Optimal unambiguous filtering of a quantum state: An instance in mixed state discrimination

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

[1]*Department of Physics, Hunter College, City University of New York, 695 Park Avenue, New York, New York 10021, USA*
[2]*Institut für Physik, Humboldt-Universität zu Berlin, Newtonstrasse 15, D-12489 Berlin, Germany*
(Received 17 December 2004; published 11 April 2005)

Deterministic discrimination of nonorthogonal states is forbidden by quantum measurement theory. However, if we do not want to succeed all the time, i.e., allow for inconclusive outcomes to occur, then unambiguous discrimination becomes possible with a certain probability of success. A variant of the problem is set discrimination: the states are grouped in sets and we want to determine to which particular set a given pure input state belongs. We consider here the simplest case, termed quantum state filtering, when the N given nonorthogonal states $\{|\psi_1\rangle, \ldots, |\psi_N\rangle\}$ are divided into two sets and the first set consists of one state only while the second consists of all of the remaining states. We present the derivation of the optimal measurement strategy, in terms of a generalized measurement (positive-operator-valued measure), to distinguish $|\psi_1\rangle$ from the set $\{|\psi_2\rangle, \ldots, |\psi_N\rangle\}$ and the corresponding optimal success and failure probabilities. The results, but not the complete derivation, were presented previously [Phys. Rev. Lett. **90**, 25901 (2003)] as the emphasis there was on appplication of the results to probabilistic quantum algorithms. We also show that the problem is equivalent to the discrimination of a pure state and an arbitrary mixed state.

I. INTRODUCTION

In quantum information and quantum computing the carrier of information is a quantum system and information is encoded in its state. In the simplest case the system lives in a two-dimensional Hilbert space and the two basis vectors are conveniently asssociated with the logical 0 and 1. Such a two-level system is called a qubit. However, it is not necessary to restrict our attention to qubits; d-dimensional systems, or qudits, can also be used to store quantum information. Reading out the information from the quantum system is tantamount to identifying the state it is in where the state itself might be the output of a quantum channel or the result of a quantum computation.

We want to find the optimum measurement that extracts information about the state. This problem is different from the usual textbook measurement as no ensemble averaging is involved nor are we interested in the average value of some physical observable. Every time a system reaches the final step we want to determine its state. Since the state of a system is not an observable in quantum mechanics, this sounds at first as if it is an impossible task. There are ways around it, however. Quantum processors are designed in such a way that their output is a member of a set of *known* states, so we are facing the more modest problem of determining which of these states was realized. If the possible target states are mutually orthogonal this is an easy task: we just set up detectors along the corresponding orthogonal directions and determine which one clicks (assuming perfect detectors, of course). However, if the target states are not mutually orthogonal the problem is still difficult and optimization with respect to some reasonable criteria leads, in general, to highly nontrivial measurement strategies. Finding the optimal measurement strategy is the subject of state discrimination. An overview of the state-of-the-art in the area of state discrimination can be found, for example, in our recent review [1] so here we just recall the immediate preliminaries.

One possible criterion is that no error is permitted, i.e., the states have to be discriminated unambiguously. Quantum measurement theory tells us that it is impossible to unambiguously discriminate between nonorthogonal quantum states with unit probability of success so we have to settle for less. If we do not require that we succeed every time, then unambiguous discrimination becomes possible. When the attempt fails, an inconclusive answer is returned. The optimal strategy is the one that minimizes the average probability of failure. Interest in unambiguous state discrimination was renewed by the suggestion to use nonorthogonal quantum states in certain secure quantum cryptographic protocols, in order to establish a secure key. A particularly clear example, based on a two-state procedure, was developed by Bennett [2].

In most of the previous work discrimination among all members of a set of states was considered. Subsequently, we turned our attention to the following variant of the problem. Instead of discriminating among all states, we ask what happens if we just want to discriminate between subsets of them. In this class of problems we know that a given system is prepared in one of N known nonorthogonal quantum states, but we do not know which one. We want to assign the state of this system to one or the other of two complementary subsets of the set of the N given states where one subset has M elements and the other has $N-M$ ($M \leqslant N/2$). Since the subsets are not mutually orthogonal, the assignment cannot be done with a 100% probability of success. For the case that the assignment is to be performed with minimum error, the solution has been found for arbitrary M and N under the restriction that the Hilbert space spanned by the states is two dimensional [3]. For the case that the assignment is required to be unambiguous, at the expense of allowing inconclusive results to occur, the probability of which is minimized, the problem has been solved for $M=1$, $N=3$ in Ref. [4]. We refer to either case as quantum state filtering, a term that we

coined in Ref. [3], when $M=1$ and $N \geq 3$. The solution presented in Ref. [4] can be generalized in a straightforward manner to arbitrary N. In our recent work we presented the exact analytical solution, but not its derivation, of the unambiguous quantum state filtering problem—the case of $M=1$ and N arbitrary, with no restriction on the states—and employed it to develop a quantum algorithm [5].

