Discrimination of two mixed quantum states with maximum confidence and minimum probability of inconclusive results

Ulrike Herzog
Institut für Physik, Humboldt-Universität Berlin, Newtonstrasse 15, D-12489 Berlin, Germany
(Received 16 December 2008; published 19 March 2009)

We study an optimized measurement that discriminates two mixed quantum states with maximum confidence for each conclusive result, thereby keeping the overall probability of inconclusive results as small as possible. When the rank of the detection operators associated with the two different conclusive outcomes does not exceed unity, we obtain a general solution. As an application, we consider the discrimination of two mixed qubit states. Moreover, for the case of higher-rank detection operators we give a solution for particular states. The relation of the optimized measurement to other discrimination schemes is also discussed.

DOI: 10.1103/PhysRevA.79.032323 PACS number(s): 03.67.Hk, 03.65.Ta, 42.50.−p

I. INTRODUCTION

Quantum state discrimination [1–3] lies at the heart of quantum communication and quantum cryptography. Since information is encoded into states of a quantum system, these states have to be distinguished when the information is read out. In the standard discrimination problem, the quantum system is prepared in a certain state that belongs to a finite set of given states which occur with known prior probabilities. When the states are nonorthogonal, they cannot be distinguished perfectly and therefore discrimination strategies have been developed which are optimized with respect to various criteria. The most prominent of these are the discrimination with minimum error [4] and the optimum unambiguous discrimination, originally introduced for two pure states [5,6]. If unambiguous discrimination errors are not allowed, at the expense of admitting a certain fraction of inconclusive results, where the measurement fails to give a definite answer. In general, a variety of measurements may lead to unambiguous, that is, error-free discrimination. The optimum measurement is defined as the one that minimizes the overall probability of inconclusive results.

Unambiguous discrimination is not always possible. When the states in the given set are pure, they must be linearly independent [7], and when they are mixed, the supports [8] of their density operators must be different in order to distinguish them without error [9–18]. For the case that some or all states in the set cannot be unambiguously discriminated, recently Croke *et al.* [19,20] introduced the strategy of discriminating them with maximum possible confidence. When a state can be unambiguously distinguished, the confidence in the respective measurement outcome is defined to be equal to one, otherwise it is smaller. As for unambiguous discrimination, also for the maximum-confidence discrimination the measurement is in general not unique [19] and additional optimization criteria can be applied.

In this paper, we consider the discrimination of two mixed quantum states. We investigate the optimized measurement that distinguishes between them with the maximum confidence for each of the two distinct outcomes, thereby keeping the probability of inconclusive results, where the measurement fails to give a definite answer, as small as possible. Our treatment generalizes previous results [13–15] derived for the optimum unambiguous discrimination of two mixed quantum states. The paper is organized as follows. Section II provides the general description of a measurement for discriminating two mixed quantum states with maximum confidence. In Sec. III the specific measurement that achieves this goal with minimum overall failure probability is investigated and applications are given, considering also the relation to optimum unambiguous discrimination and to discrimination with minimum error. Section IV concludes the paper with a discussion and a summary.

II. GENERAL MAXIMUM-CONFIDENCE MEASUREMENT FOR TWO MIXED STATES

We suppose that a quantum system is prepared in the given mixed states ρ_1 and ρ_2 with the prior probabilities η_1 and η_2, respectively, where $\eta_1+\eta_2=1$. We want to perform a measurement in order to infer from a single outcome whether the state of the system was ρ_1 or ρ_2. In general, the discrimination made upon this inference may be erroneous and inconclusive results may also occur. A complete discrimination measurement is described by three positive detection operators Π_1, Π_2, and $\Pi_?$ summing up to the identity operator I_d in the d-dimensional joint Hilbert space $\mathcal{H}_d$ spanned by the eigenstates of ρ_1 and ρ_2 belonging to nonzero eigenvalues [1–3], that is,

$$\Pi_? = I_d - \Pi_1 - \Pi_2 \geq 0, \quad \Pi_1 \geq 0, \quad \Pi_2 \geq 0. \tag{1}$$

The probability that a system prepared in the state ρ_k is inferred to be in the state ρ_j is given by $\mathrm{Tr}(\rho_k\Pi_j)$ with $j,k=1,2$, while $\mathrm{Tr}(\rho_k\Pi_?)$ is the probability that the measurement fails and yields an inconclusive result. The overall failure probability Q of the discrimination measurement then reads as

$$Q = \mathrm{Tr}(\rho\Pi_?) = 1 - \mathrm{Tr}(\rho\Pi_1) - \mathrm{Tr}(\rho\Pi_2), \tag{2}$$

where we have introduced the density operator

$$\rho = \eta_1\rho_1 + \eta_2\rho_2, \tag{3}$$

characterizing the total information about the quantum system. When all detection operators are projectors, the measurement is a von Neumann measurement, otherwise it is a generalized measurement based on a positive operator-valued measure (POVM). From the detection operators Π_j, schemes for realizing the measurement can be obtained [21].

The confidence in the conclusive measurement outcome j, which we shall denote by C_j, has been introduced [19] as the conditional probability $P(\rho_j|j)=P(\rho_j,j)/P(j)$ that the state ρ_j was indeed prepared given that the outcome j is detected. In our case, we have

$$C_j = \frac{\eta_j \mathrm{Tr}(\rho_j \Pi_j)}{\mathrm{Tr}(\rho\Pi_j)} = \frac{\eta_j \mathrm{Tr}(\rho_j \Pi_j)}{\eta_1 \mathrm{Tr}(\rho_1 \Pi_j) + \eta_2 \mathrm{Tr}(\rho_2 \Pi_j)}, \tag{4}$$

with $j=1,2$. Here $P(\rho_j,j)=\eta_j\mathrm{Tr}(\rho_j\Pi_j)$ is the joint probability that the state ρ_j was prepared and the detector j clicks, and $P(j)=\mathrm{Tr}(\rho\Pi_j)$ is the total probability for the detection of the outcome j. In other words, the confidence C_j is the ratio between the number of instances when the outcome j is correct and the total number of instances when the outcome j is detected. Similar to Ref. [19], we define the positive operators

$$\tilde{\rho}_j = \eta_j \rho^{-1/2}\rho_j\rho^{-1/2}, \quad \tilde{\Pi}_j = \frac{\rho^{1/2}\Pi_j\rho^{1/2}}{\mathrm{Tr}(\rho\Pi_j)} \tag{5}$$

