

**Optimal single-shot strategies for discrimination of quantum measurements**Michal Sedlák<sup>1,2</sup> and Mário Ziman<sup>2,3</sup><sup>1</sup>*Department of Optics, Palacký University, 17. listopadu 1192/12, CZ-771 46 Olomouc, Czech Republic*<sup>2</sup>*Institute of Physics, Slovak Academy of Sciences, Dúbravská cesta 9, 845 11 Bratislava, Slovakia*<sup>3</sup>*Faculty of Informatics, Masaryk University, Botanická 68a, 60200 Brno, Czech Republic*

(Received 4 August 2014; published 10 November 2014)

We study discrimination of  $m$  quantum measurements in the scenario when the unknown measurement with  $n$  outcomes can be used only once. We show that ancilla-assisted discrimination procedures provide a nontrivial advantage over simple (ancilla-free) schemes for perfect distinguishability and we prove that inevitably  $m \leq n$ . We derive necessary and sufficient conditions of perfect distinguishability of general binary measurements. We show that the optimization of the discrimination of projective qubit measurements and their mixtures with white noise is equivalent to the discrimination of specific quantum states. In particular, the optimal protocol for discrimination of projective qubit measurements with fixed failure rate (exploiting maximally entangled test state) is described. While minimum-error discrimination of two projective qubit measurements can be realized without any need of entanglement, we show that discrimination of three projective qubit measurements requires a bipartite probe state. Moreover, when the measurements are not projective, the non-maximally entangled test states can outperform the maximally entangled ones. Finally, we rephrase the unambiguous discrimination of measurements as quantum key distribution protocol.

DOI: [10.1103/PhysRevA.90.052312](https://doi.org/10.1103/PhysRevA.90.052312)

PACS number(s): 03.67.-a, 03.65.Ta, 03.65.Wj

**I. INTRODUCTION**

Quantum theory is statistical, hence, any distinction in the performance of quantum devices is based on statistical reasoning. However, if the set of possibilities is restricted to a finite number of alternatives (communication being the best example), the observations of individual experimental outcomes represent a nontrivial information. For example, in communication Alice encodes a letter  $a$  by selecting a preagreed preparation procedure associated with a quantum state  $\rho_a$ . In each communication round, Bob is trying to estimate which preparation was selected by Alice to recover the submitted letter  $a$ . How often Alice and Bob succeed is the research subject of optimal state discrimination (see, for instance, Chap. 11 of [1] or [2] for an overview).

Since the seminal works of Holevo and Helstrom [3,4], a lot of research effort was invested on the various aspects of discrimination problems, finding its applications in quantum communication, quantum cryptography, but also in quantum computation. For instance, in its essence the famous Grover's search algorithm [5] solves the question of optimal and efficient discrimination of quantum oracles representing the database elements. No doubt the discrimination problems represent one of the central conceptual questions of quantum physics with both practical and foundational implications. The solutions provide a natural quantitative measure of difference or similarity of quantum devices with clearly justified operational meaning. Recently, the variant of discrimination problem was used to argue the philosophical objectivity of quantum wave function [6]. It might seem that investigation of single-shot scenarios for discrimination of measurements is a bit artificial question because measurements can be naturally employed many times. Certainly, the same can be said (typically) also for sources of quantum systems, nevertheless, in this case the concept of a single copy is considered to be well established. Let us point out that there are examples of measurements for which single copy makes

a perfect sense. For example, a photographic plate (being part of Stern-Gerlach apparatus) cannot be (strictly speaking) employed again in the measurement. Moreover, except of the mentioned operational (but mathematical) motivation for discrimination problems, one meets with the discrimination questions in cases when random selection of devices plays some role. This typically happens in quantum communication and computation applications.

In comparison with the case of states [7–10] and processes [11–21], the discrimination of quantum measurements is rather unexplored. In Ref. [22], authors have shown that any pair of projective measurements can be perfectly discriminated in a finite number of runs. The question of discrimination of measurements with unlabeled outcomes has been addressed in Refs. [23,24] and experimental realizations of measurement discrimination protocols have been reported in Refs. [25,26].

This paper addresses the question of optimal discrimination of quantum measurements. It is organized as follows. The problem is formulated in Secs. II and III. In Sec. IV, we study the conditions of perfect distinguishability. Further, we continue with discrimination of quantum filters in Sec. V. We reduce this problem to discrimination of projective measurements which is then investigated in Sec. VI. Finally, in Sec. VII we solve unambiguous discrimination of two trine measurements demonstrating that non-maximally entangled states can outperform maximally entangled ones. Section VIII contains the summary of the results.

**II. PROBLEM SETTING**

Suppose we are given a measurement device we want to identify, however, we can use it only once. We will assume that a nontrivial prior knowledge on potential alternatives is given. In the simplest case we are distinguishing among two alternatives: the apparatus performs either a measurement

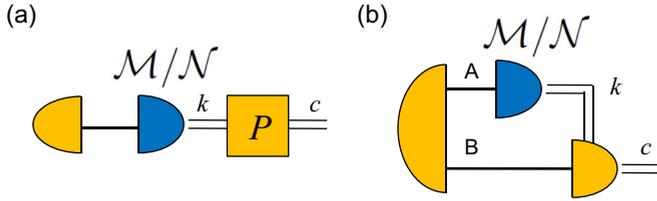


FIG. 1. (Color online) Possible approaches to discrimination of a quantum measurement. Part (a) depicts the simple discrimination scheme, in which a test state  $\rho$  is prepared, afterward measured with the unknown measurement device and based on the obtained outcome  $k$  the measurement device is identified. Part (b) depicts the general discrimination scheme, in which a bipartite state is prepared; one part of it is measured with the unknown measurement device and the other part by a known ancillary measurement chosen conditionally on the actual outcome  $k$  recorded in the unknown measurement.

$\mathcal{M}$  or a measurement  $\mathcal{N}$  with *a priori* probabilities  $\eta_{\mathcal{M}}$ ,  $\eta_{\mathcal{N}}$ , respectively. Our goal is to design a test that would (optimally) identify the unknown measurement device. We will distinguish between two types of tests: *simple (ancilla-free)* and *general (ancilla-assisted)* experimental setting (see Fig. 1). The simple scheme consists of the preparation of a test state probing the unknown measurement device and of a post-processing assignment of a conclusion for each individual measurement outcome. In contrast, the most general (ancilla-assisted) scheme begins with a preparation of a bipartite quantum state part of which is then measured by the unknown measurement device. The obtained outcome  $k$  is used to determine the measurement of the remaining part of the bipartite system. Finally, based on the recorded outcomes the guess on the identity of the unknown measurement device is made.

If the measurements  $\mathcal{M}$ ,  $\mathcal{N}$  are not in a specific mutual relation allowing for perfect discrimination, it is obvious that all the conclusions can not be always valid. The way how the imperfections are evaluated and processed is then used for definition of optimality. Each conclusion is characterized by the probability of being wrong (error probability). To evaluate the reliability of the conclusions of the discrimination test, we use the following three quantities: (i) *error probability*  $p_e$  defined as the average error probability (over all conclusive outcomes); (ii) *failure probability*  $p_f$  given as a total probability of inconclusive outcomes; (iii) *success probability*  $p_s = 1 - p_f - p_e$ .

When all the outcomes are conclusive, i.e.,  $p_f = 0$ , and we optimize the average success probability, then we speak about *minimum-error* discrimination strategy. On the other side of the spectrum of discrimination problems we find the *unambiguous* discrimination, for which  $p_e = 0$ , but inconclusive outcomes are possible. The optimality is achieved when  $p_f$  is minimized. In this paper, we will consider also intermediate variations of the discrimination problems, which include the mentioned strategies as “extremal” cases. In particular, we will consider maximization of average success probability  $p_s$  for a fixed value of the failure probability  $p_f$ .

In such case we say we implement *discrimination with fixed failure rate*. We say the measurements can be *perfectly discriminated* if  $p_f$  and  $p_e$  can vanish simultaneously.

### III. MATHEMATICAL FRAMEWORK

Let us denote by  $\mathcal{H}_d$  the  $d$ -dimensional Hilbert space of the quantum system under consideration. The measurement device  $\mathcal{M}$  is a positive-operator-valued measure (POVM) assigning a positive operator  $M_j$  for each  $j \in \Omega = \{1, \dots, n\}$  (we will not consider measurements with infinite or uncountable number of outcomes). We can represent any measurement as a specific measure-and-prepare channel  $\mathcal{M}$ :

$$\mathcal{M}(\rho) = \sum_j \text{tr}[M_j \rho] |j\rangle\langle j| \quad (1)$$

mapping states of  $\mathcal{H}_d$  into diagonal density operators on  $\mathcal{H}_n$  (probability distributions on  $\Omega$ ), where  $\mathcal{H}_n$  is  $n$ -dimensional Hilbert space spanned by fixed orthonormal basis  $\{|j\rangle\}$ . Further, we will assume that all the measurements we want to discriminate have the same number of outcomes. Using this representation of measurements, the problem of discrimination can be reformulated as a special case of (quantum-classical) channel discrimination, hence, the general results obtained for (single-shot) discrimination of channels can be directly translated into the language of measurements.