In this paper we fill in the gaps and derive the solution that we used in Ref. [5]. The paper is organized as follows. In Sec. II, based on simple but rigorous arguments, we derive the main analytical solution to the optimal positive-operator-valued measurement (POVM) problem. It invokes Neumark's theorem in order to develop a physical implementation of a generalized measurement. In Sec. III, we investigate the region of validity of the POVM solution and show that outside this region standard von Neumann projective measurements can be used to perform optimal unambiguous discrimination. In Sec. IV, we connect the problem of quantum state filtering to the unambiguous discrimination of mixed quantum states and show that our solution can be viewed as the discrimination of a pure state and an arbitrary mixed state. In Sec. V we give an alternative derivation of the optimal measurement which is based on considering the geometry of the Hilbert space, and it is closer in spirit to the standard approach to POVMs. Of course, the results are identical to those of the previous sections. A brief discussion of recent experimental progress and conclusions are given in Sec. VI.

II. DERIVATION OF THE OPTIMAL POVM

Suppose we are given a quantum system prepared in the state $|\psi\rangle$, which is guaranteed to be a member of the set of N known nonorthogonal states $\{|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_N\rangle\}$, but we do not know which one. We denote by η_i the *a priori* probability that the system was prepared in the state $|\psi_i\rangle$. We want to find a procedure that will unambiguously assign the state of the quantum system to one or the other of two complementary subsets of the set of the N given nonorthogonal quantum states, either $\{|\psi_1\rangle\}$ or $\{|\psi_2\rangle, \ldots, |\psi_N\rangle\}$. For unambiguous discrimination the procedure has to be error free, i.e., it may fail to give us any information about the state, and if it fails, it must let us know that it has, but if it succeeds, it should never give us a wrong answer. Clearly, this is a variant of the unambiguous state discrimination problem, and we shall refer to such a procedure as quantum state filtering without error. We find that, in contrast to the unambiguous state discrimination problem, this will be possible even if $|\psi_1\rangle$ is not linearly independent from the set $\{|\psi_2\rangle, \ldots, |\psi_N\rangle\}$.

According to the quantum theory of measurement, the states cannot be discriminated perfectly if they are not mutually orthogonal. Thus, if we are given $|\psi_i\rangle$, we will have some probability p_i to correctly assign it to one of the subsets and, correspondingly, some failure probability $q_i = 1 - p_i$ to obtain an inconclusive answer. The average probabilities of success P and failure $Q = 1 - P$ to correctly assign the states $|\psi_i\rangle$, $i = 1, \ldots, N$ are

$$P = \sum_i^N \eta_i p_i, \tag{2.2}$$

$$Q = \sum_i \eta_i q_i, \tag{2.1}$$

respectively. Our objective is to find the set $\{q_i\}$ that minimizes the average probability of failure Q or, equivalently, the set $\{p_i\}$ that maximizes the average probability of success P.

The procedure we shall use is a so-called "generalized measurement," based on POVM [6]. Using Neumark's theorem, a POVM can be implemented in the following way [7]. We first embed the system in a larger Hilbert space $\mathcal{K}$ consisting of the original system space $\mathcal{H}$ and an auxiliary Hilbert space called the ancilla $\mathcal{A}$. We take $\mathcal{K}$ to be a tensor product $\mathcal{K} = \mathcal{H} \otimes \mathcal{A}$. Then we introduce an interaction between the system and ancilla corresponding to a unitary evolution on this larger space. The unitary evolution entangles the system degrees of freedom with those of the ancilla. Finally, a projective measurement is performed on the extra degrees of freedom. Due to the entanglement, a click in the ancilla detectors will also transform the state of the original system in a general way. We choose this resulting transformation of the system states to be the most appropriate for our filtering purposes.

In order to accomodate N states the dimension of the system space $\mathcal{H}$, d, need be no more than N, i.e., $d \leq N$. Equality holds when all of the vectors $|\psi_i\rangle$ are linearly independent. We will use N as the dimensionality of $\mathcal{H}$ in the following treatment. The dimension of the ancilla $\mathcal{A}$ is a key point in obtaining the optimal solution and we will consider it next.

The input state of the system is one of the vectors $|\psi_i\rangle$, which is now a vector in the subspace $\mathcal{H}$ of the total space $\mathcal{K}$, so that

$$|\psi_i^{\mathcal{K}}\rangle_{\text{in}} = |\psi_i^{\mathcal{H}}\rangle |\phi_0^{\mathcal{A}}\rangle, \tag{2.2}$$

where $|\phi_0^{\mathcal{A}}\rangle$ is the initial state of the ancilla (same for all inputs). Following the general procedure outlined in the previous paragraph for the generalized measurement we now apply a unitary transformation U that entangles the system with the ancilla degrees of freedom. As a result, the input vector transforms into the state $|\psi_i^{\mathcal{K}}\rangle_{\text{out}}$. This state can be expanded using a basis $\{|m_k^{\mathcal{A}}\rangle\}$ for $\mathcal{A}$. For the purposes of optimum unambiguous discrimination between the two sets, we want three different outcomes when a projective measurement is performed on the ancilla: one that tells us that the input was a state from the first set, one that tells us that it was from the second set, and one that tells us that the discrimination failed. Thus, we require the ancilla to be three-dimensional ($k = 1, 2, 3$), as explained below, yielding

$$|\psi_i^{\mathcal{K}}\rangle_{\text{out}} \equiv U|\psi_i^{\mathcal{K}}\rangle_{\text{in}} = \delta_{i,1}|\psi_1'^{\mathcal{H}}\rangle|m_1^{\mathcal{A}}\rangle + (1 - \delta_{i,1})|\psi_i'^{\mathcal{H}}\rangle|m_2^{\mathcal{A}}\rangle + |\psi_i''^{\mathcal{H}}\rangle|m_3^{\mathcal{A}}\rangle. \tag{2.3}$$

