2\Theta) \leq \eta_1 \leq 1/(1+\cos^2\Theta)$. If η_1 is smaller than the lower boundary, the POVM goes over to the first von Neumann measurement and if η_1 exceeds the upper boundary the POVM goes over to the second von Neumann measurement. This can be easily seen from Eqs. (18) and (19) since $p_1 = 1 - q_1 = 0$ for $q_1 = 1$ and Π_0 becomes a projection along $|\psi_1\rangle$ (and correspondingly for $p_2 = 0$).

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

$$Q^{opt} = \begin{cases} Q^{POVM} & \text{if } \frac{\cos^2\Theta}{1+\cos^2\Theta} \leq \eta_1 \leq \frac{1}{1+\cos^2\Theta} \,, \\ Q_1 & \text{if } \eta_1 < \frac{\cos^2\Theta}{1+\cos^2\Theta} \,, \\ Q_2 & \text{if } \frac{1}{1+\cos^2\Theta} < \eta_1 \,. \end{cases} \tag{26}$$

The optimum detection operators are given by

$$\begin{aligned} \Pi_1 &= \frac{1 - q_1^{opt}}{\sin^2\Theta} |\psi_2^\perp\rangle\langle\psi_2^\perp| \,, \\ \Pi_2 &= \frac{1 - q_2^{opt}}{\sin^2\Theta} |\psi_1^\perp\rangle\langle\psi_2^\perp| \,. \end{aligned} \tag{27}$$

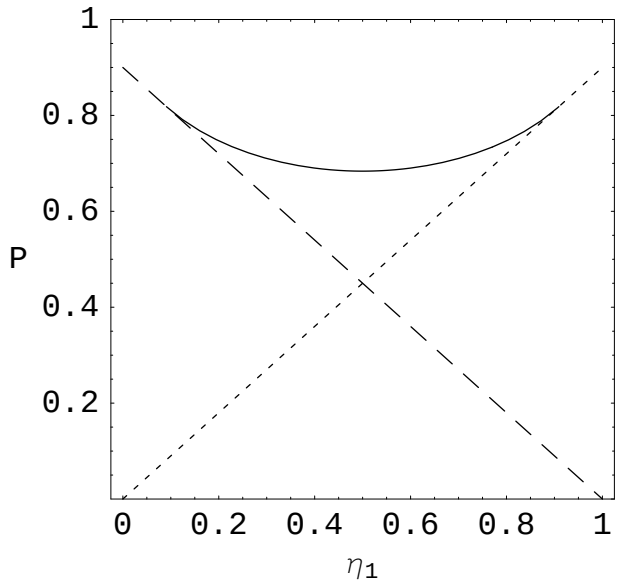


Figure 1. Succes probability, $P = 1 - Q$, vs. the prior probability, η_1. Dashed line: $P_1 = 1 - Q_1$, dotted line: $P_2 = 1 - Q_2$, solid line: $P^{POVM} = 1 - Q^{POVM}$. For the figure we used the following representative value: $|\langle\psi_1|\psi_2\rangle|^2 = 0.1$. For this the optimal success probability, $P_{opt} = 1 - Q_{opt}$ is given by $P_1 = 1 - Q_1$ for $0 < \eta_1 < 0.09$, by $P^{POVM} = 1 - Q^{POVM}$ for $0.09 \leq \eta_1 \leq 0.9$ and by $P_2 = 1 - Q_2$ for $0.9 < \eta_1$.

These expressions show explicitly that $\Pi_1 = 0$ and Π_2 is the projector $|\psi_1^{\perp}\rangle\langle\psi_1^{\perp}|$ when $q_1^{opt} = 1$ and $q_2^{opt} = \cos^2\Theta$, i.e. the POVM goes over smoothly into a projective measurement at the lower boundary and, similarly, into the other von Neumann projective measurement at the upper boundary.

Fig. 1 displays the failure probabilities, Q_1, Q_2, and Q^{POVM} vs. η_1 for a fixed value of the overlap, $\cos^2\Theta$.