In what follows, we demonstrate mathematical formulation of a discrimination problem for its simplest version when the measurement device is guaranteed to be one of two known alternatives. Generalization to any number of measurements is straightforward. Let us denote by  $\mathcal{T}$  the test procedure we use to discriminate between a pair of measurements  $\mathcal{M}$  and  $\mathcal{N}$  ( $\mathcal{N}$  corresponding to POVM elements  $\{N_j\}$ ). We denote by  $p(c|\mathcal{M}, \mathcal{T})$  the conditional probability that if the measurement  $\mathcal{M}$  was tested by the test procedure  $\mathcal{T}$ , conclusion  $c \in \{\mathcal{M}, \mathcal{N}, f\}$  was obtained. Here,  $c = f$  marks that the procedure has failed and  $c = \mathcal{M}$ ,  $c = \mathcal{N}$  corresponds to identification of the measurement device as  $\mathcal{M}$ ,  $\mathcal{N}$ , respectively. The fact that the test procedure  $\mathcal{T}$  fails with fixed probability  $p_f$  can be mathematically stated as

$$p_f = \eta_{\mathcal{M}} p(f|\mathcal{M}, \mathcal{T}) + \eta_{\mathcal{N}} p(f|\mathcal{N}, \mathcal{T}). \quad (2)$$

We define

$$\begin{aligned} p_e &= \eta_{\mathcal{M}} p(\mathcal{N}|\mathcal{M}, \mathcal{T}) + \eta_{\mathcal{N}} p(\mathcal{M}|\mathcal{N}, \mathcal{T}), \\ p_s &= \eta_{\mathcal{M}} p(\mathcal{M}|\mathcal{M}, \mathcal{T}) + \eta_{\mathcal{N}} p(\mathcal{N}|\mathcal{N}, \mathcal{T}) \end{aligned} \quad (3)$$

the probability of error and the probability of success and clearly  $p_s + p_e + p_f = 1$ . Our goal is to maximize the probability of success  $p_s$  for a fixed value of the failure rate  $p_f$ . This is equivalent to minimization of the relative error rate  $p_e/(1 - p_f)$  or maximization of the relative success probability  $p_s/(1 - p_f)$  for fixed value of  $p_f$ .

The mathematical framework for the description of test procedures  $\mathcal{T}$  (so-called quantum testers, or process POVMs) was introduced and developed in Refs. [27–30]. In this framework, the measurements are described by Choi-Jamiolkowski operators [31,32] assigned to corresponding quantum-to-classical channels. For example, to  $\mathcal{M}$  defined in Eq. (1) we assign

$$M = (\mathcal{M} \otimes \mathcal{I})[|\phi_+\rangle\langle\phi_+|] = \sum_j |j\rangle\langle j| \otimes M_j^T,$$

where  $|\phi_+\rangle = \sum_k |k\rangle \otimes |k\rangle \in \mathcal{H}_d \otimes \mathcal{H}_d$  is the (unnormalized) maximally entangled state ( $\{|k\rangle\}$  is an orthonormal basis in  $\mathcal{H}_d$

and  $\{|j\rangle\}$  is an orthonormal basis in  $\mathcal{H}_n$ ). Any possible test  $\mathcal{T}$  is described by a set of positive operators  $\{T_c\}$  acting on  $\mathcal{H}_n \otimes \mathcal{H}_d$  such that

$$\sum_c T_c = I_n \otimes \rho, \quad (4)$$

where  $\rho$  is a density operator on  $\mathcal{H}_d$ , i.e.,  $\rho \geq 0$  and  $\text{tr}[\rho] = 1$ ).

Let us define projectors  $\pi_j \equiv |j\rangle\langle j| \otimes I_d$  reflecting the symmetry of Choi-Jamiolkowski operators of measurement  $\mathcal{M}$ , thus, satisfying the identity  $M^T = \sum_j \pi_j M^T \pi_j$ , where  $T$  denotes a transposition with respect to the basis  $\{|j\rangle \otimes |k\rangle\}$ . Then, for any measurement  $\mathcal{M}$  the conditional probabilities satisfy the following identity:

$$\begin{aligned} p(c|\mathcal{M}, \mathcal{T}) &= \text{tr}[T_c M^T] = \sum_j \text{tr}[T_c \pi_j M^T \pi_j] \\ &= \sum_j \text{tr}[\pi_j T_c \pi_j M^T] \equiv \text{tr}[\pi(T_c) M^T], \end{aligned} \quad (5)$$

implying that the test procedure formed by operators  $\{\pi(T_c) \equiv \sum_j \pi_j T_c \pi_j\}_c$  is indistinguishable from the test procedure composed of operators  $\{T_c\}_c$ . In other words, without loss of generality we may assume that the test procedure  $\mathcal{T}$  is composed of positive operators of the form

$$T_c = \sum_j |j\rangle\langle j| \otimes H_j^{(c)} \quad (6)$$

for which  $\pi(T_c) = T_c$  and for all  $j$  obeying the normalization [see Eq. (4)]

$$\sum_c H_j^{(c)} = \rho. \quad (7)$$

Consequently, the conditional probability equals

$$p(c|\mathcal{M}, \mathcal{T}) = \sum_j \text{tr}[H_j^{(c)} M_j]. \quad (8)$$

In the considered case of discrimination of a pair of measurements  $\mathcal{M}$  and  $\mathcal{N}$  (*a priori* occurring with probabilities  $\eta_{\mathcal{M}}$  and  $\eta_{\mathcal{N}}$ , respectively) we have  $c \in \{\mathcal{M}, \mathcal{N}, f\}$ , hence, the normalization explicitly reads as

$$H_j^{(\mathcal{M})} + H_j^{(\mathcal{N})} + H_j^{(f)} = \rho, \quad (9)$$

for all  $j = 1, \dots, n$ .

After deriving the above expressions, we have all the mathematical instruments required to formalize and analyze any particular measurement discrimination problem. In what follows, we present several cases in which the structure of the problem allows us to either partly simplify the choice of the normalization  $\rho$  or to completely determine it and to reduce the optimization of the discrimination of measurements to discrimination of states.

#### IV. PERFECT DISCRIMINATION

Let us first address the case of perfect discrimination. This is an intriguing quantum information theory question because the maximal number of simultaneously perfectly distinguishable measurements reveals the information ‘‘capacity’’ of measurement devices. It is known that for states this number coincides with the dimension of the Hilbert space, thus, provides its

operational meaning. There are  $d^2$  perfectly distinguishable unitary channels (e.g., Pauli operators in case of qubit) and this property is exploited in superdense coding [33] to double the information transmission rate of noiseless communication with  $d$ -level systems.

Surprisingly, a nontrivial insight on perfect discrimination comes from the results of Ref. [34], where the concept of boundariness was introduced. Based on the close relation between boundariness and minimum-error discrimination, we know that perfect discrimination is possible only between boundary elements (for details see Sec. IV of [34]). Let us stress that this feature holds also for states and channels. In particular, the results of Ref. [34] imply that for each measurement  $\mathcal{M}$  from the boundary there exists a measurement  $\mathcal{N}$  (also belonging to the boundary) such that  $\mathcal{M}$  and  $\mathcal{N}$  are perfectly distinguishable.

The channel representation of measurements makes the problem of discrimination of observables a special case of the channel discrimination. It follows [35] that for the minimum-error discrimination of equiprobable measurements  $\mathcal{M}$  and  $\mathcal{N}$  the optimal-error probability is given by the following formula:

$$p_e = \frac{1}{2}(1 - \frac{1}{2}\|\mathcal{M} - \mathcal{N}\|_{\text{cb}}), \quad (10)$$

where  $\|\dots\|_{\text{cb}}$  denotes the completely bounded (CB) norm [36]. In general, it is difficult to evaluate this norm because it requires inspection of the behavior of the map when tensorized with identity channel  $\mathcal{I}_k$  on  $k$ -dimensional Hilbert space  $\mathcal{H}_k$ :

$$\|\mathcal{M} - \mathcal{N}\|_{\text{cb}} = \max_{k \in \mathbb{N}, \rho \geq 0, \text{tr}[\rho]=1} \|[(\mathcal{M} - \mathcal{N}) \otimes \mathcal{I}_k](\rho)\|_{\text{tr}},$$

where  $\|X\|_{\text{tr}} = \text{tr}|X|$  denotes the trace norm. Unfortunately, the following example demonstrates that although the measurements represent a special type of channels (with classical outputs), the perfect discrimination can not be, in general, restricted to simple (ancilla-free) schemes only, i.e.,

$$\|\mathcal{M} - \mathcal{N}\|_{\text{cb}} > \max_{\rho \geq 0, \text{tr}[\rho]=1} \|\mathcal{M}(\rho) - \mathcal{N}(\rho)\|_{\text{tr}}. \quad (11)$$