and obtain from Eq. (4) the confidences

$$C_j = \mathrm{Tr}(\tilde{\rho}_j\tilde{\Pi}_j). \tag{6}$$

Let us write the operator $\tilde{\rho}_1$ as

$$\tilde{\rho}_1 = \nu_{\max}^{(1)}\sum_{k=1}^{m}|\nu_k\rangle\langle\nu_k| + \nu_{\min}^{(1)}\sum_{k=m+1}^{m+n}|\nu_k\rangle\langle\nu_k| + \sum_{k=m+n+1}^{d}\nu_k^{(1)}|\nu_k\rangle\langle\nu_k|, \tag{7}$$

where the eigenstates $\{|\nu_k\rangle\}$ with $\langle\nu_k|\nu_{k'}\rangle=\delta_{kk'}$ form a d-dimensional orthonormal basis in $\mathcal{H}_d$. Here $\nu_{\max}^{(1)}$ and $\nu_{\min}^{(1)}$ are the largest and smallest eigenvalues of $\tilde{\rho}_1$, respectively, and m and n denote their degrees of degeneracy. From Eqs. (5) and (3) we get

$$\tilde{\rho}_1 + \tilde{\rho}_2 = \rho^{-1/2}\rho\rho^{-1/2} = I_d, \tag{8}$$

showing that the eigenvalues of $\tilde{\rho}_1$ and $\tilde{\rho}_2$ do not exceed 1. From

$$\tilde{\rho}_2 = I_d - \tilde{\rho}_1 = \sum_{k=1}^{d}|\nu_k\rangle\langle\nu_k| - \tilde{\rho}_1, \tag{9}$$

we conclude that the eigenstates belonging to the smallest eigenvalue of $\tilde{\rho}_1$, given by $\nu_{\min}^{(1)}$, are associated with the largest eigenvalue of $\tilde{\rho}_2$, given by $\nu_{\max}^{(2)}=1-\nu_{\min}^{(1)}$ and vice versa.

We consider a measurement that achieves the maximum possible confidences $C_1^{\max}$ and $C_2^{\max}$ for the discrimination of each of the two given states. By representing $\tilde{\Pi}_j$ with the help of the orthonormal basis $\{|\nu_k\rangle\}$, it follows from Eqs. (6), (7), and (9) that the operators $\tilde{\Pi}_j$ maximizing C_j for $j=1,2$ take the form

$$\tilde{\Pi}_1 = \sum_{k,k'=1}^{m}\alpha_{kk'}|\nu_k\rangle\langle\nu_{k'}|, \quad \tilde{\Pi}_2 = \sum_{k,k'=m+1}^{m+n}\beta_{kk'}|\nu_k\rangle\langle\nu_{k'}|, \tag{10}$$

where due to $\mathrm{Tr}\,\tilde{\Pi}_j=1$, we have to require that

$$\sum_{k=1}^{m}\alpha_{kk}=1, \quad \sum_{k=m+1}^{m+n}\beta_{kk}=1. \tag{11}$$

These operators yield the maximum confidences

$$C_1^{\max}=\nu_{\max}^{(1)}, \quad C_2^{\max}=\nu_{\max}^{(2)}=1-\nu_{\min}^{(1)}, \tag{12}$$

corresponding to the largest eigenvalues of the operators $\tilde{\rho}_1$ and $\tilde{\rho}_2$, respectively, in accordance with Ref. [19]. Using Eq. (12) we obtain the general relation

$$C_1^{\max}+C_2^{\max}=1+\nu_{\max}^{(1)}-\nu_{\min}^{(1)}>1, \tag{13}$$

where we took into account that the case of all eigenvalues of $\tilde{\rho}_1$ being identical is excluded since it would correspond to $\rho_1=\rho_2$.

From Eq. (5) it becomes obvious that the operators $\tilde{\Pi}_j$ and ρ define the detection operators Π_j only up to an arbitrary constant c_j and additional optimization criteria can be applied [19]. Using Eq. (10), the general structure of the detection operators discriminating ρ_1 and ρ_2 with maximum confidence thus reads as

$$\Pi_1 = c_1\sum_{k,k'=1}^{m}\alpha_{kk'}\rho^{-1/2}|\nu_k\rangle\langle\nu_{k'}|\rho^{-1/2}, \tag{14}$$

$$\Pi_2 = c_2\sum_{k,k'=m+1}^{m+n}\beta_{kk'}\rho^{-1/2}|\nu_k\rangle\langle\nu_{k'}|\rho^{-1/2}. \tag{15}$$

In order to determine the constants c_1 and c_2 as well as the matrix elements $\alpha_{kk'}$ and $\beta_{kk'}$, we consider the probability of inconclusive results given by Eq. (2), which is equivalent to $Q=1-c_1-c_2$, where Eq. (11) has been taken into account. It is our aim to find the operators Π_1 and Π_2 described by Eqs. (14) and (15), that minimize Q on the constraint that the positivity conditions expressed in Eq. (1) must hold.

At this point, we can establish the link between the above considerations and the problem of unambiguous discrimination. Since errors are not allowed, the condition $\mathrm{Tr}(\rho_1\Pi_2)=0$ has to be fulfilled for any detection operator Π_2 that unambiguously indicates the presence of the state ρ_2, and Eq. (4) then yields the confidence $C_2=1$. Equation (12) shows that $C_2^{\max}=1$ requires $\nu_{\min}^{(1)}=0$ which implies that $\mathrm{rank}(\rho_1)<d=\mathrm{rank}(\rho)$ [8], where $\rho=\eta_1\rho_1+\eta_2\rho_2$. Hence the support of ρ_2 must contain states that do not belong to the support of ρ_1, or – in other words – the kernel [8] of ρ_1 must not be zero. Similarly, only for $\nu_{\min}^{(2)}=1-\nu_{\max}^{(1)}=0$ the state ρ_1 can be unambiguously distinguished, meaning that ρ_2 must have a nonzero kernel. We thus have rederived the conditions that have to be fulfilled when individual unambiguous discrimination of the two mixed states is feasible.