The above result is very satisfying from a physical point of view. The POVM delivers a lower failure probability in its entire range of existence than either of the two von Neumann measurements. At the boundaries of this range it merges smoothly with the one von Neumann measurement that has a lower failure probability at that point. Outside this range the state preparation is dominated by one of the states and the optimal measurement becomes a von Neumann projective measurement, using the state that is prepared less frequently as its failure direction. It should be noted that unambiguous discrimination was first suggested by Ivanovic [9], the POVM leading to Eq. (23) when $\eta_1 = \eta_2 = 1/2$ was found by Dieks [10] and Peres proved its optimality [11], therefore it is called the IDP limit in this case. For arbitrary prior probabilities Jaeger and Shimony derived Eq. (23) [12].

3.2. Minimum error discrimination of two quantum states

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

Let us state the optimization problem a little more precisely. In the most general case, we want to distinguish, with minimum probability of error, among N given states of a quantum system (where $N \geq 2$). The states are given by the density operators ρ_j $(j = 1, 2, \ldots, N)$ and the j^{th} state occurs with the given a priori probability η_j, such that $\sum_{j=1}^{N} \eta_j = 1$. The measurement can be formally described with the help of a POVM, where the POVM elements, Π_j, correspond to the possible measurement outcomes. They are defined in such a way that $\mathrm{Tr}(\rho\Pi_j)$ is the probability to infer the state of the system to be ρ_j if it has been prepared in a state ρ. Since the probability is a real non-negative number, the detection operators once again have to be positive-semidefinite. In the error-minimizing measurement scheme the measurement is required to be exhaustive and conclusive in the sense that in each single case one of the N possible states is identified with certainty and inconclusive results do not occur. This leads to the requirement

$$\sum_{j=1}^{N} \Pi_j = I_{D_S}, \tag{28}$$

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

$$P_{\text{err}} = 1 - P_{\text{corr}} = 1 - \sum_{j=1}^{N} \eta_j \mathrm{Tr}(\rho_j \Pi_j) \tag{29}$$

with $\sum_j \eta_j = 1$. Here we introduced the probability P_{corr} that the guess is correct. In order to find the minimum-error measurement strategy, one has to determine the POVM that minimizes the value of P_{err} under the constraint given by Eq. (28). By inserting these optimum detection operators into Eq. (29), the minimum error probability $P_{\text{err}}^{\min} \equiv P_E$ is determined. The explicit solution to the error-minimizing problem is not trivial and analytical expressions have been derived only for a few special cases.

For the case that only two states are given, either pure or mixed, the minimum error probability, P_E, was derived in the mid 70s by Helstrom in the framework of quantum detection and estimation theory. We find it more instructive to start by analyzing the two-state minimum-error measurement with the help of an alternative method that allows us to gain immediate insight into the structure of the optimum detection operators, without applying variational techniques. Starting from Eq. (29) and making use of the relations $\eta_1 + \eta_2 = 1$ and $\Pi_1 + \Pi_2 = I_{D_S}$ that have to be fulfilled by the a priori probabilities and the detection operators, respectively, we see that the total probability to get an erroneous result in the measurement is given by

$$P_{\text{err}} = 1 - \sum_{j=1}^{2} \eta_j \mathrm{Tr}(\rho_j \Pi_j) = \eta_1 \mathrm{Tr}(\rho_1 \Pi_2) + \eta_2 \mathrm{Tr}(\rho_2 \Pi_1). \tag{30}$$

This can be alternatively expressed as

$$P_{\text{err}} = \eta_1 + \mathrm{Tr}(\Lambda \Pi_1) = \eta_2 - \mathrm{Tr}(\Lambda \Pi_2), \tag{31}$$

where we introduced the Hermitian operator

$$\Lambda = \eta_2 \rho_2 - \eta_1 \rho_1 = \sum_{k=1}^{D_S} \lambda_k |\phi_k\rangle\langle\phi_k|. \tag{32}$$