*Example 1 (Perfect discrimination without simple scheme).*

Let us consider a pair of symmetric three-outcome qubit measurements

$$\begin{aligned} \mathcal{M} : M_1 &= \frac{2}{3}|0\rangle\langle 0|, M_2 = \frac{2}{3}|v_+\rangle\langle v_+|, M_3 = \frac{2}{3}|v_-\rangle\langle v_-|, \\ \mathcal{N} : N_1 &= \frac{2}{3}|1\rangle\langle 1|, N_2 = \frac{2}{3}|v_+^\perp\rangle\langle v_+^\perp|, N_3 = \frac{2}{3}|v_-^\perp\rangle\langle v_-^\perp|, \end{aligned} \quad (12)$$

where  $|v_\pm\rangle = \frac{1}{2}|0\rangle \pm \frac{\sqrt{3}}{2}|1\rangle$ ,  $|v_\pm^\perp\rangle = \frac{\sqrt{3}}{2}|0\rangle \mp \frac{1}{2}|1\rangle$ . Applying these measurements on one part of a singlet state  $|\psi_-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$  of two qubits, the other part (ancilla) is projected into two orthogonal states for any of the (equiprobable) outcomes  $j$ . For example, outcome  $j = 2$  heralds the ancilla state  $|v_+^\perp\rangle$  in case of measurement  $\mathcal{M}$  and state  $|v_+\rangle$  in case of  $\mathcal{N}$ . Thus, for  $j = 2$  the perfect discrimination can be achieved by distinguishing orthogonal states  $|v_+\rangle$ ,  $|v_+^\perp\rangle$ . Similarly, for  $j = 1$  ( $j = 3$ ) we would have states  $|0\rangle$ ,  $|1\rangle$  ( $|v_-\rangle$ ,  $|v_-^\perp\rangle$ ), respectively. We conclude that measurements  $\mathcal{M}$ ,  $\mathcal{N}$  can be perfectly discriminated using general (ancilla-assisted) scheme. It remains to show there is no ancilla-free scheme for perfect discrimination. Let us denote by  $\mu_j = \text{tr}[M_j \rho]$  and

$v_j = \text{tr}[N_j \varrho]$  the probabilities of outcomes  $j$  given the probe state is  $\varrho$ . For ancilla-free scheme the perfect discrimination happens if and only if each outcome  $j$  is associated either with conclusion  $\mathcal{M}$  or  $\mathcal{N}$  and on top of that the probability of each outcome is nonvanishing for at most one of the measurements, i.e.,  $\sum_j \mu_j v_j = 0$  (being equivalent to the conditions  $\mu_j v_j = 0$  for each  $j$ ). For the considered pair of measurements, it follows that always at least two of the outcomes have nonvanishing probabilities, thus, the necessary condition for perfect discrimination ( $\sum_j \mu_j v_j = 0$ ) can not be satisfied. In conclusion, for the discrimination of measurements, the use of ancilla provides a nontrivial advantage in comparison with simple schemes.

### A. Binary measurements

In this section, we will focus on perfect discrimination of two outcomes (binary) measurements. Suppose measurements  $\mathcal{M}, \mathcal{N}$  are described by effects  $M_1, M_2$  ( $M_1 + M_2 = I$ ) and  $N_1, N_2$  ( $N_1 + N_2 = I$ ), respectively. The following theorem provides a simple criterion for binary measurements being perfectly distinguishable. Moreover, it justifies that their perfect discrimination is achievable by simple schemes.

*Theorem 1.* A pair of two outcome measurements  $\mathcal{M}$  and  $\mathcal{N}$  can be perfectly discriminated if and only if there exists a state  $|\psi\rangle$  such that

$$\langle \psi | M_j | \psi \rangle = 1 \quad \text{and} \quad \langle \psi | N_j | \psi \rangle = 0, \quad (13)$$

for either  $j = 1$  or  $2$ .

*Proof.* We start by proving sufficiency of the condition. Suppose  $j = 1$ , i.e., the identities  $\mu_1 = \langle \psi | M_1 | \psi \rangle = 1$  and  $\nu_1 = \langle \psi | N_1 | \psi \rangle = 0$  hold. The normalization implies  $\mu_2 = 0$  and  $\nu_2 = 1$ , hence, applying the unknown measurement on the probe state  $|\psi\rangle$  and recording the outcome 1 we may conclude with certainty that the measurement is  $\mathcal{M}$ . Similarly, the observation of the outcome 2 implies the unknown measurement is  $\mathcal{N}$ , thus, the perfect discrimination is achieved. The argumentation for the case  $j = 2$  is analogous, only interpretation of the observed outcomes is switched. This proves the sufficiency of the identities (13).

Let us proceed and prove their necessity. First, we will show that perfect discrimination conditions for binary measurements  $\mathcal{M}$  and  $\mathcal{N}$ ,

$$0 = p(\mathcal{N} | \mathcal{M}, \mathcal{T}) = \text{tr}[H_1^{(\mathcal{N})} M_1] + \text{tr}[H_2^{(\mathcal{N})} M_2], \quad (14)$$

$$0 = p(\mathcal{M} | \mathcal{N}, \mathcal{T}) = \text{tr}[H_1^{(\mathcal{M})} N_1] + \text{tr}[H_2^{(\mathcal{M})} N_2],$$

imply  $\text{tr}[\varrho(M_1 + N_1)] = 1$ . Since the trace of a product of two positive operators is non-negative, all four traces in the above equation vanish. In particular, the condition  $\text{tr}[H_2^{(\mathcal{N})} M_2] = 0$  implies  $\text{tr}[H_2^{(\mathcal{N})} M_1] = \text{tr}[H_2^{(\mathcal{N})}(I - M_2)] = \text{tr}[H_2^{(\mathcal{N})}]$ . Similarly,  $\text{tr}[H_2^{(\mathcal{M})} N_2] = 0$  implies  $\text{tr}[H_2^{(\mathcal{M})} N_1] = \text{tr}[H_2^{(\mathcal{M})}]$ . Using the identity

$$H_2^{(\mathcal{M})} + H_2^{(\mathcal{N})} = \rho \quad (15)$$

implied by normalization (7) with  $\text{tr}[\rho] = 1$  we obtain the condition

$$\text{tr}[H_2^{(\mathcal{N})} M_1] + \text{tr}[H_2^{(\mathcal{M})} N_1] = 1. \quad (16)$$

Due to Eq. (15) we have  $H_2^{(\mathcal{N})} \leq \rho$  and  $H_2^{(\mathcal{M})} \leq \rho$ , hence,

$$1 = \text{tr}[H_2^{(\mathcal{N})} M_1] + \text{tr}[H_2^{(\mathcal{M})} N_1] \leq \text{tr}[\varrho(M_1 + N_1)]. \quad (17)$$

Trace of a product of two positive operators vanishes if and only if the supports of the two operators are orthogonal. This implies [Eqs. (14) and (7)]

$$H_1^{(\mathcal{N})} = \lambda \tilde{M}_1^\perp, \quad H_1^{(\mathcal{M})} = (1 - \lambda) \tilde{N}_1^\perp, \quad (18)$$

where  $\tilde{M}_1^\perp, \tilde{N}_1^\perp$  are density operators with supports orthogonal to  $M_1, N_1$ , respectively, and  $\lambda = \text{tr}[H_1^{(\mathcal{N})}] \in [0, 1]$ . Clearly, due to normalization

$$\rho = H_1^{(\mathcal{M})} + H_1^{(\mathcal{N})} = \lambda \tilde{M}_1^\perp + (1 - \lambda) \tilde{N}_1^\perp.$$

Using the identities  $\tilde{M}_1^\perp M_1 = O$  and  $\tilde{N}_1^\perp N_1 = O$  we obtain

$$\begin{aligned} \text{tr}[\rho(M_1 + N_1)] &= (1 - \lambda) \text{tr}[\tilde{N}_1^\perp M_1] + \lambda \text{tr}[\tilde{M}_1^\perp N_1] \\ &\leq \max\{\text{tr}[\tilde{N}_1^\perp M_1], \text{tr}[\tilde{M}_1^\perp N_1]\}. \end{aligned} \quad (19)$$

Since  $\tilde{M}_1^\perp, \tilde{N}_1^\perp$  are states and  $M_1, N_1$  are effects (i.e.,  $M_1, N_1 \leq I$ ) it follows that  $\text{tr}[\rho(M_1 + N_1)] \leq 1$ . Combining this inequality with Eq. (17) we may conclude that perfect discrimination implies  $\text{tr}[\varrho(M_1 + N_1)] = 1$ , thus, the upper and lower bounds are both saturated. For upper bound this requires an existence either of a state  $\rho = \tilde{M}_1^\perp$  such that  $\text{tr}[\tilde{M}_1^\perp N_1] = 1$ ,  $\text{tr}[\tilde{M}_1^\perp M_1] = 0$ , or of a state  $\rho = \tilde{N}_1^\perp$  such that  $\text{tr}[\tilde{N}_1^\perp M_1] = 1$ ,  $\text{tr}[\tilde{N}_1^\perp N_1] = 0$ . Finally, let us stress that the state  $\rho$  can be always chosen to be a pure state  $|\psi\rangle$  being an eigenvector of  $M_1$  (case  $j = 1$ ) or  $N_1$  (case  $j = 2$ ) associated with eigenvalue 1 and simultaneously belonging to the kernel of operators  $N_1, M_1$ , respectively. ■

As a consequence of this theorem, a pair of two-outcome (binary) measurements can be perfectly discriminated only if one of the POVM elements of  $\mathcal{M}$  (say  $M_1$ ) has eigenvalue one in a subspace in which  $N_1$  has eigenvalue zero, i.e.,  $M_1 |\psi\rangle = |\psi\rangle$  and  $N_1 |\psi\rangle = 0$ . In particular, for binary qubit measurements the perfect distinguishability implies the following form of observables:

$$\begin{aligned} M_1 &= |\varphi\rangle\langle\varphi| + q|\varphi^\perp\rangle\langle\varphi^\perp|, & M_2 &= (1 - q)|\varphi^\perp\rangle\langle\varphi^\perp|, \\ N_1 &= r|\varphi^\perp\rangle\langle\varphi^\perp|, & N_2 &= |\varphi\rangle\langle\varphi| + (1 - r)|\varphi^\perp\rangle\langle\varphi^\perp|, \end{aligned}$$

where  $0 \leq r, q \leq 1$  and  $|\varphi\rangle, |\varphi^\perp\rangle$  form an orthonormal basis of  $\mathcal{H}_2$ . The optimal probe state reads  $|\psi\rangle = |\varphi\rangle$ .