In the following we drop the upper index $\mathcal{H}$ and $\mathcal{A}$ if it does not lead to confusion. We also note that the states $|\psi_i'^{\mathcal{H}}\rangle$ and $|\psi_i''^{\mathcal{H}}\rangle$ are not normalized. From the construction of the output state we see that the first outcome is compatible with the first input, the second with an input state from the second set and the third outcome is compatible with both inputs. We

might want to require that $|\psi_1'\rangle$ be distinguishable from $|\psi_2'\rangle, \ldots, |\psi_N'\rangle$, yielding the condition

$$\langle \psi_1' | \psi_j' \rangle = 0 \qquad (2.4)$$

for $j = 2, \ldots, N$ (in general, i runs from 1 to N and j from 2 to N). Strictly speaking, though, this condition is only convenient but not necessary.

Now, a state selective measurement is performed on the ancilla that projects $|\psi_i^{\mathcal{K}}\rangle_{\text{out}}$ onto one of the basis vectors $|m_i\rangle$ ($i = 1, 2, 3$). If it projects $|\psi_i^{\mathcal{K}}\rangle_{\text{out}}$ onto $|m_1\rangle$ or $|m_2\rangle$, the procedure succeeds, because we can unambiguously assign the input to one or the other set. The probability to get this outcome, if the input state is $|\psi_i\rangle$, is

$$p_i = \langle \psi_i' | \psi_i' \rangle. \qquad (2.5)$$

If the measurement projects $|\psi_i^{\mathcal{K}}\rangle_{\text{out}}$ onto $|m_3\rangle$, the procedure fails because it conditionally transforms all input system states into the output that cannot be distinguished. The probability of this outcome, if the input state is $|\psi_i\rangle$, is

$$q_i = \langle \psi_i'' | \psi_i'' \rangle. \qquad (2.6)$$

From the unitarity of the transformation in Eq. (2.3) the relation

$$p_i + q_i = 1 \qquad (2.7)$$

immediately follows, by taking the scalar product of the two sides with their adjoints.

The nature of the problem we are trying to solve imposes a number of other constraints and requirements on the output vectors. Let us first consider the set of system states associated with a click in the $|m_3\rangle$ detector $\{|\psi_i''\rangle\}$, which we also call failure vectors. If they were linearly independent, we could apply a further state discrimination procedure to them [8], contrary to our assumption that this direction is associated with an inconclusive outcome. Therefore, the optimal procedure should lead to failure vectors to which we cannot successfully apply a state discrimination procedure, implying that they are linearly dependent. In fact, more is true and it is easy to show that they must be collinear by demonstrating that the contrary leads to contradiction. To this end, let us assume that the failure vectors are *not* collinear. Then at least one of the the failure vectors $|\psi_j''\rangle$ will have a component in the direction that is perpendicular to $|\psi_1''\rangle$ in $\mathcal{H}$. We can set up a detector in the system Hilbert space projecting onto this direction and a click of the detector will tell us that our input state was not $|\psi_1\rangle$ but one of the other $N-1$ states. Thus, contrary to our assumption that the third dimension of the ancilla is associated with the inconclusive outcome, further discrimination is possible. Hence, the failure vectors must be collinear.

Next, we take the scalar product between $|\psi_1^{\mathcal{K}}\rangle_{\text{out}}$ and $|\psi_j^{\mathcal{K}}\rangle_{\text{out}}$. Using Eq. (2.3) and the fact that U is unitary lead to the conditions

$$\langle \psi_1'' | \psi_j'' \rangle = \langle \psi_1 | \psi_j \rangle \qquad (j > 1). \qquad (2.8)$$

Our objective is to find the optimal $|\psi_1''\rangle$ and $|\psi_j''\rangle$ which satisfy Eqs. (2.5)–(2.8) and maximize the success probability P. We shall now explore the consequences of the conclusion

that $|\psi_i''\rangle$ ($i = 1, \ldots, N$) are collinear, i.e., the failure space, a subspace of $\mathcal{H}$, is one dimensional. If $|\psi_0\rangle$ is the basis vector spanning this Hilbert space then, taking Eq. (2.6) into account, we can write the failure vectors as

$$|\psi_i''\rangle = \sqrt{q_i} e^{\chi_i} |\psi_0\rangle, \qquad (2.9)$$

where χ_i is the phase of $|\psi_i''\rangle$. Substituting this representation of the failure vectors in Eq. (2.8) gives

$$\langle \psi_1 | \psi_j \rangle = \sqrt{q_1 q_j} e^{i(\chi_j - \chi_1)}, \qquad (2.10)$$

which determines the phases for $j = 2, \ldots, N$.