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

$$\begin{aligned} \lambda_k &< 0 \qquad \text{for} \qquad 1 \leq k < k_0, \\ \lambda_k &> 0 \qquad \text{for} \qquad k_0 \leq k \leq D, \\ \lambda_k &= 0 \qquad \text{for} \qquad D < k \leq D_S. \end{aligned} \tag{33}$$

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

$$P_{\rm err} = \eta_1 + \sum_{k=1}^{D_S} \lambda_k \langle\phi_k|\Pi_1|\phi_k\rangle = \eta_2 - \sum_{k=1}^{D_S} \lambda_k \langle\phi_k|\Pi_2|\phi_k\rangle. \tag{34}$$

Our optimization task now consists in determining the specific operators Π_1, or Π_2, respectively, that minimize the right-hand side of Eq. (34) under the constraint that

$$0 \leq \langle\phi_k|\Pi_j|\phi_k\rangle \leq 1 \qquad (j = 1, 2) \tag{35}$$

for all eigenstates $|\phi_k\rangle$. The latter requirement is due to the fact that $\mathrm{Tr}(\rho\Pi_j)$ denotes a probability for any ρ. From this constraint and from Eq. (34) it immediately follows that the smallest possible error probability, $P_{\rm err}^{\rm min} \equiv P_E$, is achieved when the detection operators are chosen in such a way that the equations $\langle\phi_k|\Pi_1|\phi_k\rangle = 1$ and $\langle\phi_k|\Pi_2|\phi_k\rangle = 0$ are fulfilled for eigenstates belonging to negative eigenvalues, while eigenstates corresponding to positive eigenvalues obey the equations $\langle\phi_k|\Pi_1|\phi_k\rangle = 0$ and $\langle\phi_k|\Pi_2|\phi_k\rangle = 1$. Hence the optimum detection operators can be written as

$$\Pi_1 = \sum_{k=1}^{k_0-1} |\phi_k\rangle\langle\phi_k|, \qquad \Pi_2 = \sum_{k=k_0}^{D_S} |\phi_k\rangle\langle\phi_k|, \tag{36}$$

where the expression for Π_2 has been supplemented by projection operators onto eigenstates belonging to the eigenvalue $\lambda_k = 0$, in such a way that $\Pi_1 + \Pi_2 = I_{D_S}$. Obviously, provided that there are positive as well as negative eigenvalues in the spectral decomposition of Λ, the minimum-error measurement for discriminating two quantum states is a von Neumann measurement that consists in performing projections onto the two orthogonal subspaces spanned by the set of states $\{|\phi_1\rangle, \ldots, |\phi_{k_0-1}\rangle\}$, on the one hand, and $\{|\phi_{k_0}\rangle, \ldots, |\phi_{D_S}\rangle\}$, on the other hand. An interesting special case arises when negative eigenvalues do not exist. In this case it follows that $\Pi_1 = 0$ and $\Pi_2 = I_{D_S}$ which means that the minimum error probability can be achieved by always guessing the quantum system to be in the state ρ_2, without performing any measurement at all. Similar considerations hold true in the absence of positive eigenvalues so a measurement does not always aid minimum-error discrimination. By inserting the optimum detection operators into Eq. (31) the minimum error probability is found to be

$$P_E = \eta_1 - \sum_{k=1}^{k_0-1} |\lambda_k| = \eta_2 - \sum_{k=k_0}^{D} |\lambda_k|. \tag{37}$$

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

$$P_E = \frac{1}{2}\left(1 - \sum_k |\lambda_k|\right) = \frac{1}{2}\left(1 - \mathrm{Tr}|\Lambda|\right), \tag{38}$$

where $|\Lambda| = \sqrt{\Lambda^\dagger\Lambda}$. Together with Eq. (29) this immediately yields the well-known Helstrom formula for the minimum error probability in discriminating ρ_1 and ρ_2,

$$P_E = \frac{1}{2}\left(1 - \mathrm{Tr}|\eta_2\rho_2 - \eta_1\rho_1|\right) = \frac{1}{2}\left(1 - \|\eta_2\rho_2 - \eta_1\rho_1\|\right). \tag{39}$$