### B. More than two measurements

Let us get back to the questions raised at the beginning of this section. What is the maximum number  $m$  of perfectly distinguishable measurements of  $d$ -dimensional quantum system? How is this number related to the dimension? In what follows, we give an example exhibiting a rather surprising fact that  $m$  can be arbitrary, irrelevant of the system's dimension.

Consider  $m$  measurements  $\mathcal{M}_l$  ( $l = 1, \dots, m$ ), each of them with  $n \geq m$  outcomes (labeled as before by  $j = 1, \dots, n$ ) associated with effects

$$M_{lj} = \begin{cases} |\varphi\rangle\langle\varphi| & \text{if } j = l, \\ x_{lj}(I - |\varphi\rangle\langle\varphi|) & \text{if } j \neq l, \end{cases}$$

where  $0 < x_{lj} < 1$  and  $\sum_j x_{lj} = 1$ . Using a test state  $|\psi\rangle = |\varphi\rangle$ , the outcome  $j = l$  of measurement  $\mathcal{M}_l$  is observed with certainty, hence, observation of the outcome  $j$  perfectly identifies the measurement  $\mathcal{M}_{l=j}$ . This is an example of  $m$  perfectly distinguishable measurements. Let us stress that the dimension of the system is not specified and also that no ancilla is needed. Let us also note that the choice of  $x_{lj}$  (for  $j \neq l$ ) is arbitrary, thus, the measurements  $\mathcal{M}_l$  are not just mutually relabeled measurements.

The following proposition relates the maximal number of perfectly distinguishable measurements  $m$  with the number of outcomes  $n$ .

*Proposition 1.* If  $n$ -outcome quantum measurements  $\mathcal{M}_1, \dots, \mathcal{M}_m$  can be perfectly discriminated, then  $m \leq n$ .

*Proof.* Similarly, as for the discrimination of two measurements, we can write the conditional probability as  $p(c|\mathcal{M}_l, \mathcal{T}) = \sum_j \text{tr}[H_j^{(c)} M_{lj}]$ . Since for perfect discrimination the inconclusive outcome cannot occur, thus,  $c \neq f$ , we use  $c \in \{1, \dots, m\}$  indicating the measurement  $\mathcal{M}_c$ . The operators  $H_j^{(c)}$  must fulfill the normalization identity  $\sum_{k=1}^m H_j^{(k)} = \rho$  for  $\forall j \in \Omega$ . A test  $\mathcal{T}$  perfectly distinguishes measurements  $\{\mathcal{M}_l\}$  if and only if for all  $l$  the following identity holds:  $p(c|\mathcal{M}_l, \mathcal{T}) = \sum_j \text{tr}[H_j^{(c)} M_{lj}] = \delta_{cl}$ .

Let us introduce positive operators  $E_{cj} \equiv \rho^{-1/2} H_j^{(c)} \rho^{-1/2}$  and  $Q_{lj} \equiv \rho^{1/2} M_{lj} \rho^{1/2}$  satisfying the identities

$$\sum_c E_{cj} = \Pi_\rho \quad \text{and} \quad \sum_{j=1}^n Q_{lj} = \rho \quad (20)$$

for all  $j$  and  $l$ , respectively. We denote by  $\Pi_\rho$  the projector onto the support of  $\rho$ . Then,

$$\begin{aligned} p(l|\mathcal{M}_l, \mathcal{T}) &= \sum_j \text{tr}[E_{lj} Q_{lj}] \\ &\leq \sum_j \text{tr}[\Pi_\rho Q_{lj}] = \sum_j \text{tr}[Q_{lj}] = 1, \end{aligned} \quad (21)$$

where we used that  $E_{lj} \leq \Pi_\rho$ ,  $Q_{lj} \leq \rho \leq \Pi_\rho$ , Eq. (20), and  $\text{tr}[\rho] = 1$ . It follows that the condition  $p(l|\mathcal{M}_l, \mathcal{T}) = 1$  holds only if  $\text{tr}[E_{lj} Q_{lj}] = \text{tr}[Q_{lj}]$  for all  $l, j$ . Since  $0 \leq E_{lj} \leq \Pi_\rho$  and  $0 \leq Q_{lj} \leq \Pi_\rho$ , this is equivalent to the requirement  $E_{lj} \geq \Pi_{lj}$ , where  $\Pi_{lj}$  denotes a projector onto a support of the operator  $Q_{lj}$ . Consequently, the multiplicity  $\kappa_{lj}$  of eigenvalue 1 in the spectral decomposition of  $E_{lj}$  has to be at least the rank of  $\Pi_{lj}$ , i.e.,

$$\forall l, j, \quad \kappa_{lj} \geq \text{tr}[\Pi_{lj}]. \quad (22)$$

Denote by  $D = \text{tr}[\Pi_\rho]$  the dimension of the support of  $\rho$ . Due to Eq. (20) we have  $\sum_j \text{tr}[\Pi_{lj}] \geq D$  because the rank of the sum of positive operators  $Q_{lj}$  is at most the sum of the ranks of its parts. Combining this with Eq. (22) and summing over  $l$  we obtain

$$\sum_l \sum_j \kappa_{lj} \geq mD. \quad (23)$$

On the other hand, taking into the account the normalization from Eq. (20) and inequality  $\kappa_{lj} \leq \text{tr}[E_{lj}]$ , it follows that

$\sum_l \kappa_{lj} \leq D$  and, consequently,

$$\sum_j \sum_l \kappa_{lj} \leq nD. \quad (24)$$

Combining inequalities (23) and (24), we get  $m \leq n$ . ■

For nondegenerate projective measurements, a rank-one projector corresponds to each outcome. By definition, such measurement has  $n = d$  outcomes, where  $d$  is the dimension of  $\mathcal{H}_d$ . The above Proposition 1 implies there are at most  $d$  perfectly distinguishable nondegenerate projective measurements.

## V. QUANTUM FILTERS

A projective two-outcome measurement  $\mathcal{M}$  is called a *quantum filter* if one of its outcomes is described by rank-one projection. Discrimination of a pair of quantum filters  $\mathcal{M}, \mathcal{N}$  can appear in two different variations depending on assignment of rank-one operators to particular labels: either the same outcome is described by rank-one operators, or exclusive outcomes are associated with rank-one operators for  $\mathcal{M}$  and  $\mathcal{N}$ .

Let us start with the first case and set the outcome labeled as “1” to be the one described by rank-one projector, i.e.,

$$\begin{aligned} \mathcal{M} : M_1 &= |\varphi\rangle\langle\varphi|, \quad M_2 = I - M_1; \\ \mathcal{N} : N_1 &= |\psi\rangle\langle\psi|, \quad N_2 = I - N_1. \end{aligned} \quad (25)$$

The reduction theorem formulated in the following section reduces this problem to discrimination of qubit projective measurements by identifying a relevant two-dimensional subspace of  $\mathcal{H}_d$ . In particular, the statement of the theorem is more general and allows us to identify irrelevant subspace for discrimination of arbitrary measurements.

### A. Reduction theorem

Consider a pair of  $n$ -outcome measurements  $\mathcal{M}$  and  $\mathcal{N}$  on  $d$ -dimensional Hilbert space represented by POVMs  $\{M_j\}$  and  $\{N_j\}$ , respectively. Suppose that  $\forall j$   $Q_j$  is a (largest) projector such that  $Q_j \leq M_j$  and  $Q_j \leq N_j$ . Due to POVM normalization, the projectors  $Q_j$  are mutually orthogonal, i.e.,  $Q_j Q_k = \delta_{jk} Q_j$ . We may define a projector  $P = \sum_j Q_j$  and “measurements”  $\tilde{\mathcal{M}}$  and  $\tilde{\mathcal{N}}$  with POVM elements  $\{\tilde{M}_j \equiv M_j - Q_j\}$  and  $\{\tilde{N}_j \equiv N_j - Q_j\}$ , respectively, and normalized to  $I - P$ . Let us stress that  $\tilde{M}_j = (I - P)M_j(I - P)$ . The following theorem shows that the subspace determined by the support of  $P$  plays no role and the original discrimination problem is equivalent to discrimination of measurements  $\tilde{\mathcal{M}}, \tilde{\mathcal{N}}$  defined on the subspace  $\tilde{\mathcal{H}} \equiv (I - P)\mathcal{H}_d \subset \mathcal{H}_d$  relevant for the discrimination.