When the density operators of both states have nonvanishing kernels, the maximum-confidence discrimination is equivalent to unambiguous discrimination. However, when only the kernel of the first state is nonzero while the kernel of the second one vanishes, the usual measurement for unambiguous discrimination delivers an inconclusive result in the presence of the first state. In this case, the measurement scheme of unambiguous discrimination differs from a maximum-confidence measurement since the latter distinguishes also the first state with a certain nonzero confidence, thereby admitting errors to occur.

III. OPTIMIZED MEASUREMENT WITH MINIMUM FAILURE PROBABILITY

A. Solution for states where rank$(\Pi_1,\Pi_2)\leq 1$

1. General solution

In the following, we want to determine the specific discrimination measurement that achieves the maximum confidences $C_1^{\max}$ and $C_2^{\max}$ given by Eq. (12), with the lowest possible overall failure probability Q. First we restrict ourselves to the simplest case, where neither the largest nor the smallest eigenvalue of $\tilde{\rho}_1$, and consequently also of $\tilde{\rho}_2$, are degenerate; that is,

$$\tilde{\rho}_1 = \nu_{\max}^{(1)}|\nu_1\rangle\langle\nu_1| + \nu_{\min}^{(1)}|\nu_2\rangle\langle\nu_2| + \sum_{k=3}^{d}\nu_k^{(1)}|\nu_k\rangle\langle\nu_k|. \quad (16)$$

Using Eqs. (14) and (15) with $m=n=1$, the detection operators warranting the maximum confidences $C_j^{\max}$ for discriminating the states can be written as

$$\Pi_1 = c_1\rho^{-1/2}|\nu_1\rangle\langle\nu_1|\rho^{-1/2} = a|v\rangle\langle v|, \quad (17)$$

$$\Pi_2 = c_2\rho^{-1/2}|\nu_2\rangle\langle\nu_2|\rho^{-1/2} = b|w\rangle\langle w|, \quad (18)$$

where we introduced the normalized states

$$|v\rangle = \frac{\rho^{-1/2}|\nu_1\rangle}{\sqrt{\langle\nu_1|\rho^{-1}|\nu_1\rangle}}, \quad |w\rangle = \frac{\rho^{-1/2}|\nu_2\rangle}{\sqrt{\langle\nu_2|\rho^{-1}|\nu_2\rangle}}. \quad (19)$$

Here $\rho=\eta_1\rho_1+\eta_2\rho_2$, and a and b are some constants that have to be determined. Our task is to minimize the failure probability resulting from Eqs. (2), (17), and (18),

$$Q = 1 - a\langle v|\rho|v\rangle - b\langle w|\rho|w\rangle, \quad (20)$$

on the constraint that the eigenvalues of the operator $\Pi_1+\Pi_2$ are smaller than 1, as required by Eq. (1). A simple calculation shows that the latter eigenvalues are $\lambda_{1/2}=\frac{1}{2}[a+b\pm\sqrt{(a-b)^2+4ab|\langle v|w\rangle|^2}]$ and that they both do not exceed 1 if $a+b\leq 1+ab(1-|\langle v|w\rangle|^2)$. In order to obtain the smallest possible failure probability, we take the equality sign to hold and substitute the resulting expression $b=(1-a)/[1-a(1-|\langle v|w\rangle|^2)]$ into Eq. (20). Upon minimizing the resulting function $Q(a)$, we find that the minimum failure probability is reached when $a=a_o$ and $b=b_o$ with

$$a_o = \frac{1-\sqrt{\frac{\rho_{ww}}{\rho_{vv}}}|\langle v|w\rangle|}{1-|\langle v|w\rangle|^2}, \quad b_o = \frac{1-\sqrt{\frac{\rho_{vv}}{\rho_{ww}}}|\langle v|w\rangle|}{1-|\langle v|w\rangle|^2}, \quad (21)$$

where $\rho_{vv}=\langle v|\rho|v\rangle$ and $\rho_{ww}=\langle w|\rho|w\rangle$. Due to the positivity condition expressed in Eq. (1), the constants a_o and b_o represent a physical solution only in the parameter region where $0\leq a_o,b_o\leq 1$; while outside this region they have to be replaced by their values at the boundaries in order to get the optimum solution. Thus we obtain

$$\begin{aligned} a_{\rm opt}&=1, \quad b_{\rm opt}=0 \quad \text{if } \sqrt{\tfrac{\rho_{ww}}{\rho_{vv}}}\leq|\langle v|w\rangle|,\\ a_{\rm opt}&=a_o, \quad b_{\rm opt}=b_o \quad \text{if } |\langle v|w\rangle|\leq\sqrt{\tfrac{\rho_{ww}}{\rho_{vv}}}\leq\tfrac{1}{|\langle v|w\rangle|}, \\ a_{\rm opt}&=0, \quad b_{\rm opt}=1 \quad \text{if } \sqrt{\tfrac{\rho_{ww}}{\rho_{vv}}}\geq\tfrac{1}{|\langle v|w\rangle|},\end{aligned} \quad (22)$$

determining the optimum detection operators

$$\Pi_1^{\rm opt}=a_{\rm opt}|v\rangle\langle v|, \quad \Pi_2^{\rm opt}=b_{\rm opt}|w\rangle\langle w|, \quad (23)$$

and $\Pi_?^{\rm opt}=I_d-\Pi_1^{\rm opt}-\Pi_2^{\rm opt}$. The minimum failure probability $Q_{\rm opt}$ associated with a measurement achieving the maximum possible confidences $C_1^{\max}=\nu_{\min}^{(1)}$ and $C_2^{\max}=1-\nu_{\min}^{(1)}$ is obtained by substituting Eq. (22) into Eq. (20), yielding

$$Q_{\rm opt}=\begin{cases}1-\rho_{vv} & \text{if } \sqrt{\frac{\rho_{ww}}{\rho_{vv}}}\leq|\langle v|w\rangle|\\ 1-\rho_{ww} & \text{if } \sqrt{\frac{\rho_{ww}}{\rho_{vv}}}\geq\frac{1}{|\langle v|w\rangle|},\end{cases} \quad (24)$$

and, for the condition in middle line of Eq. (22),

$$Q_{\rm opt}=1-\frac{\rho_{vv}+\rho_{ww}-2\sqrt{\rho_{vv}\rho_{ww}}|\langle v|w\rangle|}{1-|\langle v|w\rangle|^2}. \quad (25)$$

When Eq. (24) applies the measurement is a von Neumann measurement, where $\Pi_1^{\rm opt}=|v\rangle\langle v|$, $\Pi_2^{\rm opt}=0$, and $\Pi_?^{\rm opt}=I_d-|v\rangle\langle v|$ if the condition in the upper line is fulfilled, while for the condition in the lower line $\Pi_1^{\rm opt}=0$, $\Pi_2^{\rm opt}=|w\rangle\langle w|$, and $\Pi_?^{\rm opt}=I_d-|w\rangle\langle w|$. On the other hand, when Eq. (25), or the middle line of Eq. (22), respectively, applies and $\langle v|w\rangle\neq 0$, the discrimination is achieved by a generalized measurement since then in Eq. (23) $a_{\rm opt}=a_o<1$ and $b_{\rm opt}=b_o<1$.