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

$$P_E = \frac{1}{2}\left(1 - \sqrt{1 - 4\eta_1\eta_2|\langle\psi_1|\psi_2\rangle|^2}\right). \tag{40}$$

This expression, which is the one found in textbooks, can be cast to the equivalent form,

$$P_E = \eta_{min}\left(1 - \frac{2\eta_{max}(1-|\langle\psi_1|\psi_2\rangle|^2)}{\eta_{max}-\eta_{min}+\sqrt{1-4\eta_{min}\eta_{max}|\langle\psi_1|\psi_2\rangle|^2}}\right) , \tag{41}$$

where η_{min} (η_{max}) is the smaller (greater) of the prior probabilities, η_1 and η_2. This form lends itself to a transparent interpretation. The first factor on the right-hand-side is what we would get if we always guessed the state that is prepared more often, without any measurement at all. Thus, the factor multiplying η_{min} is the result of the optimized measurement.

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

As we mentioned already in the introduction, the result Eq. (39) was obtained independently by Helstrom [5] and Holevo [6] and is commonly referred to as the Helstrom bound. The derivation presented here follows that of Ref. [13].

Finally, we present an interesting relation, without proof, that is always satisfied between the minimum-error probability of the minimum-error detection and the optimal failure probability of unambiguous detection [13]. It reads as

$$P_E \leq \frac{1}{2}Q^{opt} . \tag{42}$$

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

3.3. The B92 quantum key distribution protocol

Both of the state discrimination strategies discussed in the previous subsections come very nicely together in the so called B92 quantum key distribution (QKD) protocol.

In cryptography the sender is often called Alice and the receiver Bob. We will use this nomenclature in what follows. The basic task in cryptography can then be formulated as this. Alice wants to send a message to Bob, and keep it secret from everybody else. To accomplish this, she uses a code to encrypt the message. A possible code is to shift each letter of the message by a different amount. In this case Alice and Bob must share a sequence of numbers telling them the shifts. This is the key. If the key is random and is used only once the code is unbreakable. The question is, then, how to generate a secret key?

In 1992 Bennett proposed using the unambiguous discrimination of two nonorthogonal states as the basis of a form of quantum cryptography [7]. Quantum cryptography is a method of generating a secure shared key by quantum mechanical means that is discarded after being used only once. So, it is the quantum version of the one-time pad cipher and here is how it works.

(i) Alice sends a system that is prepared with equal probability either in the state $|\psi_0\rangle = |0\rangle$, which corresponds to the logical 0, or in the state $|\psi_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, which corresponds to the logical 1.

(ii) Bob applies optimum unambiguous state discrimination strategy to the state he receives. Using Eq. (23), the success probability for Bob's measurement is $P = 1 - Q^{POVM} = 1 - \frac{1}{\sqrt{2}}$.

(iii) Bob tells Alice, over a public classical channel, whether the discrimination succeeded or failed. They keep the bit if the discrimination was successful, throw it away if it failed.

After repeating this procedure many times Alice and Bob share a sequence of 0's and 1's that they can use for a key.

Why is this procedure secure? Suppose an eavesdropper, called Eve, has intercepted the particle. She cannot determine whether it is in $|\psi_0\rangle$ or $|\psi_1\rangle$. One thing she can do is to apply the optimum unambiguous state discrimination procedure. Then she will fail with a probability of $\frac{1}{\sqrt{2}} \approx 71\%$. When she does, she has no idea what state was sent, so she must guess which one to send to Bob. Since the two states are prepared with equal probability, Eve will guess half the time right and half the time wrong. This means that the probability that Bob will receive a wrong bit is $\frac{1}{2\sqrt{2}} \approx 35.5\%$. These errors can easily be detected if Alice and Bob add one more step to their protocol.

(iv) Alice and Bob publicly compare some of their bits. If there are no errors there is no eavesdropper and they keep the remaining bits. If there are errors, in the range of 35%, there is likely to be an eavesdropper. They then simply throw out all bits and try again.