Taking the magnitude of Eq. (2.10), yields

$$q_1 q_j = |\langle \psi_1 | \psi_j \rangle|^2 \qquad (j > 1). \qquad (2.11)$$

These $N-1$ conditions are a consequence of unitarity and imply that only one of the N failure probabilities can be chosen independently. If we chose q_1 as the independent one we can express the others as $q_j = |\langle \psi_1 | \psi_j \rangle|^2 / q_1$. Let $O_{ij} \equiv \langle \psi_i | \psi_j \rangle$ then the average failure probability $Q = \Sigma_i^N \eta_i q_i$ can be written explicitly as

$$Q = \eta_1 q_1 + \frac{\sum_{j=2}^{N} \eta_j |O_{1j}|^2}{q_1}. \qquad (2.12)$$

From the condition for minimum

$$\frac{dQ}{dq_1} = 0, \qquad (2.13)$$

we now find the optimal value of q_1 as

$$q_1 = \sqrt{\frac{\sum_{j=2}^{N} \eta_j |O_{1j}|^2}{\eta_1}}. \qquad (2.14)$$

Inserting this value into Eq. (2.12) finally gives

$$Q_{\text{POVM}} = 2 \sqrt{\sum_{j=2}^{N} \eta_1 \eta_j |O_{1j}|^2}. \qquad (2.15)$$

This result represents the absolute optimum for the measurement problem at hand. In the following we will investigate its range of validity and derive the complete solution that is valid for all values of the parameters.

III. LIMITATIONS OF THE POVM AND THE COMPLETE SOLUTION

The value given in Eq. (2.15) for the minimum probability of failure cannot always be realized. For it to be true, there has to exist a unitary transformation that takes $|\psi_i\rangle_{\text{in}}$ to $|\psi_i\rangle_{\text{out}}$ in Eq. (2.3). One of the consequences of unitarity is the conservation of norm which is expressed by Eqs. (2.5)–(2.7). Another consequence is the conservation of the scalar product which we only partially used in Eq. (2.8). Taking the scalar product of $|\psi_l\rangle_{\text{out}}$ with $|\psi_k\rangle_{\text{out}}$ from Eq. (2.3) leads to the generalization of Eq. (2.8)

$$\langle\psi_l|\psi_k\rangle = \langle\psi_l'|\psi_k'\rangle + \sqrt{q_l q_k}\, e^{i(\chi_k - \chi_l)}. \tag{3.1}$$

Obviously, $k=1$ and $l=j>2$ reproduces Eq. (2.8) as a special case since, according to Eq. (2.4), $\langle\psi_1'|\psi_j'\rangle=0$ for $j=2,\ldots,N$. These equations imply that

$$\langle\psi_l'|\psi_k'\rangle = \langle\psi_l|\psi_k\rangle - \sqrt{q_l q_k}\, e^{i(\chi_k - \chi_l)}. \tag{3.2}$$

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

$$M_{lk} = \langle\psi_l|\psi_k\rangle - \sqrt{q_l q_k}\, e^{i(\chi_k - \chi_l)} \tag{3.3}$$

is positive semidefinite, as discussed in detail in Ref. [9].

The matrix $M_{lk} \equiv \langle\psi_l'|\psi_k'\rangle$ has the structure

$$M = \begin{pmatrix} M^\alpha & 0 \\ 0 & M^\beta \end{pmatrix}, \tag{3.4}$$

where $M^\alpha = M_{11} = 1 - q_1$ and all other elements in the first row and first column are zero because of the condition (2.4), and $M^\beta = M_{jj'}$ for $j, j' = 2,\ldots,N$.

Thus, one of the positivity conditions is $q_1 \leqslant 1$ from the positivity of M^α. If $q_1 = 1$, then the average failure probability, which we denote by Q_α, becomes

$$Q_\alpha = \eta_1 + \sum_{j=2}^{N} \eta_j |O_{1j}|^2, \tag{3.5}$$

which follows from Eq. (2.12) with $q_1 = 1$. Note that this is the same average probability that we would obtain if we projected the state of the system we were given onto $|\psi_1\rangle$.

In order to evaluate the positivity condition for M^β we first express the second term on the right-hand side of Eq. (3.3) as

$$\sqrt{q_j q_{j'}}\, e^{i(\chi_k - \chi_l)} = \frac{\langle\psi_j|\psi_1\rangle\langle\psi_1|\psi_{j'}\rangle}{q_1}, \tag{3.6}$$

where we multiplied Eq. (2.10) with its conjugate for j'. This allows us to write

$$M_{jj'} = \langle\psi_j|\psi_{j'}\rangle - \frac{\langle\psi_j|\psi_1\rangle\langle\psi_1|\psi_{j'}\rangle}{q_1}. \tag{3.7}$$

At this point, it is convenient to introduce the following notation. We call $\{|\psi_1\rangle\}$ the α set and $\{|\psi_j\rangle|j>1\}$ the β set. Define $\mathcal{H}_\alpha$ to be the one-dimensional space that is the span of $|\psi_1\rangle$, and $\mathcal{H}_\beta$ to be the span of $\{|\psi_j\rangle|j=2,\ldots,N\}$. In addition, let P_α be the projection onto $\mathcal{H}_\alpha$, and P_β be the projection onto $\mathcal{H}_\beta$. This gives us two different decompositions of the system Hilbert space $I_\mathcal{H} = P_\alpha + \bar{P}_\alpha = P_\beta + \bar{P}_\beta$, where the overbar stands for projection onto the orthogonal complement. Then we can write, for $j, j' > 1$,