In the special case $\langle v|w\rangle=0$, the middle line of Eq. (22) always holds. We then get the operators $\Pi_1^{\rm opt}=|v\rangle\langle v|$, $\Pi_2^{\rm opt}=|w\rangle\langle w|$, and $\Pi_?^{\rm opt}=I_d-|v\rangle\langle v|-|w\rangle\langle w|$ which describe a von Neumann measurement with the resulting failure probability $Q_{\rm opt}=1-\rho_{vv}-\rho_{ww}$. For $d=2$, this means that $\Pi_?^{\rm opt}=0$ and inconclusive results do not occur.

It is interesting to relate the maximum-confidence measurement with minimum failure probability to the measurement strategy of minimum-error discrimination [4], where $\Pi_?=0$. Since in this case $\Pi_2=I_d-\Pi_1$, the error probability $P_{\rm err}=\eta_1{\rm Tr}(\rho_1\Pi_2)+\eta_2{\rm Tr}(\rho_2\Pi_1)=1-\eta_1{\rm Tr}(\rho_1\Pi_1)-\eta_2{\rm Tr}(\rho_2\Pi_2)$ can be written as

$$P_{\rm err}=\eta_1+{\rm Tr}(\Lambda\Pi_1) \quad \text{with} \quad \Lambda=\eta_2\rho_2-\eta_1\rho_1, \quad (26)$$

or $\Lambda=\rho-2\eta_1\rho_1$, respectively, due to Eq. (3). The error probability takes its minimum $P_E=\frac{1}{2}(1-{\rm Tr}|\Lambda|)$ [4], when $\Pi_1=\Pi_1^E$, where

$$\Pi_1^E=\sum_{i(\lambda_i<0)}|\lambda_i\rangle\langle\lambda_i| \quad \text{with} \quad \Lambda=\sum_{i=1}^{d}\lambda_i|\lambda_i\rangle\langle\lambda_i| \quad (27)$$

and $\langle\lambda_i|\lambda_j\rangle=\delta_{ij}$ [22,23]. In other words, in a minimum-error measurement Π_1^E projects onto the subspace spanned by all eigenstates of Λ that belong to negative eigenvalues λ_i, while $\Pi_2^E=I_d-\Pi_1^E$. In the next paragraph, we derive the conditions that have to be fulfilled when discrimination with the minimum error probability is achieved by the same measurement like the maximum-confidence discrimination.

Before proceeding, we note that our general solution given by Eqs. (22)–(25) comprises the optimum unambiguous discrimination of two arbitrary mixed quantum states with one-dimensional kernels [9]. This case arises when in Eq. (16) $\nu_{\max}^{(1)}=1$ and $\nu_{\min}^{(1)}=0$. Indeed, since because of Eq. (9) then also $\nu_{\min}^{(2)}=1-\nu_{\max}^{(1)}=0$, it follows that the operators $\tilde{\rho}_1$ and $\tilde{\rho}_2$ and consequently also the supports of the operators ρ_1 and ρ_2 have the rank $d-1$ if ρ has the rank d, the two kernels thus being one dimensional.

2. Discrimination of two mixed qubit states

As an important application, we consider the maximum-confidence discrimination of two arbitrary qubit states ρ_1 and ρ_2 that are defined in the same two-dimensional Hilbert space and occur with the prior probabilities η_1 and $\eta_2=1-\eta_1$, respectively. Equation (16) then takes the form

$$\tilde{\rho}_1 = \eta_1\rho^{-1/2}\rho_1\rho^{-1/2} = \nu_{\max}^{(1)}|\nu_1\rangle\langle\nu_1| + \nu_{\min}^{(1)}|\nu_2\rangle\langle\nu_2| \quad (28)$$

and determines the maximum confidences $C_1^{\max}=\nu_{\max}^{(1)}$ and $C_2^{\max}=1-\nu_{\min}^{(1)}$, as well as the orthonormal states $|\nu_1\rangle$ and $|\nu_2\rangle$. Since $\rho=\eta_1\rho_1+\eta_2\rho_2$ is a rank-two operator, the matrix elements of ρ^{-1} can be easily expressed by the matrix elements of ρ. Equations (24) and (25), characterizing the minimum failure probability achievable in maximum-confidence discrimination, are then transformed into

$$Q_{\rm opt} = \begin{cases} 1-\dfrac{\det(\rho)}{\langle\nu_2|\rho|\nu_2\rangle} & \text{if } |\langle\nu_1|\rho|\nu_2\rangle| \geq \langle\nu_2|\rho|\nu_2\rangle \\ 1-\dfrac{\det(\rho)}{\langle\nu_1|\rho|\nu_1\rangle} & \text{if } |\langle\nu_1|\rho|\nu_2\rangle| \geq \langle\nu_1|\rho|\nu_1\rangle \\ 2|\langle\nu_1|\rho|\nu_2\rangle| & \text{else.} \end{cases} \quad (29)$$