*Theorem 2.* Suppose that  $\mathcal{T}$  and  $\tilde{\mathcal{T}}$  are optimal solutions to discrimination with fixed failure rate  $p_f = \tilde{p}_f$  between pairs of measurements  $\mathcal{M}, \mathcal{N}$  and  $\tilde{\mathcal{M}}, \tilde{\mathcal{N}}$ , respectively. Then,  $p_s = \tilde{p}_s$  or, equivalently,  $p_e = \tilde{p}_e$ . Moreover, optimal  $\mathcal{T}$  can be chosen such that  $\mathcal{T}|_{\tilde{\mathcal{H}}} = \tilde{\mathcal{T}}$  and vice versa (i.e., given optimal  $\mathcal{T}$  optimal test  $\tilde{\mathcal{T}}$  can be chosen as  $\mathcal{T}|_{\tilde{\mathcal{H}}} = \tilde{\mathcal{T}}$ ). ■

*Proof.* See Appendix. ■

Let us formulate consequences of the above theorem for quantum filters. We define a two-dimensional Hilbert space

$\tilde{\mathcal{H}}$  as a linear span of vectors  $|\varphi\rangle, |\psi\rangle$ . Clearly,  $M_2 \geq P$  and  $N_2 \geq P$ , where  $P$  is a projector onto the subspace  $\tilde{\mathcal{H}}^\perp$ . Using the reduction Theorem 2, the discrimination of filters can be solved by finding the solution to discrimination of projective qubit measurements

$$\begin{aligned} \tilde{\mathcal{M}} : \quad \tilde{M}_1 &= |\varphi\rangle\langle\varphi|, \quad \tilde{M}_2 = |\varphi^\perp\rangle\langle\varphi^\perp|; \\ \tilde{\mathcal{N}} : \quad \tilde{N}_1 &= |\psi\rangle\langle\psi|, \quad \tilde{N}_2 = |\psi^\perp\rangle\langle\psi^\perp|, \end{aligned} \quad (26)$$

where  $|\psi^\perp\rangle, |\varphi^\perp\rangle$  are vectors from  $\tilde{\mathcal{H}}$  orthogonal to  $|\psi\rangle, |\varphi\rangle$ , respectively. The solution to this problem is given in Sec. VI. Let us stress that similar reasoning applies also to the case of discrimination among  $m$  quantum filters (with  $M_{1i}$  being rank-one projectors). In such case, the problem is equivalent to discrimination of  $m$  quantum filters on  $m$ -dimensional subspace of  $\mathcal{H}_d$ .

Finally, we discuss the other possible assignment of outcomes for two quantum filters, i.e., the case

$$\begin{aligned} \mathcal{M} : M_1 &= |\varphi\rangle\langle\varphi|, \quad M_2 = I - M_1; \\ \mathcal{N} : N_1 &= I - N_2, \quad N_2 = |\psi\rangle\langle\psi|, \end{aligned} \quad (27)$$

when rank-one projections correspond to different outcomes. If the dimension of  $\mathcal{H}_d$  is two, the problem coincides with discrimination of two projective qubit measurements, which we solve in the next section. Otherwise, there exists a state  $|\phi\rangle$  orthogonal to  $|\psi\rangle$  and  $|\varphi\rangle$ . Measuring  $|\phi\rangle$  with  $\mathcal{M}$  we always get outcome 2, while  $\mathcal{N}$  will always produce outcome 1. Thus, for  $\dim \mathcal{H}_d \geq 3$  any pair of (different) quantum filters (27) is always perfectly distinguishable.

## VI. PROJECTIVE QUBIT MEASUREMENTS

In this section, we shall analyze discrimination of projective qubit measurements, i.e., measurements such as  $\mathcal{M}$  described by effects  $M_1 = |\varphi\rangle\langle\varphi|$  and  $M_2 = I - M_1 = |\varphi^\perp\rangle\langle\varphi^\perp|$  for some orthonormal basis  $\{|\varphi\rangle, |\varphi^\perp\rangle\}$  of  $\mathcal{H}_2$ . As we declared in Sec. III, our goal is to maximize probability of success  $p_s$  for a fixed failure probability  $p_f$ .

### A. Binary discrimination problem

Let us start with the simplest case, when our goal is to discriminate among a pair of projective measurements

$$\begin{aligned} \mathcal{M} : \quad M_1 &= |\varphi\rangle\langle\varphi|, \quad M_2 = |\varphi^\perp\rangle\langle\varphi^\perp|; \\ \mathcal{N} : \quad N_1 &= |\psi\rangle\langle\psi|, \quad N_2 = |\psi^\perp\rangle\langle\psi^\perp|. \end{aligned} \quad (28)$$

Suppose  $\mathcal{T}$  is a test procedure specified by operators  $H_1^{(c)}, H_2^{(c)}$  with  $c \in \{\mathcal{M}, \mathcal{N}, f\}$  such that for all  $j$   $\sum_c H_j^{(c)} = \varrho$  and  $\mathcal{T}$  leads to certain values of  $p_s, p_e$ , and  $p_f$ . Further, we will exploit the reflection symmetry of the problem. In particular, let us denote by  $\Gamma$  the *universal NOT* transformation  $X \mapsto X^\perp = \text{tr}[X]I - X$  for any operator  $X$ . In  $\mathcal{H}_2$ , this map is positive (not completely positive) and trace preserving. Moreover,  $\text{tr}[\Gamma(X)Y] = \text{tr}[X\Gamma(Y)]$  for all operators  $X, Y$  and  $\Gamma^2 = \mathcal{I}$ .

By properties (positivity) of  $\Gamma$  it follows that operators

$$H_1^{(c)} = \Gamma(H_2^{(c)}) \quad \text{and} \quad H_2^{(c)} = \Gamma(H_1^{(c)}) \quad (29)$$

form a valid test procedure  $\mathcal{T}'$  with normalization

$$\sum_c H_1^{(c)} = \sum_c H_2^{(c)} = \Gamma(\varrho).$$

For conditional probabilities, we find

$$\begin{aligned} p(c|\mathcal{M}, \mathcal{T}') &= \text{tr}[\Gamma(H_2^{(c)}) |\varphi\rangle\langle\varphi| + \Gamma(H_1^{(c)}) |\varphi^\perp\rangle\langle\varphi^\perp|] \\ &= \text{tr}[H_2^{(c)} \Gamma(|\varphi\rangle\langle\varphi|) + H_1^{(c)} \Gamma(|\varphi^\perp\rangle\langle\varphi^\perp|)] \\ &= \text{tr}[H_2^{(c)} |\varphi^\perp\rangle\langle\varphi^\perp| + H_1^{(c)} |\varphi\rangle\langle\varphi|] \\ &= p(c|\mathcal{M}, \mathcal{T}) \end{aligned} \quad (30)$$

and analogously  $p(c|\mathcal{N}, \mathcal{T}') = p(c|\mathcal{N}, \mathcal{T})$ . In other words, both test procedures  $\mathcal{T}$  and  $\mathcal{T}'$  determine the same probabilities  $p_s, p_e$ , and  $p_f$ , thus they both perform equally well in the considered discrimination problem. Moreover, any convex combination, in particular  $\tilde{\mathcal{T}} = \frac{1}{2}\mathcal{T} + \frac{1}{2}\mathcal{T}'$  of these tests, results in the same probabilities  $p_s, p_e$ , and  $p_f$ . This allows us to reduce the set of considered test procedures and to fix their normalization  $\varrho$  without loss of generality because the normalization of  $\tilde{\mathcal{T}}$  is independent of the test  $\mathcal{T}$  and reads as

$$\begin{aligned} \forall j \quad \sum_c \tilde{H}_j^{(c)} &= \sum_c \frac{1}{2}(H_j^{(c)} + H_j^{(c)}) \\ &= \frac{1}{2}[\varrho + \Gamma(\varrho)] = \frac{1}{2}I. \end{aligned} \quad (31)$$

Moreover,

$$\begin{aligned} \Gamma(\tilde{H}_1^{(c)}) &= \Gamma[\frac{1}{2}(H_1^{(c)} + H_1^{(c)})] \\ &= \frac{1}{2}(H_2^{(c)} + H_2^{(c)}) = \tilde{H}_2^{(c)}, \end{aligned} \quad (32)$$

so the considered test procedures  $\tilde{\mathcal{T}}$  are completely specified by operators for a single outcome, i.e., by positive operators  $\tilde{H}_1^{(c)}$  and by their normalization condition (31). Using this fact, we obtain formulas

$$\begin{aligned} p(c|\mathcal{M}, \tilde{\mathcal{T}}) &= \text{tr}[\tilde{H}_1^{(c)} |\varphi\rangle\langle\varphi| + \Gamma(\tilde{H}_1^{(c)}) |\varphi^\perp\rangle\langle\varphi^\perp|] \\ &= \text{tr}[2\tilde{H}_1^{(c)} |\varphi\rangle\langle\varphi|] \equiv \text{tr}[E_c |\varphi\rangle\langle\varphi|], \\ p(c|\mathcal{N}, \tilde{\mathcal{T}}) &= \text{tr}[2\tilde{H}_1^{(c)} |\psi\rangle\langle\psi|] \equiv \text{tr}[E_c |\psi\rangle\langle\psi|], \end{aligned} \quad (33)$$

where we defined positive operators  $E_c \equiv 2\tilde{H}_1^{(c)}$  for each  $c \in \{\mathcal{M}, \mathcal{N}, f\}$ . Let us stress that Eq. (31) implies

$$E_{\mathcal{M}} + E_{\mathcal{N}} + E_f = I.$$

In other words, the positive operators  $E_{\mathcal{M}}, E_{\mathcal{N}}, E_f$  form a POVM coinciding with a measurement discriminating pure states  $|\psi\rangle, |\varphi\rangle$ . Indeed, using Eqs. (2), (3), and (33), we can express  $p_s, p_e$ , and  $p_f$  as

$$\begin{aligned} p_s &= \eta_{\mathcal{M}} \langle\varphi|E_{\mathcal{M}}|\varphi\rangle + \eta_{\mathcal{N}} \langle\psi|E_{\mathcal{N}}|\psi\rangle, \\ p_e &= \eta_{\mathcal{M}} \langle\varphi|E_{\mathcal{N}}|\varphi\rangle + \eta_{\mathcal{N}} \langle\psi|E_{\mathcal{M}}|\psi\rangle, \\ p_f &= \eta_{\mathcal{M}} \langle\varphi|E_f|\varphi\rangle + \eta_{\mathcal{N}} \langle\psi|E_f|\psi\rangle. \end{aligned} \quad (34)$$

Thus, we managed to reduce the discrimination of projective qubit measurements (in any version) to discrimination of pure states. In particular, we may formulate the following theorem.