$$M_{jj'} = \langle\psi_j|P_\beta|\psi_{j'}\rangle - \frac{\langle\psi_j|P_\beta|\psi_1\rangle\langle\psi_1|P_\beta|\psi_{j'}\rangle}{q_1}$$
$$= \left\langle\psi_j\left|\left[P_\beta - \frac{|\psi_1^\parallel\rangle\langle\psi_1^\parallel|}{q_1}\right]\right|\psi_{j'}\right\rangle, \tag{3.8}$$

where $|\psi_1^\parallel\rangle = P_\beta|\psi_1\rangle$ is the component of $|\psi_1\rangle$ in $\mathcal{H}_\beta$. This leads to a further decomposition of the $\mathcal{H}_\beta$ subspace, $P_\beta = P_\beta^\parallel + P_\beta^\perp$, where $P_\beta^\parallel \equiv |\psi_1^\parallel\rangle\langle\psi_1^\parallel|/\langle\psi_1^\parallel|\psi_1^\parallel\rangle$ and $P_\beta^\perp$ is the projec-

tion onto the orthogonal subspace of $\mathcal{H}_\beta$. Thus, M^β is positive semidefinite if $q_1 \geqslant \langle\psi_1^\parallel|\psi_1^\parallel\rangle$. When $q_1 = \langle\psi_1^\parallel|\psi_1^\parallel\rangle$, the failure probability is

$$Q_\beta = \eta_1\langle\psi_1^\parallel|\psi_1^\parallel\rangle + \frac{\sum_{j=2}^{N}\eta_j|O_{1j}|^2}{\langle\psi_1^\parallel|\psi_1^\parallel\rangle}. \tag{3.9}$$

This is the same failure probability that is obtained by projecting each quantum system we are given onto $P_\beta^\parallel$.

Combining the conditions for the positivity of M^α and M^β, we find that the POVM solution is valid if

$$\langle\psi_1^\parallel|\psi_1^\parallel\rangle \leqslant q_1 \leqslant 1. \tag{3.10}$$

In view of Eq. (2.11) this condition ensures that all failure probabilities will be bounded by similar inequalities

$$\langle\psi_j^\parallel|\psi_j^\parallel\rangle \leqslant q_j \leqslant 1, \tag{3.11}$$

where we introduced the notation $|\psi_j^\parallel\rangle = P_\alpha|\psi_j\rangle$ for the component of any state from the second set in $\mathcal{H}_\alpha$.

The boundaries for the validity of the POVM solution (3.10) [or Eq. (3.11)], can be expressed in terms of the independent parameters of the problem. The *a priori* probability that the input state is from the α set is $\eta_\alpha \equiv \eta_1$, and the *a priori* probability that it is from the β set is $\eta_\beta \equiv 1 - \eta_\alpha (= 1 - \eta_1)$. Next, we introduce the renormalized *a priori* probabilities, $\eta_j' = \eta_j/\eta_\beta$, for $j > 1$. In terms of these renormalized quantities we can write q_1 for the optimal POVM, Eq. (2.14), as

$$q_1 = \sqrt{\frac{(1 - \eta_1)\sum_{j=2}^{N}\eta_j'|O_{1j}|^2}{\eta_1}}. \tag{3.12}$$

Substitution into Eq. (3.10) yields upper and lower bounds for the *a priori* probability of the state to be filtered,

$$\frac{S}{S+1} \leqslant \eta_1 \leqslant \frac{S}{S + |\langle\psi_1^\parallel|\psi_1^\parallel\rangle|^2}, \tag{3.13}$$

with

$$S \equiv \sum_{j=2}^{N}\eta_j'|O_{1j}|^2. \tag{3.14}$$

Within these bounds the POVM solution is valid.

Summarizing Eqs. (2.15), (3.5), and (3.9) and taking Eq. (3.13) into account, we can write the optimal solution as

$$Q^{\text{opt}} = \begin{cases} Q_{\text{POVM}} & \text{if } \dfrac{S}{1+S} \leqslant \eta_1 \leqslant \dfrac{S}{S + |\langle\psi_1^\parallel|\psi_1^\parallel\rangle|^2}, \\[2mm] Q_\alpha & \text{if } \eta_1 < \eta_l \equiv \dfrac{S}{1+S}, \\[2mm] Q_\beta & \text{if } \eta_1 > \eta_u \equiv \dfrac{S}{S + |\langle\psi_1^\parallel|\psi_1^\parallel\rangle|^2}, \end{cases} \tag{3.15}$$

representing our main result. In the intermediate range of η_1 the optimal failure probability Q_{POVM} is achieved by a generalized measurement or POVM. Outside this region, the op-

timal failure probabilities Q_α and Q_β are realized by standard von Neumann measurements, corresponding to two different orthogonal decompositions of $\mathcal{H}$. For $\eta_1 < \eta_l$, $I_{\mathcal{H}} = P_\alpha + \bar{P}_\alpha$. A click of the P_α detector corresponds to failure because it can have its origin in either of the two subsets and a click in the orthogonal directions uniquely assigns the input state to the β set. For $\eta_1 > \eta_u$, $I_{\mathcal{H}} = P_\beta^{\parallel} + P_\beta^{\perp} + \bar{P}_\beta$. A click of the $P_\beta^{\parallel}$ detector corresponds to failure because it can have its origin in either of the two subsets, a click of the $\bar{P}_\beta$ detector uniquely assigns the input state to the α set and a click of the $P_\beta^{\perp}$ detector uniquely assigns the input state to the β set. At the boundaries of their respective regions of validity, the optimal measurements transform into one another continuously. In its range of validity the POVM performs better than either one of the two possible von Neumann measurements.