Here the relation $\langle\nu_1|\rho|\nu_1\rangle+\langle\nu_2|\rho|\nu_2\rangle=\mathrm{Tr}\,\rho=1$ has been used, and $\det(\rho)=\langle\nu_1|\rho|\nu_1\rangle\langle\nu_2\rho|\nu_2\rangle-|\langle\nu_1|\rho|\nu_2\rangle|^2$. The optimum detection operators are determined by

$$\begin{aligned} &a_{\rm opt}=1, \quad b_{\rm opt}=0 \quad \text{if } |\langle\nu_1|\rho|\nu_2\rangle| \geq \langle\nu_1|\rho|\nu_1\rangle \\ &a_{\rm opt}=0, \quad b_{\rm opt}=1 \quad \text{if } |\langle\nu_1|\rho|\nu_2\rangle| \geq \langle\nu_2|\rho|\nu_2\rangle \\ &a_{\rm opt}=a_o, \quad b_{\rm opt}=b_o \quad \text{else,} \end{aligned}$$

where

$$a_o = \frac{1-\frac{|\langle\nu_1|\rho|\nu_2\rangle|}{\langle\nu_1|\rho|\nu_1\rangle}}{1-\frac{|\langle\nu_1|\rho|\nu_2\rangle|^2}{\langle\nu_1|\rho|\nu_1\rangle\langle\nu_2|\rho|\nu_2\rangle}}, \quad b_o = \frac{1-\frac{|\langle\nu_1|\rho|\nu_2\rangle|}{\langle\nu_2|\rho|\nu_2\rangle}}{1-\frac{|\langle\nu_1|\rho|\nu_2\rangle|^2}{\langle\nu_1|\rho|\nu_1\rangle\langle\nu_2|\rho|\nu_2\rangle}} \quad (30)$$

and they follow from $\Pi_1^{\rm opt}=a_{\rm opt}|v\rangle\langle v|$ and $\Pi_2^{\rm opt}=b_{\rm opt}|w\rangle\langle w|$, where $|u\rangle$ and $|v\rangle$ are defined in Eq. (19).

The special case $\langle\nu_1|\rho|\nu_2\rangle=0$ or $\langle v|w\rangle=0$, respectively, deserves a separate discussion. For $d=2$ it implies that $|\nu_1\rangle$ and $|\nu_2\rangle$ are eigenstates of ρ, or, equivalently, $[\rho,\tilde{\rho}_1]=0$ and thus also $[\rho_1,\rho_2]=0$. Equation (19) then reduces to $|v\rangle=|\nu_1\rangle$, $|w\rangle=|\nu_2\rangle$, and we arrive at

$$\Pi_1^{\rm opt}=|\nu_1\rangle\langle\nu_1|, \quad \Pi_2^{\rm opt}=|\nu_2\rangle\langle\nu_2|, \quad \Pi_?^{\rm opt}=0. \quad (31)$$

Let us relate this measurement to the minimum-error measurement. For $[\rho_1,\rho_2]=0$ and $d=2$, we find from Eqs. (26), (28), and (12) that $\Lambda=\lambda_1|\nu_1\rangle\langle\nu_1|+\lambda_2|\nu_2\rangle\langle\nu_2|$ with

$$\lambda_1=\langle\nu_1|\rho|\nu_1\rangle(1-2C_1^{\max}), \quad \lambda_2=\langle\nu_2|\rho|\nu_2\rangle(2C_2^{\max}-1), \quad (32)$$

since $\Lambda=\rho(1-2\tilde{\rho}_1)$ for $[\rho,\rho_1]=0$. From Eq. (27) it becomes obvious that for $C_1^{\max}>0.5$ and $C_2^{\max}>0.5$ the detection operators for minimum-error discrimination are $\Pi_1^E=|\nu_1\rangle\langle\nu_1|$ and $\Pi_2^E=|\nu_2\rangle\langle\nu_2|$ which coincide with the optimum detection operators in Eq. (31). On the other hand, if either $C_1^{\max}$ or $C_2^{\max}$ is smaller than 0.5, we conclude with the help of Eq. (13) that either $\Pi_1^E=0$ or $\Pi_1^E=I_d$. This means that the minimum probability of errors arises without any measurement at all, just by always guessing the presence of the most probable state [24].

As an example for $[\rho_1,\rho_2]=0$ or $\langle\nu_1|\rho|\nu_2\rangle=0$, respectively, we treat the discrimination between the completely mixed qubit state $\rho_1=I_2/2$, occurring with the prior probability $\eta_1=1-\eta_2$, and a given mixed qubit state ρ_2, occurring with the prior probability η_2. We then have to distinguish between the states

$$\rho_1=\frac{I_2}{2}, \quad \rho_2=p|\psi\rangle\langle\psi|+(1-p)\frac{I_2}{2}, \quad (33)$$

with $0<p\leq 1$, where we took into account that any mixed qubit state ρ_2 can always be written in the form given in Eq. (33). Loosely speaking, the parameter p characterizes the purity of the qubit state ρ_2, since for $p=1$ it is pure and for $p=0$ it is completely mixed. By applying Eqs. (12) and (28)–(30), we obtain the maximum confidences and the associated minimum failure probability for discriminating the states,

$$C_1^{\max}=\frac{1-\eta_2}{1-p\eta_2}, \quad C_2^{\max}=\frac{\eta_2(1+p)}{1+p\eta_2}, \quad Q^{\rm opt}=0. \quad (34)$$

The corresponding optimized measurement is the projection measurement with

$$\Pi_1^{\rm opt}=|\psi^\perp\rangle\langle\psi^\perp|, \quad \Pi_2^{\rm opt}=|\psi\rangle\langle\psi|, \quad \Pi_?^{\rm opt}=0, \quad (35)$$

where $|\psi^\perp\rangle$ is the normalized state that is orthogonal to $|\psi\rangle$, that is, $I_2=|\psi\rangle\langle\psi|+|\psi^\perp\rangle\langle\psi^\perp|$. Using Eq. (32) we find that for $(2+p)^{-1}<\eta_2<(2-p)^{-1}$ these detection operators are identical with those of the minimum-error measurement. When η_2 lies outside this range, however, the minimum probability of errors is obtained when simply the state with the largest prior probability is guessed to be present, without performing a measurement.

In the special case $p=1$, the example given in Eq. (33) corresponds to the discrimination between the pure state $\rho_2=|\psi\rangle\langle\psi|$ and a mixed state ρ_1, a problem that is also known as quantum state filtering and that has been treated with respect to minimum-error discrimination [25], optimum unambiguous discrimination [26,27], and maximum-confidence discrimination [20]. When $|\psi\rangle$ lies within the support of ρ_1, the measurement for optimum unambiguous discrimination is a von Neumann measurement with $\Pi_1=|\psi^\perp\rangle\langle\psi^\perp|$, $\Pi_2=0$, and $\Pi_?=|\psi\rangle\langle\psi|$ [27]. In our case, it yields the failure probability $Q=\frac{1}{2}\eta_1+\eta_2$ and the confidences $C_1=1$ and $C_2=0$, in contrast to the measurement described by Eq. (35), where for $p=1$ we get $Q=0$, $C_1^{\max}=1$, and $C_2^{\max}=2\eta_2/(1+\eta_2)$.