However, Eve can do better. Her goal, besides learning as much as possible about the key, is also to introduce as few errors as possible. There are unavoidable errors in any communication scheme, partly due to the imperfections of the communication channel and partly due to the imperfect detection. Eve's goal is to remain below this unavoidable noise level in order to avoid being detected. So, suppose she has intercepted the particle but she now chooses the minimum error strategy to determine which state was sent. Using Eq. (40), her error rate will now be $\frac{1}{2}(1 - \frac{1}{\sqrt{2}}) \approx 14.6\%$ which is much less than the error rate that she introduces if she uses the unambiguous discrimination strategy. In addition, she still learns the key with a fidelity of about 85%. However, even this rather low error rate can still be detected if Alice and Bob modify the last step of their protocol.

(iv′) Alice and Bob publicly compare some of their bits. If there are no errors there is no eavesdropper and they keep the remaining bits. If there are errors, in the range of 14%, there is likely to be an eavesdropper. They then simply throw out all bits and try again.

This requirement is much more stringent than the one in Step (iv) of the original protocol. It is still possible to detect the presence of an eavesdropper but the requirements on the channel quality and detector efficiency are much more demanding than in the case when Eve uses the same strategy as Bob. So, here we had an example where one state discrimination strategy is optimal for the intended recipient and the other for the eavesdropper and to analyze the worst case scenario for Alice and Bob we have to consider all of their possibilities. There are many other QKD protocols but this one is perhaps the clearest example of how important optimal detection strategies are for quantum communication.

4. Recent developments

It is interesting to note that the two main discrimination strategies have evolved very differently from the beginning. Minimum error discrimination (also known as hypothesis testing or best guess discrimination) considered pure and mixed states on equal footing from its introduction. However, it is very difficult to generalize this strategy to more than two states, except for some very special highly symmetrical cases. Unambiguous discrimination, on the other hand, dealt with pure states for a long time after it was introduced. It is precisely the area of the unambiguous discrimination of mixed states that has seen a significant amount of progress recently.

Optimum unambiguous discrimination between two mixed states is an issue of ongoing theoretical research [14–20]. In contrast to minimum-error discrimination, there does not exist a compact formula expressing the minimum probability of inconclusive results, i. e. the minimum failure probability, for unambiguously discriminating two mixed states that are completely arbitrary. However, analytical solutions can be obtained for certain special classes of density operators, including the cases that are of interest for this paper. In a seminal paper Rudolph, Spekkens and Turner [14] clearly established the principles behind the unambiguous discrimination of mixed states. In order to understand their reasoning, we need to introduce some terminology. We call the support of a mixed state density operator the subspace spanned by its eigenvectors belonging to nonzero eigenvalues. The kernel of the mixed

state density operator is the orthogonal complement to its the support. Any measurement in the kernel of a state with a positive outcome unambiguously identifies the other state. Of course, this is not necessarily optimal. In the same paper it has also been shown that the fidelity of the two density operators represents an absolute lower bound for the failure probability of the unambiguous discrimination. Built on these results we have presented exact analytical expressions for an important special case in [20] and found conditions for when the fidelity bound can be saturated.

These results can be directly applied to a novel Quantum Key Distribution (QKD) scheme that can be regarded as a generalization of the B92 protocol, discussed in Sec. 3.3. We have recently proposed such a scheme [21]. It is a QKD scheme based on communicating via quantum patterns rather than via known pure sates. In this scheme Bob's ability to distill a noise-free string of qubits, shared also by Alice, is virtually unaffected but the ability of the eavesdropper, Eve, to learn the proper sequence is vastly diminished as compared to the B92 protocol. In particular, the minimum amount of noise that Eve is bound to introduce goes from the 14.7% of the B92 protocol to about 35.5% in the present one, providing a dramatic increase in security. Although this figure of merit is still coming short of that of the theoretical optimum of 50%, the scheme offers additional advantages. For example, it does not require a shared reference frame between Alice and Bob as it uses only the symmetry features of the patterns for communication and, furthermore, is robust against unitary errors. QKD protocols that similarly do not require a shared reference frame between Alice and Bob and are robust against unitary errors have been proposed recently but they require much more resources and offer a lower figure of merit [22]. The obvious advantages make our scheme secure and attractive from a practical point of view.