Finally, we want to point to an interesting feature of the solution. The results hold true even when the first input state $|\psi_1\rangle$ lies entirely in $\mathcal{H}_\beta$. In this case the two von Neumann decompositions coincide and the range of validity of the POVM solution shrinks to zero. A click in the P_α detector corresponds to failure since it can originate from either of the two subsets and a click in one of the detectors along the orthogonal directions unambiguously identifies an input from the β set.

IV. SET DISCRIMINATION AS DISCRIMINATION OF MIXED STATES

In this section we shall establish a connection between quantum state filtering and the discrimination of mixed states. In fact, we will show that filtering is equivalent to the problem of discrimination between a pure state (a rank 1 mixed state) and an arbitrary (rank N) mixed state. Thus filtering can be regarded as an instance of mixed state discrimination.

It is possible to express a number of the quantities in the solution in a more compact way. Since we do not want to resolve the individual states in the two sets, the states in a set can be given an ensemble description. To make the connection between the set discrimination and the ensemble viewpoint, we define two density matrices

$$\rho_\alpha = |\psi_1\rangle\langle\psi_1|,$$

$$\rho_\beta = \sum_{j=2}^{N} \eta_j' |\psi_j\rangle\langle\psi_j|, \qquad (4.1)$$

where the primed quantities have been introduced in connection with Eq. (3.12). The *a priori* probabilities of these states are given by $\eta_\alpha = \eta_1$ and $\eta_\beta = 1 - \eta_1$, respectively. Since these density matrices completely characterize the sets all results should be expressible in terms of them. Indeed, we have immediately that

$$S = \langle\psi_1|\rho_\beta|\psi_1\rangle = \text{Tr}(\rho_\alpha\rho_\beta). \qquad (4.2)$$

We ultimately want to find a compact expression for the optimal failure probabilities. We can express Q_α in terms of ρ_α, ρ_β, and P_α as

$$Q_\alpha = \eta_\alpha + \eta_\beta \,\text{Tr}(\rho_\alpha\rho_\beta)$$
$$= \eta_\alpha + \eta_\beta S = \eta_\alpha \,\text{Tr}(P_\alpha\rho_\alpha P_\alpha) + \eta_\beta \,\text{Tr}(P_\alpha\rho_\beta P_\alpha). \qquad (4.3)$$

The last expression, although superfluous, makes it explicit that in this case the measurement is a von Neumann projection on the one-dimensional subspace $\mathcal{H}_\alpha$.

Similarly, we can express Q_β in terms of the density matrices and $P_\beta^{\parallel}$ as

$$Q_\beta = \eta_\alpha\langle\psi_1^{\parallel}|\psi_1^{\parallel}\rangle + \eta_\beta \frac{\text{Tr}(\rho_\alpha\rho_\beta)}{\langle\psi_1^{\parallel}|\psi_1^{\parallel}\rangle}$$
$$= \eta_\alpha\langle\psi_1^{\parallel}|\psi_1^{\parallel}\rangle + \eta_\beta \frac{S}{\langle\psi_1^{\parallel}|\psi_1^{\parallel}\rangle}$$
$$= \eta_\alpha \,\text{Tr}(P_\beta^{\parallel}\rho_\alpha P_\beta^{\parallel}) + \eta_\beta \,\text{Tr}(P_\beta^{\parallel}\rho_\beta P_\beta^{\parallel}). \qquad (4.4)$$

The last expression, although again superfluous, makes it explicit that in this case the measurement is a von Neumann projection on the one-dimensional subspace $\mathcal{H}_{\tilde{\beta}}$.

Finally, Q_{POVM} can be written in terms of the density matrices as

$$Q_{\text{POVM}} = 2\sqrt{\eta_\alpha\eta_\beta S} = 2\sqrt{\eta_\alpha\eta_\beta \,\text{Tr}(\rho_\alpha\rho_\beta)}. \qquad (4.5)$$

Since all of the failure probabilities can be expressed in terms of invariant expressions of the density matrices only, we have just shown that filtering is equivalent to the optimal unambiguous discrimination between a rank 1 mixed state (a pure state) and an arbitrary mixed state, providing the simplest example for discrimination between mixed states.

Before leaving the realm of mixed state discrimination we want to point to an interesting connection to earlier work. We notice that the fidelity F between a pure state $|\psi\rangle$ and a mixed state ρ is given by [10]

$$F(|\psi\rangle\langle\psi|, \rho) = \sqrt{\langle\psi|\rho|\psi\rangle} = \sqrt{\text{Tr}(|\psi\rangle\langle\psi|\rho)}. \qquad (4.6)$$

The optimal POVM failure probability can be written as

$$Q_{\text{POVM}} = 2\sqrt{\eta_\alpha\eta_\beta}F(\rho_\alpha, \rho_\beta). \qquad (4.7)$$

This coincides with the lower bound on the optimal failure probability found by Rudolph *et al.* [11], constructively proving that, for this case, the lower bound can be saturated in the range of validity of the optimal POVM.