Our second example refers to the case $[\rho_1,\rho_2]\neq 0$ or $\langle\nu_1|\rho|\nu_2\rangle\neq 0$, respectively. We suppose equal prior probabilities of the two states and take also their purities to be the same, assuming that

$$\rho_j = p|\psi_j\rangle\langle\psi_j| + (1-p)\frac{I_2}{2} \quad (j=1,2), \tag{36}$$

with $0\le\langle\psi_1|\psi_2\rangle<1$ and $0<p\le 1$. Without lack of generality, we put $I_2=|0\rangle\langle 0|+|1\rangle\langle 1|$ and

$$|\psi_{1/2}\rangle = \cos\frac{\gamma}{2}|0\rangle \pm \sin\frac{\gamma}{2}|1\rangle \quad (0<\gamma<\pi/2), \tag{37}$$

where $|0\rangle$ and $|1\rangle$ are two orthonormal basis states and $\cos\gamma=\langle\psi_1|\psi_2\rangle$. With $\eta_1=\eta_2=0.5$, Eqs. (28)–(30) together with Eqs. (12) and (23) yield the eigenstates of $\tilde{\rho}_1$, $|\nu_{1,2}\rangle=\frac{1}{\sqrt{2}}(|0\rangle\pm|1\rangle)$, and the maximum confidences and associated minimum failure probabilities

$$C_1^{\max}=C_2^{\max}=\frac{1}{2}+\frac{p\sin\gamma}{2\sqrt{1-p^2\cos^2\gamma}}, \quad Q_{\text{opt}}=p\cos\gamma, \tag{38}$$

as well as the optimum detection operators

$$\Pi_1^{\text{opt}}=\frac{|v\rangle\langle v|}{1+p\cos\gamma}, \quad \Pi_2^{\text{opt}}=\frac{|w\rangle\langle w|}{1+p\cos\gamma},$$

and $\Pi_?^{\text{opt}}=I_2-\Pi_1^{\text{opt}}-\Pi_2^{\text{opt}}$. Here $|v\rangle$ and $|w\rangle$ are the normalized states

$$|v/w\rangle=\frac{1}{\sqrt{2}}(\sqrt{1-p\cos\gamma}|0\rangle \pm \sqrt{1+p\cos\gamma}|1\rangle), \tag{39}$$

which are nonorthogonal since $p\neq 0$. Clearly, the detection operators are not projectors and the measurement therefore is a generalized measurement. For $p=1$, it reduces to the well-known measurement for the optimum unambiguous discrimination of two equally probable nonorthogonal pure states [5] and the maximum confidences are equal to 1, while their limiting value for $p\to 0$ is equal to 0.5. For fixed p, the minimum failure probability associated with the measurement decreases with growing angle γ (cf. Fig. 1), while the maximum confidences increase and tend to $(1+p)/2$ for $\gamma\to\pi/2$.

By exploiting Eq. (27), we find that the minimum-error discrimination of the two equiprobable states defined in Eq. (36) is achieved by a projective measurement with $\Pi_{1/2}^E=|\nu_{1/2}\rangle\langle\nu_{1/2}|$, where again $|\nu_{1,2}\rangle=\frac{1}{\sqrt{2}}(|0\rangle\pm|1\rangle)$. Using these detection operators in Eq. (4), we get the confidences $C_1^E=C_2^E=\frac{1}{2}(1+p\sin\gamma)$ in a minimum-error measurement which are clearly smaller than the confidences given in Eq. (38) and arising from a maximum-confidence measurement.

B. Case of higher-rank detection operators

When the rank of the detection operators represented by Eqs. (14) and (15) is larger than one, minimizing the probability Q of inconclusive results is in general a highly nontrivial optimization problem because the positivity constraints in Eq. (1) impose a set of complicated conditions. However, when the given density operators allow to separate the problem into independent optimizations in orthogonal two-dimensional subspaces of the joint Hilbert space, an analytical solution can be easily obtained by applying the results for discriminating two mixed qubit states. This is analogous to the separation into orthogonal two-dimensional subspaces that has been used previously for investigating the optimum unambiguous discrimination of two mixed states [13–15]. In the following we treat a simple example.

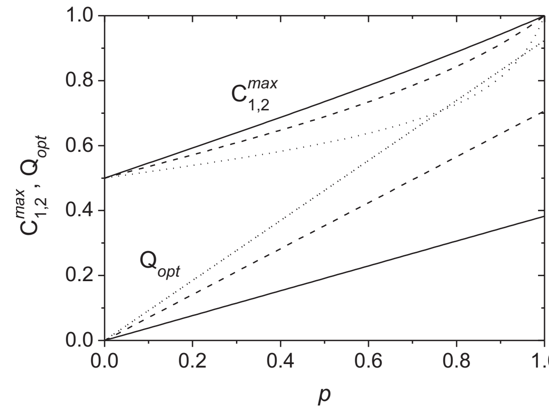

FIG. 1. Maximum confidence $C_{1,2}^{\max}$ and the associated minimum failure probability Q_{opt} for discriminating two equally probable qubit states having the same purity p [cf. Eq. (36)]. The parameters are $\gamma=3\pi/8$ (full line), $\gamma=\pi/4$ (dashed line), and $\gamma=\pi/8$ (dotted line) [cf. Eq. (37)].