Our proposal is based on the programmable state discriminator for the unambiguous discrimination between unknown quantum states [23]. Prior to this device it was widely held that only *known* states can be discriminated. In this device the *unknown* states are provided as program (qubits A and B) and a copy of one of the unknown states as data (qubit C) at the input. More specifically, Alice prepares one of the following three-qubit states

$$\begin{aligned} |\Phi_0\rangle &= |\psi_1\rangle_A|\psi_2\rangle_B|\psi_1\rangle_C \,, \\ |\Phi_1\rangle &= |\psi_1\rangle_A|\psi_2\rangle_B|\psi_2\rangle_C \,, \end{aligned} \tag{43}$$

at random. Here $|\psi_1\rangle$ and $|\psi_2\rangle$ can be completely arbitrary qubit states,

$$|\psi_1\rangle = \cos(\theta_1/2)|0\rangle + e^{i\phi_1}\sin(\theta_1/2)|1\rangle \,, \tag{44}$$

$$|\psi_2\rangle = \cos(\theta_2/2)|0\rangle + e^{i\phi_2}\sin(\theta_2/2)|1\rangle \,, \tag{45}$$

that remain unknown to both Alice and Bob throughout the entire discrimination process. In fact, they can even change from one preparation to the next. What matters is that the pattern remains the same: the state of the last qubit matches the state of the first qubit, as in $|\Phi_0\rangle$, or it matches the state of the second qubit, as in $|\Phi_1\rangle$. Furthermore, we assume that both inputs, $|\Phi_0\rangle$ and $|\Phi_1\rangle$, are prepared with probability 1/2. Alice and Bob are never required to learn the states $|\psi_1\rangle$ and $|\psi_2\rangle$. For the transmission and reception all that matters is the comparison of the last qubits to the other two.

The device then performs an optimum unambiguous discrimination of the two inputs given in (43). Details of how this is done can be found in [23] and [24]. Here we just recall the main results which were obtained by exploiting obvious symmetry features of the two inputs. Namely, the first one is symmetric in qubits A and C, while the second one is symmetric in qubits B and C. If we set up a projective measurement along the antisymmetric subspace of qubits A and C (and the orthogonal subspace) a click in the detector along the antisymmetric subspace will unambiguously identify the input as $|\Phi_0\rangle$ since $|\Phi_1\rangle$ has no component in this subspace. Similarly, a detector along the antisymmetric subspace of B and C unambiguously identifies the input as $|\Phi_1\rangle$ since $|\Phi_0\rangle$ has no component in this subspace. These two antisymmetric subspaces are not orthogonal, therefore one needs again two different detector settings. One that projects on the antisymmetric subspace of A and C and the orthogonal subspace, and another that projects on the antisymmetric subspace of B and C and the orthogonal subspace. Bob again chooses

randomly between these two detector settings, that is performs a flip-flop measurement. The success probability of such a measurement is 1/8 when averaged over all possible input states [23]. Once again, however, the optimum measurement is a POVM with an average probability of success of 1/6 over all possible qubits $|\psi_1\rangle$ and $|\psi_2\rangle$ transmitted by Alice [23].

An eavesdropper, Eve, could, in principle, perform the same measurement as Bob. Her success rate would then be 1/6 and in the remaining 5/6 of the cases she would have to guess the input pattern. If half the time she guesses right she will introduce an error rate of $(5/6)/2 = 5/12 = 0.417$. This is large, so her presence could be easily detected. However, once again, it is more advantageous for Eve to perform a minimum error identification. The minimum error probability for these input patterns was found in [24]. In the case when the preparation of the two patterns is equally likely, the minimum error probability is $\frac{1}{2}(1 - \sqrt{3}/6) = 0.356$. This is much higher than the corresponding figure of merit for the B92 protocol, by a factor of about 2.5. While it is still below the theoretical optimum of 50%, this figure is high enough for generating practical interest [25].