V. GEOMETRICAL INTERPRETATION OF THE OPTIMAL MEASUREMENTS

In this section we show that for a complete description of the POVM one does not need to invoke Neumark's theorem. In fact, a complete description is possible without ever leaving the Hilbert space of the system and enlarging it with the ancilla degrees of freedom [6]. Of course, Neumark's theorem is still useful when it comes to a physical implementation of the POVM.

The POVM will have three possible measurement results, one that corresponds to $|\psi_1\rangle$, one that corresponds to the β set and one that corresponds to failure. In order to describe the measurement, we introduce the quantum detection opera-

tors Π_1, Π_2, and Π_0, also called POVM elements, corresponding the three possible measurement results. We then have that $\langle\psi_1|\Pi_1|\psi_1\rangle=p_1(\equiv p_\alpha)$ is the probability of successfully identifying $|\psi_1\rangle$, $\langle\psi_1|\Pi_0|\psi_1\rangle=q_1(\equiv q_\alpha)$ is the probability of failing to identify $|\psi_1\rangle$, $\langle\psi_j|\Pi_2|\psi_j\rangle=p_j$ is the probability of successfully assigning $|\psi_j\rangle$ (for $j=2,\ldots,N$) to the β set, and $\langle\psi_j|\Pi_0|\psi_j\rangle=q_j$ is the probability of failing to assign $|\psi_j\rangle$. For later purposes, we also introduce $p_\beta=\Sigma_2^N\eta_j'p_j$ and $q_\beta=1-p_\beta$. For unambiguous filtering we then require $\langle\psi_j|\Pi_1|\psi_j\rangle=\langle\psi_1|\Pi_2|\psi_1\rangle=0$ (for $j=2,\ldots N$). We want these possibilities to be exhaustive,

$$\Pi_1+\Pi_2+\Pi_0=I, \tag{5.1}$$

where I is the identity in $\mathcal{H}$. The probabilities are always real and non-negative which implies that the quantum detection operators are non-negative. The conditions of positivity and unambiguous filtering require that

$$\Pi_1|\psi_j\rangle=0,$$

$$\Pi_2|\psi_1\rangle=0, \tag{5.2}$$

for $j=2,\ldots,N$.

In order to find the form of the POVM elements explicitly, it is useful to define the subspace $\mathcal{H}_1$ to be the linear span of the, in general nonorthogonal but linearly independent, vectors $|\psi_1\rangle$ and $|\psi_1^\parallel\rangle$. Note that $\mathcal{H}_1^\perp\subseteq\mathcal{H}_\alpha^\perp$ where $\perp$ denotes the orthogonal complement in $\mathcal{H}$. The two POVM elements Π_1 and Π_2 will be related to two different orthogonal decompositions of $\mathcal{H}_1$. Indeed, the first of the above requirements immediately gives us the form of Π_1. We must have

$$\Pi_1=c_1|e_1\rangle\langle e_1|, \tag{5.3}$$

where $|e_1\rangle$ is the unit vector in $\mathcal{H}_1$ that is orthogonal to $|\psi_1^\parallel\rangle$. The constant $0\leq c_1\leq 1$ remains to be determined.

The second requirement tells us that the support of Π_2 is contained in $\mathcal{H}_\alpha^\perp$, the subspace orthogonal to $\mathcal{H}_\alpha$. We can learn more about Π_2 by looking at the failure operator which, from Eq. (5.1), is given as

$$\Pi_0=I-\Pi_1-\Pi_2. \tag{5.4}$$

This operator must be positive, and we want the failure probabilities to be as small as possible. For a normalized vector $|v\rangle\in\mathcal{H}_1^\perp$, we have

$$\langle v|\Pi_0|v\rangle=1-\langle v|\Pi_2|v\rangle. \tag{5.5}$$

This will achieve the minimum value consistent with the positivity of Π_0, which is 0, if $\Pi_2|v\rangle=|v\rangle$. This means that we can express Π_2 as

$$\Pi_2=P_1\Pi_2P_1+\bar{P}_1, \tag{5.6}$$

where P_1 is the projection onto $\mathcal{H}_1$, and $\bar{P}_1=I-P_1$. The appearance of the projector $\bar{P}_1$ in the POVM element is a consequence of what is called the reduction theorem in Ref. [12]. Define $|e_2\rangle$ to be the normalized vector in $\mathcal{H}_1$ that is orthogonal to $|\psi_1\rangle$. The second requirement in Eq. (5.2) implies that $P_1\Pi_2P_1=c_2|e_2\rangle\langle e_2|$, where $0\leq c_2\leq 1$. Combining our results for the different parts of Π_2, we have that

$$\Pi_2=c_2|e_2\rangle\langle e_2|+\bar{P}_1. \tag{5.7}$$

It is also possible to express the, as yet undetermined, constants c_1 and c_2 in terms of the success or failure probabilities for the sets. From the definition of these probabilities, given at the beginning of this section, we find

$$c_1=\frac{1-q_\alpha}{1-\|\psi_1^\parallel\|^2},$$

$$c_2=\frac{1-\frac{\|\psi_1^\parallel\|^2}{S}q_\beta}{1-\|\psi_1^\parallel\|^2}. \tag{5.8}$$