We consider the discrimination of two mixed states defined in a d-dimensional joint Hilbert space with d being an even number and described by the density operators,

$$\rho_j=\frac{2p}{d}\sum_{k=1}^{d/2}|r_k^{(j)}\rangle\langle r_k^{(j)}| + (1-p)\frac{I_d}{d} \quad (j=1,2), \tag{40}$$

with $0<p\le 1$ and $|r_k^{(1,2)}\rangle=\cos\frac{\gamma_k}{2}|0\rangle_k\pm\sin\frac{\gamma_k}{2}|1\rangle_k$, where for $k\neq k'$ any two basis states labeled by k and k' are mutually orthogonal. The identity operator then takes the form $I_d=\sum_{k=1}^{d/2}(|0\rangle_k\langle 0|_k+|1\rangle_k\langle 1|_k)$. For simplicity, we suppose equal prior probabilities of the two states, $\eta_1=\eta_2=\frac{1}{2}$. We then get $\tilde{\rho}_1=\frac{1}{2}\rho^{-1/2}\rho_1\rho^{-1/2}$ with the spectral decomposition

$$\tilde{\rho}_1=\sum_{k=1}^{d/2}(\nu_k^{(+)}|\nu_k^{(+)}\rangle\langle\nu_k^{(+)}| + \nu_k^{(-)}|\nu_k^{(-)}\rangle\langle\nu_k^{(-)}|), \tag{41}$$

where the eigenvalues and eigenstates are

$$\nu_k^{(\pm)}=\frac{1}{2}\pm\frac{p\sin\gamma_k}{2\sqrt{1-p^2\cos^2\gamma_k}}, \quad |\nu_k^{(\pm)}\rangle=\frac{|0\rangle_k\pm|1\rangle_k}{\sqrt{2}}, \tag{42}$$

with $1\le k\le d/2$. If we denote the largest of the angles γ_k by γ, we obtain with the help of Eq. (12) the maximum confidences

$$C_1^{\max} = C_2^{\max} = \frac{1}{2} + \frac{p \sin\gamma}{2\sqrt{1-p^2\cos^2\gamma}} \quad (\gamma = \max\{\gamma_k\}). \tag{43}$$

In the special case $p=1$, where $C_1^{\max}=C_2^{\max}=1$, the maximum-confidence discrimination with minimum failure probability is equivalent to optimum unambiguous discrimination. The latter measurement has been derived previously and yields for our example the minimum failure probability $Q_{\rm opt}^{(p=1)}=\frac{2}{d}\Sigma_{k=1}^{d/2}\cos\gamma_k$ [14,15]. For $p=1$ the operator $\tilde{\rho}_1$ has only the eigenvalues 0 and 1, each being $d/2$-fold degenerate, and the optimum detection operators Π_1 and Π_2 therefore have the rank $d/2$.

Here we are interested in the case that the largest eigenvalue of $\tilde{\rho}_1$ may be degenerate also for $p<1$, thus leading to higher-rank detection operators for the maximum-confidence discrimination. We assume that

$$\gamma_k = \gamma \quad \text{for } k=1,\ldots,m, \tag{44}$$

$$\gamma_k < \gamma \quad \text{for } k=m+1,\ldots,\tfrac{d}{2}. \tag{45}$$

Using the eigenstates of $\tilde{\rho}_1$ and the explicit expression resulting for $\rho=\frac{1}{2}(\rho_1+\rho_2)$, the general ansatz for the detection operators in maximum-confidence discrimination given by Eqs. (14) and (15) can be rewritten as

$$\Pi_1 = \sum_{k,k'=1}^{m} a_{kk'}|v_k^{(\gamma)}\rangle\langle v_{k'}^{(\gamma)}|, \quad \Pi_2 = \sum_{k,k'=1}^{m} b_{kk'}|w_k^{(\gamma)}\rangle\langle w_{k'}^{(\gamma)}|, \tag{46}$$

where in analogy to Eq. (39),

$$|v_k^{(\gamma)}/w_k^{(\gamma)}\rangle = \sqrt{\frac{1-p\cos\gamma}{2}}|0\rangle_k \pm \sqrt{\frac{1+p\cos\gamma}{2}}|1\rangle_k. \tag{47}$$

The expression for the failure probability [Eq. (2)] then yields $Q=1-\frac{1}{d}(1-p^2\cos^2\gamma)\Sigma_{k=1}^{m}(a_{kk}+b_{kk})$. Since due to our special choice of the density operators, the pairs of states $\{|v_k^{(\gamma)}\rangle,|w_k^{(\gamma)}\rangle\}$ with different values of k span mutually orthogonal two-dimensional subspaces; the minimization of Q under the positivity constraints for the detection operators can be separated into m independent two-dimensional problems. We find that Q takes its minimum $Q_{\rm opt}$, when in Eq. (46) $a_{kk'}=a_{kk}\delta_{kk'}$ and $b_{kk'}=b_{kk}\delta_{kk'}$, and in analogy to the derivation of Eq. (38) we arrive at

$$\Pi_1^{\rm opt} = \sum_{k=1}^{m} \frac{|v_k^{(\gamma)}\rangle\langle v_k^{(\gamma)}|}{1+p\cos\gamma}, \quad \Pi_2^{\rm opt} = \sum_{k=1}^{m} \frac{|w_k^{(\gamma)}\rangle\langle w_k^{(\gamma)}|}{1+p\cos\gamma}. \tag{48}$$

From these operators we get $Q_{\rm opt}=1-\frac{2m}{d}(1-p\cos\gamma)$. Clearly, for fixed m the maximum confidences given in Eq. (43) require a minimum overall failure probability $Q_{\rm opt}$ which grows with increasing dimensionality d.

We still remark that in certain cases it might be desirable to perform a different measurement where all two-dimensional subspaces contribute to the conclusive results, yielding somewhat reduced confidences but a considerably lower failure probability. In particular, for

$$\Pi_1^{av} = \sum_{k=1}^{d/2} \frac{|v_k^{(\gamma_k)}\rangle\langle v_k^{(\gamma_k)}|}{1+p\cos\gamma_k}, \quad \Pi_2^{av} = \sum_{k=1}^{d/2} \frac{|w_k^{(\gamma_k)}\rangle\langle w_k^{(\gamma_k)}|}{1+p\cos\gamma_k}, \tag{49}$$

we obtain from Eqs. (2) and (4) the probability of inconclusive results $Q_{av}=\frac{2p}{d}\Sigma_{k=1}^{d/2}\cos\gamma_k$ and the confidences

$$C_1^{av} = C_2^{av} = \frac{1}{2} + \frac{p\sum_{k=1}^{d/2}\sin\gamma_k\sqrt{\frac{1-p\cos\gamma_k}{1+p\cos\gamma_k}}}{2\sum_{k=1}^{d/2}(1-p\cos\gamma_k)}. \tag{50}$$

In general, whenever eigenvalues other than the smallest and largest one occur in the spectral decomposition of the operator $\tilde{\rho}_1$, it might be worthwhile in some cases to replace the maximum confidence strategy by a balanced strategy yielding a somewhat smaller confidence at a drastically reduced probability of inconclusive results.