*Theorem 3.* The problem of optimal discrimination with fixed failure rate  $p_f$  of projective qubit measurements  $\mathcal{M}$  and  $\mathcal{N}$  (determined by vector states  $|\varphi\rangle, |\psi\rangle$ , respectively) is

mathematically equivalent to an optimal discrimination with fixed failure rate  $p_f$  of pure states  $|\varphi\rangle, |\psi\rangle$ .

Suppose a POVM associated with effects  $E_c$  with  $c \in \{\varphi, \psi, f\}$  is the optimal solution (for details see [7–9]) for the discrimination with a fixed failure rate  $p_f$  of pure states  $|\varphi\rangle, |\psi\rangle$ . Then, the optimal discrimination [see Eq. (28)] of projective qubit measurements  $\mathcal{M}$  and  $\mathcal{N}$  (determined by vector states  $|\varphi\rangle, |\psi\rangle$ , respectively) can be implemented as follows. We prepare a maximally entangled state  $|\phi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$  of two qubits. We measure one of the qubits by the unknown measurement that we want to identify. If the outcome 1 is observed, then we perform the measurement of  $\{E_c^T\}$  on the second qubit. If we observe outcome 2, then the second qubit is measured by POVM  $\{\Gamma(E_c)^T = \Gamma(E_c^T)\}$ . It is straightforward to verify that this procedure results in conditional probabilities given in Eq. (33) if we associate conclusions  $c$  as  $\varphi \leftrightarrow \mathcal{M}$ ,  $\psi \leftrightarrow \mathcal{N}$ . Let us remind that the reduction Theorem 2 described in the previous section implies that the same procedure can be used to discriminate (optimally) quantum filters.

In the following, we illustrate what the results on optimal discrimination of two states imply for the discrimination of two qubit measurements.

*Example 2 (Minimum-error discrimination).* By definition we set  $p_f = 0$ . The formula for discrimination of two pure states is well known due to seminal works of Helstrom and Holevo [3,4]. Let us denote by  $\eta \equiv \eta_{\mathcal{M}}$  the *a priori* probability for  $|\varphi\rangle$  ( $\mathcal{M}$ ) and  $1 - \eta \equiv \eta_{\mathcal{N}}$  being the *a priori* probability for  $|\psi\rangle$  ( $\mathcal{N}$ ). The optimal POVM consists of elements  $E_\varphi = |\alpha\rangle\langle\alpha|$ ,  $E_\psi = |\beta\rangle\langle\beta|$  being projectors onto positive and negative eigensubspaces of operator  $\Delta = (1 - \eta)|\psi\rangle\langle\psi| - \eta|\varphi\rangle\langle\varphi|$ , respectively. The optimal (minimal) probability of error is given by the famous Helstrom's formula

$$p_e = \frac{1}{2}[1 - \sqrt{1 - 4\eta(1 - \eta)|\langle\psi|\varphi\rangle|^2}]. \quad (35)$$

The test procedure described above helps us to design the optimal discrimination of a pair of associated measurements  $\mathcal{M}$  and  $\mathcal{N}$ . However, let us note that the ancilla is not really necessary to achieve the optimality (which is in accordance with the discussion at the beginning of Sec. IV). Indeed, it is sufficient to prepare a test state  $|\alpha\rangle$ . Observing outcome 1 we conclude that the tested measurement was  $\mathcal{N}$  and otherwise we conclude it was  $\mathcal{M}$ . Alternatively, one can also exploit the test state  $|\beta\rangle$  and inverting the interpretation of the outcomes, we achieve again the optimal value of the error probability  $p_e$ .

*Example 3 (Unambiguous discrimination).* By definition, we call the discrimination unambiguous if  $p_e = 0$ . The solution to unambiguous discrimination of two pure states was found first for equal prior probabilities by Ivanovic [37], Dieks [38], and Peres [39] and later by Jaeger and Shimony [40] for the general situation. The solution has three regimes depending on the relation between the prior probability  $\eta$  and overlap  $F = |\langle\psi|\varphi\rangle|$ :

$$p_f = \begin{cases} \eta + (1 - \eta)F^2, & (1 + F^2)\eta \leq F^2; \\ 2\sqrt{\eta(1 - \eta)}F, & F^2 \leq (1 + F^2)\eta \leq 1; \\ 1 - \eta + \eta F^2, & (1 + F^2)\eta \geq 1. \end{cases} \quad (36)$$

If the priors are very unbalanced (first and last intervals), then one of the states is never detected, so the optimal measurement

has two outcomes and is projective. In the intermediate regime, when the priors are “comparable” all three outcomes have nonzero probability of appearance.

This means that also for discrimination of projective qubit measurements we will have three regimes defined by the same conditions. For the regime of comparable prior probabilities, it is clear that we need to use the ancillary measurement since we need three outcomes. An intuitive scheme for achieving the optimal performance is based on preparing a singlet state  $(|01\rangle - |10\rangle)/\sqrt{2}$  of two qubits. Application of the unknown measurement on one part of the state projects (depending on the identity of the measurement) the other part into a state  $|\varphi^\perp\rangle$  or  $|\psi^\perp\rangle$  in case of outcome 1 and into state  $|\varphi\rangle$  or  $|\psi\rangle$  in case of outcome 2. These two pairs of states have the same overlap ( $|\langle\varphi|\psi\rangle| = |\langle\varphi^\perp|\psi^\perp\rangle| = F$ ) and we can discriminate within the pairs using the optimal unambiguous pure state discrimination by Jaeger and Shimony. Thanks to equal overlap in case of outcome 1, outcome 2 and also on average we fail with the probability  $p_f$  given in Eq. (36).

For the remaining (unbalanced) regimes, the optimal performance can be achieved also by directly measuring the single partite state with the unknown measurement. If  $(1 + F^2)\eta_{\mathcal{M}} \geq 1$ , then we prepare  $|\psi^\perp\rangle$  and the outcome 1 unambiguously indicates that the unknown measurement is  $\mathcal{M}$ , whereas the outcome 2 is inconclusive and means that the test failed. Similarly, if  $(1 + F^2)\eta_{\mathcal{M}} \leq F^2$ , then we use  $|\varphi^\perp\rangle$  as the test state and outcome 1 unambiguously identifies the measurement  $\mathcal{N}$ .

*Example 4 (Noisy qubit measurements).* Suppose  $\mathcal{M}, \mathcal{N}$  are defined as convex combinations of a projective measurement and a trivial observable generating the uniform distribution of outcomes independently of the measured state, i.e.,

$$\begin{aligned} M_1 &= \mu|\varphi\rangle\langle\varphi| + \frac{1 - \mu}{2}I, & M_2 &= \mu|\varphi^\perp\rangle\langle\varphi^\perp| + \frac{1 - \mu}{2}I, \\ N_1 &= \nu|\psi\rangle\langle\psi| + \frac{1 - \nu}{2}I, & N_2 &= \nu|\psi^\perp\rangle\langle\psi^\perp| + \frac{1 - \nu}{2}I. \end{aligned} \quad (37)$$

As the key symmetry  $\Gamma(M_1) = M_2, \Gamma(N_1) = N_2$  holds, we can directly generalize the arguments used before and conclude that the optimal test procedure is characterized by POVM elements  $E_{\mathcal{M}}, E_{\mathcal{N}}, E_f$ , which thanks to this symmetry define  $\tilde{H}_1^{(c)} = \frac{1}{2}E_c, \tilde{H}_2^{(c)} = \frac{1}{2}\Gamma(E_c)$ . We find

$$\begin{aligned} p_s &= \eta_{\mathcal{M}} \text{tr}[E_{\mathcal{M}}M_1] + \eta_{\mathcal{N}} \text{tr}[E_{\mathcal{N}}N_1], \\ p_e &= \eta_{\mathcal{M}} \text{tr}[E_{\mathcal{N}}M_1] + \eta_{\mathcal{N}} \text{tr}[E_{\mathcal{M}}N_1], \\ p_f &= \eta_{\mathcal{M}} \text{tr}[E_fM_1] + \eta_{\mathcal{N}} \text{tr}[E_fN_1]. \end{aligned}$$

Let us stress that operators  $M_1, N_1$  are positive and have trace one, so they correspond to mixed quantum states. Thus, for measurements  $\mathcal{M}, \mathcal{N}$  defined by POVM elements from Eq. (37), we reexpressed the problem as discrimination with fixed failure rate  $p_f$  among two mixed states  $M_1, N_1$ . Such problems were studied in [10] and an upper bound on the success probability was derived. Notice that for  $\mu\nu \neq 0$ , the unambiguous discrimination of measurements is not possible because the states  $M_1, N_1$  have completely overlapping supports [41]. For the minimum-error discrimination, the

optimal-error rate and optimal POVM  $\{E_c\}$  can be again acquired easily from the work of Helstrom [4].