Our final task is to choose c_1 and c_2 as large as possible (this will minimize the failure probabilities) consistent with the requirement that Π_0 be positive. Since Π_0 is a simple 2 by 2 matrix in $\mathcal{H}_1$, the corresponding eigenvalue problem can be solved analytically. Non-negativity of the eigenvalues leads, after some tedious but straightforward algebra, to the condition

$$q_\alpha q_\beta=S\equiv\mathrm{Tr}(\rho_\alpha\rho_\beta). \tag{5.9}$$

Note that this condition is consistent with Eq. (2.11). Multiplying Eq. (2.11) with η_j' and taking the sum over j leads to the above condition. The task then is to find the minimum of the average failure probability

$$Q=\eta_\alpha q_\alpha+\eta_\beta q_\beta, \tag{5.10}$$

under the constraint of Eq. (5.9). This, once again, gives the solution (3.15), found via the Neumark approach. In particular, we obtain the optimum values of the failure probabilities as

$$q_\alpha=\sqrt{\frac{\eta_\beta}{\eta_\alpha}S}, \quad q_\beta=\sqrt{\frac{\eta_\alpha}{\eta_\beta}S}. \tag{5.11}$$

Inserting these values in Eq. (5.8) gives us the explicit expressions for the optimal POVM elements. More importantly, the positivity conditions of c_1 and c_2 give us the range of existence of the POVM solution. Obviously, for $c_1>0$ we have to require $q_\alpha<1$ and for $c_2>0$ we have to require $q_\beta<S/\|\psi_1^\parallel\|^2$. Combining these with Eq. (5.9) we obtain that the POVM solution is valid in the interval

$$\|\psi_1^\parallel\|^2\leq q_\alpha\leq 1, \quad S\leq q_\beta\leq\frac{S}{\|\psi_1^\parallel\|^2}, \tag{5.12}$$

which is of course identical to our earlier findings.

From these results, it is now very easy to see what happens at the boundaries. When $q_\alpha=1$ and $q_\beta=S$ we have $c_1=0$ and $c_2=1$ and the POVM degenerates into projective von Neumann measurements corresponding to the second decomposition of $\mathcal{H}_1$. $P_1\Pi_2P_1=|e_2\rangle\langle e_2|$ will be part of Π_2 for successfully identifying an input from the β set and $\Pi_0=|\psi_1\rangle\langle\psi_1|$ becomes a projector for failure, so the input $|\psi_1\rangle$ will be missed completely. Conversely, when $q_\alpha=\|\psi_1^\parallel\|^2$ and $q_\beta=S/\|\psi_1^\parallel\|^2$ we have $c_1=1$ and $c_2=0$ and the POVM degenerates into projective von Neumann measurements corresponding to the first decomposition of $\mathcal{H}_1$. Now, we have

$\Pi_1 = |e_1\rangle\langle e_1|$ for successfully identifying the input as being from the α set and $\Pi_0 = |\psi_1^{\parallel}\rangle\langle\psi_1^{\parallel}|/\langle\psi_1^{\parallel}|\psi_1^{\parallel}\rangle$ becomes a projector for failure. In this later case both types of input can be identified. Finally, we note that from these considerations it is clear that for the implementation of Neumark's theorem only the subspace $\mathcal{H}_1$ has to be entangled with the ancilla, giving further directions for an experimental realization.

VI. CONCLUSIONS

The usual problem considered when trying to unambiguously discriminate among quantum states is to correctly identify which state a given system is in when one knows the set of possible states in which it can be prepared. Here we have considered a related problem that can lead to further generalizations and applications in quantum information and quantum computing. The set of N possible states is divided into two subsets, and we only want to know to which subset the quantum state of our given system belongs. We considered the simplest instance of this problem, the situation in which we are trying to discriminate between a set containing one quantum state and another containing the remaining $N-1$ states. A method for finding the optimal strategy for discriminating between these two sets was presented, and explicit analytical solutions were given. For the special case of $N=3$, which we treated earlier, we proposed a quantum optical implementation of the optimal POVM strategy based on linear optical devices only [4]. Since our original proposal the experiment has been performed, and the results are in perfect agreement with our theoretical predictions [13].

One application of these results is the development of quantum algorithms [5]. A more detailed consideration of these and related problems is left for a subsequent publication [14].

[1] J. A. Bergou, U. Herzog, and M. Hillery, Lect. Notes Phys. **649**, 417 (2004).

[2] C. H. Bennett, Phys. Rev. Lett. **68**, 3121 (1992).

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

[4] Y. Sun, J. A. Bergou, and M. Hillery, Phys. Rev. A **66**, 032315 (2002).

[5] J. A. Bergou, U. Herzog, and M. Hillery, Phys. Rev. Lett. **90**, 257901 (2003); quant-ph/0209007.

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

[7] M. A. Neumark, Izv. Akad. Nauk SSSR, Ser. Mat. **4**, 277 (1940).

[8] A. Chefles, Phys. Lett. A **239**, 339 (1998).

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

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

[11] T. Rudolph, R. W. Spekkens, and P. S. Turner, Phys. Rev. A **68**, 010301(R) (2003).

[12] Ph. Raynal, N. Lütkenhaus, and S. van Enk, Phys. Rev. A **68**, 022308 (2003).

[13] M. Mohseni, A. Steinberg, and J. A. Bergou, quant-ph/0401002; Phys. Rev. Lett. **93**, 200403 (2004).

[14] J. A. Bergou and M. Hillery (in preparation).