The scheme has further advantages over existing protocols. It utilizes symmetry only and a successful identification of the input patterns corresponds to a successful projection onto one of the antisymmetric subspaces of the two inputs. This subspace is invariant, it is the same in any basis. Therefore, there is no need for a shared reference frame between the two communicating parties. The scheme is also robust against unitary errors that affect the two inputs in such a way that the patterns are still preserved. This happens, for example, if all qubits undergo the same local unitary transformation as this leaves the input patterns invariant. In the following Table we summarize the features of the present case and compare them to the benchmarks of the B92 protocol.

Table 1. Comparison of the current proposal to the benchmarks of the B92 protocol.

	B92	QKD via patterns
Bob's success probability	0.29	0.17
Eve's minimal error rate	0.147	0.355
Shared reference frame	Yes	No
Robust against unitary errors	No	Yes

Obviously, in spite of the fact that we use more qubits (three times the number of quantum resources), the most attractive features of the current proposal are those in lines 3 and 4 of the Table: the high error rate Eve is bound to introduce even if she uses the minimum error discrimination strategy, and no need for Alice and Bob to share a reference frame.

Another important consideration is the security of the scheme. To apply known security criteria we recall that it has been shown in a series of recent papers that the problem of discrimination of unknown pure quantum states is equivalent to the discrimination of known mixed states [24, 26, 27]. The two density operators, corresponding to the pure state inputs in (43) averaged over the Bloch spheres of the independent qubits, i.e. over θ_1, ϕ_1 and θ_2, ϕ_2 of (45), have been given in Ref. [24] and they satisfy the security criteria that were introduced in [28] for the case when quantum key distribution in the B92 protocol employs mixed states. Eve can get no information about the patterns without disturbing them in an essential way. Assuming perfect detectors and noise free environment the scheme is unconditionally secure. On the experimental side, optimal unambiguous discrimination has been realized optically both for pure and mixed states, in excellent agreement with the theory [29].

Thus, the scheme represents a dramatic improvement over the original B92 protocol and incorporates many desirable aspects of existing, more practical protocols. First, the ability of the eavesdropper to

obtain information is greatly reduced as compared to the original B92 protocol. It is much harder for Eve to hide in the noise that is inevitably introduced by imperfect detectors and communication channel losses. Second, the identification of the states is based on symmetry and, therefore, no shared reference frame is required for Alice and Bob to communicate. Third, the patterns are more robust against unitary errors than the patterns based on communicating via known states. For example, a simultaneous rotation of all three qubits carrying the pattern into which the 0 and 1 are encoded still preserves the pattern. Finally, using the equivalence of the pattern of unknown qubit states to that of known mixed states we were able to apply the security conditions derived in [28] to our case and show that the current scheme is unconditionally secure mathematically while it is a practically secure scheme if noise is included.

5. Summary and outlook

In this paper we have given a tutorial introduction to the theory of generalized measurements (Positive Operator Valued Measures, POVMs). In our treatment the POVM emerges as a natural and for most practical purposes completely satisfactory generalization of the standard projector valued quantum measurement, originally introduced by von Neumann. Based on Neumark's Theorem, we have shown a general method to implement POVMs experimentally.

Next, we have shown that optimal state discrimination problems are equivalent to finding measurement strategies that optimize some reasonably chosen figure of merit. The solution often leads to POVMs. As a first application we have analyzed the performance of the B92 quantum key distribution protocol and have shown that the optimal measurement strategies for Bob, the intended recipient of the key and Eve, the eavesdropper, are different.

Finally, based on recent progress in the discrimination of mixed quantum states, we have discussed an extension of the B92 protocol to communicating the key via symmetry alone. While the proposed new protocol does not significantly affect Bob's ability to distill a shared key, it drastically increases the noise introduced by the eavesdropper, thus making her presence easier to detect.