### B. General case

In general, the discrimination of more than two objects is more complicated than the discrimination among two of them. However, in our particular case, it turns out that the derivations in Sec. VI can be trivially generalized to the discrimination of  $m$  projective qubit measurements. In particular, the optimal discrimination of measurements  $\mathcal{M}_1, \dots, \mathcal{M}_m$ , each of them associated with effects  $M_{l1} = |\varphi_l\rangle\langle\varphi_l|$  and  $M_{l2} = |\varphi_l^\perp\rangle\langle\varphi_l^\perp|$ , can be designed by using the optimal measurement discriminating the pure states  $|\varphi_1\rangle, \dots, |\varphi_m\rangle$ . The optimal performance can be achieved by preparing a maximally entangled state, measuring one part of it by the unknown measurement and optimally discriminating the states of the remaining system. Thus, the optimal relation between success  $p_s$  and failure probability  $p_f$  for the discrimination of projective qubit measurement apparatuses is given by the solution of (pure) state discrimination problem. Sugimoto *et al.* [42] solved the discrimination with fixed failure rate of three symmetric states of a qubit, but in general the solution is not known. The special case of minimum-error discrimination of  $m$  pure qubit states can be solved completely using the result of Ref. [43].

As we discussed earlier in this section, the optimal discrimination of two projective qubit measurements with minimum possible error can always be realized by a simple discrimination scheme. A natural question arises as to whether a simple scheme can be utilized to perform the optimal minimum-error discrimination of  $m$  quantum measurements. The following example demonstrates that this is not the case, so also for projective qubit measurements there are situations when an ancilla-assisted scheme is necessary for optimization of minimum-error discrimination.

*Example 5 (Minimum-error discrimination of three projective qubit measurements).* Consider projective qubit measurements  $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3$  determined by states  $|0\rangle, |v_\pm\rangle = \frac{1}{2}(|0\rangle \pm \sqrt{3}|1\rangle)$ , respectively, appearing with equal prior probabilities. Due to the results of Clarke *et al.* [44], we have the minimal-error probability for discrimination of these states having pairwise the same fidelity. In this particular case  $p_e^{\text{opt}} = \frac{1}{3}$ , hence,  $p_s^{\text{opt}} = \frac{2}{3}$  and the same holds for optimal minimum-error discrimination of measurements  $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3$ . Let us denote by  $\rho$  the (ancilla-free) test state and we define  $x_l \equiv \text{tr}[M_{l1}\rho]$ . We further denote by  $q(l|j)$  the conditional probability of conclusion  $l$  given the outcome  $j$  was recorded on the unknown measurement we would like to identify. By definition,  $\sum_l q(l|j) = 1$  for both outcomes  $j = 1, 2$ . Then,

$$\begin{aligned} p_s &= \frac{1}{3} \sum_{l=1}^3 [q(l|1)x_l + q(l|2)(1-x_l)] \\ &= \frac{1}{3} \left\{ 1 + \sum_{l=1}^3 [q(l|1) - q(l|2)]x_l \right\} \\ &\leq \frac{1}{3} [1 + \max_l x_l - \min_l x_l] < \frac{2}{3} = p_s^{\text{opt}} \end{aligned} \quad (38)$$

because  $0 \leq \min_l x_l \leq \sum_l q(l|k)x_l \leq \max_l x_l \leq 1$  and for the considered operators  $M_{lk}$  we have strict inequality  $(\max_l x_l - \min_l x_l) < 1$ . In fact, the maximum probability of success  $p_s = (2 + \sqrt{3})/6$  for simple scheme strategies is achieved for pure state  $\rho = |\xi\rangle\langle\xi|$ , where  $|\xi\rangle = \cos\omega|0\rangle + \sin\omega|1\rangle$  and  $\omega \approx 0.0833\pi$ .

### VII. UNAMBIGUOUS DISCRIMINATION OF TWO TRINE MEASUREMENTS

Based on our previous analysis, it is natural to ask whether the ancilla-based test procedures with maximally entangled states are always the ones (although not the unique ones) optimizing the discrimination figures of merits. In this section, we will demonstrate an example rejecting such hypothesis.

Consider a symmetric three-outcome qubit measurement

$$\mathcal{M} : M_1 = \frac{2}{3}|0\rangle\langle 0|, M_2 = \frac{2}{3}|v_+\rangle\langle v_+|, M_3 = \frac{2}{3}|v_-\rangle\langle v_-|, \quad (39)$$

where  $|v_\pm\rangle = \frac{1}{2}|0\rangle \pm \frac{\sqrt{3}}{2}|1\rangle$ . Rotating this measurement by an angle  $\theta$  around the  $z$  axis (see Fig. 2) we obtain a measurement  $\mathcal{N}_\theta$  with POVM elements  $N_j = R_\theta M_j R_\theta^\dagger$ , where  $R_\theta = |0\rangle\langle 0| + e^{i\theta}|1\rangle\langle 1|$ . In what follows, we will show that maximally entangled states as test states do not optimize success probability for unambiguous discrimination of measurements  $\mathcal{M}$  and  $\mathcal{N}_\theta$ .

In the following, we will use the lower bound on the failure probability of unambiguous discrimination of two channels  $\mathcal{M}$  and  $\mathcal{N}$  (with Choi operators  $M$  and  $N$ , respectively):

$$p_f \geq 2\sqrt{\eta_{\mathcal{M}}\eta_{\mathcal{N}}}\text{tr}[\sqrt{M}(I \otimes \rho)\sqrt{N}], \quad (40)$$

where  $\rho \geq 0, \text{tr}[\rho] = 1$  is a normalization of the test used for the discrimination [see Eq. (4)]. The above bound was derived by Ziman *et al.* in [19] [see Eq. (16) therein]. Our aim is to evaluate the bound for any normalization  $\rho$  and to show that the bound can be saturated. This will allow us to compare attainable failure probability for unambiguous discrimination of measurements  $\mathcal{M}$  and  $\mathcal{N}_\theta$  for ancilla-based tests with maximally entangled states and those with optimal bipartite input states.

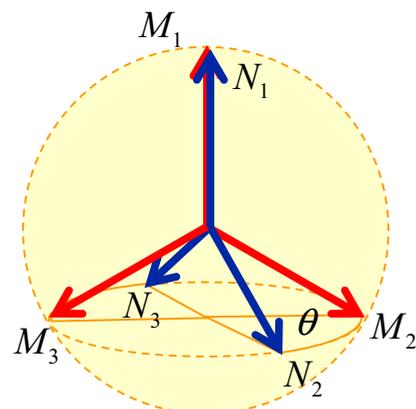


FIG. 2. (Color online) Two symmetric three-outcome qubit POVMs in the Bloch representation that are mutually rotated by angle  $\theta$ , with respect to axis  $z$ .

For uniform prior probabilities  $\eta_{\mathcal{M}} = \eta_{\mathcal{N}_\theta} = \frac{1}{2}$ , the bound (40) reads as

$$p_f \geq \frac{3}{2} \text{tr} |M(I \otimes \rho) N_\theta|, \quad (41)$$

where we used the fact that  $\sqrt{M} = \sqrt{3/2} M$ ,  $\sqrt{N_\theta} = \sqrt{3/2} N_\theta$  are the Choi operators of the measurements  $\mathcal{M}$  and  $\mathcal{N}_\theta$ , respectively.

Combining the triangle inequality for trace norm and the invariance of the norm with respect to  $\sigma_z$  rotation for the term with  $j = 3$  we can write the inequality

$$\begin{aligned} \gamma &= \text{tr} |M(I \otimes \rho) N_\theta| = \sum_j \text{tr} |M_j^T \rho N_j^T| \\ &\geq \text{tr} |M_1^T \rho N_1^T| + \text{tr} |M_2^T \rho N_2^T + \sigma_z M_3^T \rho N_3^T \sigma_z|. \end{aligned} \quad (42)$$

Using the parametrization  $\rho = q|0\rangle\langle 0| + (1-q)|1\rangle\langle 1| + z|0\rangle\langle 1| + z^*|1\rangle\langle 0|$ , the above inequality reads as

$$\gamma \geq \frac{4}{9}q + \frac{2}{9}\sqrt{q^2 + 9(1-q)^2 + 6p(1-q)\cos\theta}. \quad (43)$$

Interestingly, this expression does not depend on  $z$ , hence, the only relevant parameter of  $\rho$  is  $q$ . Combining Eqs. (41) and (43) we get

$$p_f \geq \frac{2q + \sqrt{q^2 + 9(1-q)^2 + 6q(1-q)\cos\theta}}{3}, \quad (44)$$

where  $0 \leq q \leq 1$ .