IV. DISCUSSION AND CONCLUSIONS

The measurement strategy of maximum confidence discrimination is related to another optimization strategy that was considered by Fiurášek and Ježek [28] for mixed states and that was introduced already earlier for pure states [29]. In this scheme, the average success probability to get a correct result $P_S=\Sigma_j\eta_j{\rm Tr}(\rho_j\Pi_j)$ is maximized for a given probability $Q=1-\Sigma_j{\rm Tr}(\rho\Pi_j)$ of inconclusive results. In addition, the so-called relative success rate $P_{\rm RS}=P_S/(1-Q)$ is considered [28]. Introducing $f_j={\rm Tr}(\rho\Pi_j)/(1-Q)$, where $\Sigma_j f_j=1$, and using Eq. (4), it follows that $P_{\rm RS}=\Sigma_j f_j C_j$. Hence the largest possible value of $P_{\rm RS}$ is equal to the largest of the different maximum confidences $C_j^{\max}$, $P_{\rm RS}^{\max}={\rm Max}_j\{C_j^{\max}\}$. This value is obtained in a measurement where $f_j=0$ or $\Pi_j=0$, respectively, for any state ρ_j with $C_j^{\max}<{\rm Max}_j\{C_j^{\max}\}$ which then yields an inconclusive result. For two equiprobable qubit states with the same purity given by Eq. (36), the maximum relative success rate $P_{\rm RS}^{\max}$ has been calculated in Ref. [28]. As expected from the above considerations, it coincides with the maximum confidences $C_1^{\max}=C_2^{\max}$ given in Eq. (38).

To summarize, we investigated the measurement for discriminating two mixed quantum states with maximum possible confidence for each of the two different conclusive outcomes, thereby keeping the overall probability of inconclusive results as small as possible. When the density operators of both states have nonvanishing kernels, the measurement is equivalent to optimum unambiguous discrimination. When one of the kernels is zero, however, the optimum unambiguous discrimination always fails for one of the states and thus differs from the optimized maximum-confidence measurement discriminating both states with a certain nonzero confidence. Provided that the rank of the detection operators associated with the two conclusive outcomes does not exceed unity, we obtained a general solution for the optimum measurement valid for arbitrary prior probabilities of the states. It is given by Eqs. (22)–(25) and represents our main result. As an application, we considered the discrimination of two mixed qubit states. Moreover, for the case of higher-rank detection operators, we derived a solution for particular states.

ACKNOWLEDGMENTS

The author would like to thank Janos Bergou (Hunter College, New York) for many useful discussions and for the hospitality extended to her during a visit in New York. Discussions with Oliver Benson (Humboldt-University, Berlin) and Mark Hillery (Hunter College, New York) are also gratefully acknowledged.

[1] A. Chefles, Contemp. Phys. **41**, 401 (2000).
[2] J. A. Bergou, U. Herzog, and M. Hillery, Lecture Notes in Physics Vol. 649 (Springer, Berlin, 2004), pp. 417–465.
[3] S. M. Barnett and S. Croke, Adv. Opt. Photon. **1**, 238 (2009).
[4] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).
[5] I. D. Ivanovic, Phys. Lett. A **123**, 257 (1987); D. Dieks, *ibid.* **126**, 303 (1988), A. Peres, *ibid.* **128**, 19 (1988).
[6] G. Jaeger and A. Shimony, Phys. Lett. A **197**, 83 (1995).
[7] A. Chefles, Phys. Lett. A **239**, 339 (1998).
[8] The support of a density operator is the Hilbert space spanned by its eigenvectors with nonzero eigenvalues. The kernel is the subspace orthogonal to the support and the rank of a density operator is the dimension of the support.
[9] T. Rudolph, R. W. Spekkens, and P. S. Turner, Phys. Rev. A **68**, 010301(R) (2003).
[10] P. Raynal, N. Lütkenhaus, and S. J. van Enk, Phys. Rev. A **68**, 022308 (2003).
[11] Y. C. Eldar, M. Stojnic, and B. Hassibi, Phys. Rev. A **69**, 062318 (2004).
[12] Y. Feng, R. Duan, and M. Ying, Phys. Rev. A **70**, 012308 (2004).
[13] U. Herzog and J. A. Bergou, Phys. Rev. A **71**, 050301(R) (2005).
[14] J. A. Bergou, E. Feldman, and M. Hillery, Phys. Rev. A **73**, 032107 (2006).
[15] U. Herzog, Phys. Rev. A **75**, 052309 (2007).
[16] X.-F. Zhou, Y.-S. Zhang, and G.-C. Guo, Phys. Rev. A **75**, 052314 (2007).
[17] P. Raynal and N. Lütkenhaus, Phys. Rev. A **76**, 052322 (2007).
[18] M. Kleinmann, H. Kampermann, and D. Bruß, e-print arXiv:0807.3923.
[19] S. Croke, E. Andersson, S. M. Barnett, C. R. Gilson, and J. Jeffers, Phys. Rev. Lett. **96**, 070401 (2006).
[20] S. Croke, E. Andersson, and S. M. Barnett, Phys. Rev. A **77**, 012113 (2008).
[21] J. Preskill, *Quantum Information and Computation*, Lecture Notes in Physics Vol. 229 (Cambridge University Press, Cambridge, UK, 1998).
[22] U. Herzog, J. Opt. B: Quantum Semiclassical Opt. **6**, S24 (2004).
[23] U. Herzog and J. A. Bergou, Phys. Rev. A **70**, 022302 (2004).
[24] K. Hunter, Phys. Rev. A **68**, 012306 (2003).
[25] U. Herzog and J. A. Bergou, Phys. Rev. A **65**, 050305(R) (2002).
[26] Y. Sun, J. A. Bergou, and M. Hillery, Phys. Rev. A **66**, 032315 (2002).
[27] J. A. Bergou, U. Herzog, and M. Hillery, Phys. Rev. Lett. **90**, 257901 (2003); Phys. Rev. A **71**, 042314 (2005).
[28] J. Fiurášek and M. Ježek, Phys. Rev. A **67**, 012321 (2003).
[29] A. Chefles and S. M. Barnett, J. Mod. Opt. **45**, 1295 (1998).