In summary, state discrimination is a rapidly developing subfield of quantum information theory, touching the very foundation of quantum mechanics. Perhaps the most immediate open problem is the unambiguous discrimination of two mixed states. Although the solution is known for many special cases, no closed form solution is available for two general mixed states. It is fully expected that the optimal unambiguous discrimination of two rank 2 density matrices will be solved in the near future. However, no analytical expression is expected for higher rank problems due to the fact that the dependence on the relevant parameters is highly nonlinear. On the other hand, the general case is an example of a semidefinite programming problem and as such, very efficient numerical methods can be employed to find a numerically optimized solution to the underlying measurement problem.

Acknowledgments

This research was partially supported by a grant from the Humboldt Foundation and by PSC-CUNY. Financial support of FONDECYT (Cooperacion Internacional 7050101) during a stay at the Catholic University of Santiago is gratefully acknowledged. The author is grateful to Ulrike Herzog, E. Feldman, M. Hillery and M. Orszag for many helpful discussions and fruitful collaborations.

References

[1] Nielsen M A and Chuang I L 2000 *Quantum Computation and Information* (Cambridge: Cambridge University Press)
[2] Peres A 1995 *Quantum Theory: Concepts and Methods* (Amsterdam: Kluwer)
[3] Chefles A 2000 *Contemporary Physics* **41** 401
[4] Bergou J A, Herzog U and Hillery M 2004 *Quantum State Estimation* (Lect. Notes Phys. vol 649) ed M Paris and J Řehaček (Berlin: Springer) chapter 11 pp 417-465
[5] Helstrom C W 1976 *Quantum Detection and Estimation Theory* (New York: Academic Press)
[6] Holevo A S 1979 *Theory Probab. Appl.* **23** 411
[7] Bennett C H 1992 *Phys. Rev. Lett.* **68** 3121
[8] von Neumann J 1955 *The Mathematical Foundations of Quantum Mechanics* (Princeton: Princeton University Press)

[9] Ivanovic I D 1987 *Phys. Lett. A* **123** 257
[10] Dieks D 1988 *Phys. Lett. A* **126** 303
[11] Peres A 1988 *Phys. Lett. A* **128** 19
[12] Jaeger G and Shimony A 1995 *Phys. Lett. A* **197** 83
[13] Herzog U and Bergou J 2004 *Phys. Rev. A* **70** 022302
[14] Rudolph T, Spekkens R W and Turner P S 2003 *Phys. Rev. A* **68** 010301(R)
[15] Raynal Ph Lütkenhaus N and van Enk S J 2003 *Phys. Rev. A* **68** 022308
[16] Feng Y, Duan R and Ying M 2004 *Phys. Rev. A* **70** 012308
[17] Eldar Y C, Stojnic M and Hassibi B 2004 *Phys. Rev. A* **69** 062318
[18] Herzog U and Bergou J A 2005 *Phys. Rev. A* **71** 050301(R)
[19] Raynal Ph and Lütkenhaus N 2005 *Phys. Rev. A* **72** 022342
[20] Bergou J A, Feldman E and Hillery M 2006 *Phys. Rev.A* **73** 032107
[21] Bergou J and Orszag M 2007 *in preparation*
[22] Boileau J C, Gottesman D, Laflamme R, Poulin D and Spekkens R W 2004 *Phys. Rev. Lett.* **92** 017901
[23] Bergou J and Hillery M 2005 *Phys. Rev. Lett.* **94** 160501
[24] Bergou J A, Bužek V, Feldman E, Herzog U and Hillery M 2006 *Phys. Rev. A* **73** 062334
[25] Gisin N, Ribordy G, Tittel W and Zbinden H 2002 *Rev. Mod. Phys.* **73** 145
[26] Hayashi A, Horibe M and Hashimoto T 2005 *Phys. Rev. A* **72** 052306
[27] Hayashi A, Horibe M and Hashimoto T 2006 *Phys. Rev. A* **73** 012328
[28] Koashi M and Imoto M 1996 *Phys. Rev. Lett.* **77** 2137
[29] Mohseni M, Steinberg A M and Bergou J A 2004 *Phys. Rev. Lett.* **93** 200403