Next, we consider a test procedure with normalization  $\rho = q|0\rangle\langle 0| + (1-q)|1\rangle\langle 1|$ , which saturates the above lower bound for every  $0 \leq q \leq 1$ . Consider a test state

$$|\phi_q\rangle = \sqrt{q}|00\rangle + \sqrt{1-q}|11\rangle. \quad (45)$$

Performing a trine measurement (either  $\mathcal{M}$  or  $\mathcal{N}_\theta$ ) on one of the qubits, the second one ends up either in a conditional state  $|\psi_j^{\mathcal{M}}\rangle$  or  $|\psi_j^{\mathcal{N}}\rangle$ . For  $j = 1$ , these conditional states coincide with  $|0\rangle$ , thus, this outcome is necessarily inconclusive. The pairs of states to be discriminated for outcomes 2 and 3 are mutually related by unitary transformation  $\sigma_z$ , so they have the same overlap

$$F = |\langle \psi_2^{\mathcal{M}} | \psi_2^{\mathcal{N}} \rangle| = |\langle \psi_3^{\mathcal{M}} | \psi_3^{\mathcal{N}} \rangle| = \frac{|q + e^{i\theta}(1-q)|}{3-2q}. \quad (46)$$

Using the results of Ivanovic [37], Dieks [38], and Peres [39], such pairs of pure equiprobable states can be unambiguously discriminated with failure probability equal to their overlap  $F$ . Weighting these cases by  $p_j = \langle \phi_q | M_j \otimes I | \phi_q \rangle = \langle \phi_q | N_j \otimes I | \phi_q \rangle$ , the probability of appearance of outcome  $j$ , we derive the average failure probability of the scheme

$$p_f = \frac{2}{3}q + 2\frac{3-2q}{6}F. \quad (47)$$

Inserting Eq. (46) into (47), we see that the proposed scheme saturates the lower bound on the failure probability (44) for any  $q \in [0, 1]$ . Thus, tuning  $q$  in order to minimize the failure probability of the proposed scheme simultaneously gives the lowest achievable failure probability in general. It can be shown that the minimum of the right-hand side of Eq. (44) is achieved for  $q = [9 - 2\sqrt{3} \cos(\theta/2) -$

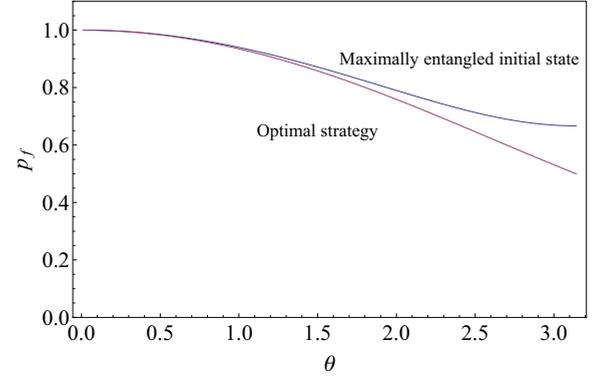


FIG. 3. (Color online) Illustration of the difference between maximally entangled and optimal input bipartite state for the discrimination of two symmetric three-outcome qubit POVMs mutually rotated by angle  $\theta$ , with respect to axis Z.

$3 \cos \theta]/(10 - 6 \cos \theta)$  implying that

$$p_f = \frac{1}{3} \left( 1 + \sqrt{3} \left| \cos \frac{\theta}{2} \right| + \frac{4 - 2\sqrt{3} |\cos \frac{\theta}{2}|}{5 - 3 \cos \theta} \right). \quad (48)$$

Finally, let us assume that the test state is any maximally entangled state. Any such test has normalization  $\rho = \frac{1}{2}I$  corresponding to  $q = \frac{1}{2}$ . Thus, by comparing the failure probability given by Eq. (44) for  $q = \frac{1}{2}$  and for  $q$  minimizing the failure probability, we can demonstrate that the use of less than maximally entangled states is needed in order to achieve the optimal performance. The difference is illustrated in Fig. 3.

## VIII. SUMMARY

In this paper, we studied discrimination of quantum measurements with finitely many outcomes in the scenario when the unknown measurement can be used only once, but use of any other resources is allowed. In particular, we investigated special instances of the discrimination with fixed failure rate. This class of problems includes perfect discrimination, minimum-error discrimination, and unambiguous discrimination.

We studied first the conditions for perfect discrimination. We have shown that the maximal number of distinguishable measurement apparatuses is bounded by the total number of outcomes  $n$ . Let us stress that the dimension of the system is irrelevant and one can find arbitrarily many qubit observables that are single-shot perfectly distinguishable. Further, we have formulated a reduction theorem excluding a subspace irrelevant for the discrimination. More precisely, we showed that any subspace common to a given outcome of both measurements is irrelevant for the discrimination. We employed this theorem to relate the discrimination of quantum filters to discrimination of projective qubit measurements.

We found that the optimization of the discrimination of projective qubit measurements is mathematically equivalent to solving discrimination of pure states. Not only the optimal success rates are the same, but also the optimal discrimination algorithm for pure state discrimination can be directly exploited for optimal discrimination of projective qubit measurements. First, we prepare a singlet state of two qubits and apply the

unknown measurement on one part of the state. This projects the second qubit into a pure state determined by the obtained outcome and the identity of the unknown measurement. Conditionally on the observed outcome, we employ the optimal state discrimination strategy to identify the projected state of the second qubit, hence, identifying the measurement used. Using this “measurement-to-state” reduction, we provide a solution to optimal minimum-error and unambiguous discrimination of projective qubit measurements. Let us note that this procedure was successfully experimentally implemented in a quantum optical system [45]. We extend this result to the case of  $m$  projective qubit measurements and, in addition, each of them may be affected by a different level of white noise. Unfortunately, we have not succeeded to formulate a similar result in more dimensional Hilbert spaces, where already the optimal discrimination of projective measurements is left open.

Our results clearly exhibit the added value of maximally entangled states although we have argued that in the case of perfect discrimination of binary measurements, the ancilla can be completely ignored and a simple scheme works as well as the entangled one. From the algebraic point of view, the measurements are channels mapping quantum (noncommutative) algebra to a classical (commutative) one, hence, the concepts of positivity and complete positivity coincide, i.e., tensor product extensions of such channels are irrelevant for judging this property. However, we were surprised to find an example exhibiting the fact that even in case of (perfect) discrimination between only a pair of measurements, the ancilla, hence, tensor product extension of the channel, provides an advantage over the simple (ancilla-free) schemes. It is an intriguing question to understand in which cases the simple scheme performs as good as the general one, and when the maximally entangled states provide the optimal discrimination strategy. We have shown an explicit example demonstrating situations in which non-maximally entangled states outperform maximally entangled ones.

In conclusion, let us mention one interesting application. The unambiguous discrimination of quantum measurements can be rephrased as conceptually novel entanglement-based quantum key distribution protocol. Indeed, suppose Alice prepares two qubits in the singlet state and sends one part of it to Bob. Bob randomly chooses to apply either a projective measurement  $\sigma_z$  or  $\sigma_x$ , each with outcomes labeled by  $\pm$ . After the measurement is accomplished, he announces publicly the outcome  $b$  to Alice, but keeps the choice of the measurement secret. Based on this outcome, Alice can perform appropriate unambiguous state discrimination either between the states  $|0\rangle, |+\rangle$  or  $|1\rangle, |-\rangle$ . If an inconclusive result is recorded, Alice and Bob discard this event from their strings. However, if Alice obtains the conclusive outcome, then she knows Bob’s choice and perfect correlation (bit of raw key) is established. Up to this point, we have just described the procedure of optimal unambiguous discrimination of measurements  $\sigma_x$  and  $\sigma_z$ . Let us stress that no one except Alice (holder of the second qubit from the singlet pair) can use the transmitted information  $b$  to learn some information on Bob’s selection. Similarly as in other quantum key distribution protocols (see [46] and the references therein), a potential eavesdropper necessarily introduces errors in the raw key. Any attempt to modify the qubit sent to Bob or the transmitted bit  $b$  leads to a degradation

of the performance of the discrimination scheme that can be verified by uncovering part of the raw key. The analysis of practical robustness of this protocol is beyond the scope of this work.

### ACKNOWLEDGMENTS

M.S. acknowledges support by the Operational Program Education for Competitiveness-European Social Fund (Project No. CZ.1.07/2.3.00/30.0004) of the Ministry of Education, Youth and Sports of the Czech Republic. M.Z. acknowledges the support of Projects No. VEGA 2/0125/13 (QUICOST), No. APVV-0646-10 (COQI), and No. GAČR P202/12/1142.

### APPENDIX: PROOF OF LEMMA 2

Suppose that a test procedure  $\mathcal{T}$  specified by operators  $H_j^{(c)}$   $c \in \{\mathcal{M}, \mathcal{N}, f\}$ ,  $j \in \Omega$  leads to certain values of  $p_s$ ,  $p_e$ , and  $p_f$ . In the first step, our aim is to design a different test procedure  $\mathcal{T}'$  that would give the same values of  $p_s$ ,  $p_e$ , and  $p_f$  that can be interpreted as a mixture of the subproblem defined in Sec. V A and a discrimination of two identical measurements.

We define  $H_j^{(c)} = \sum_{k \in \omega} Q_k H_j^{(c)} Q_k + (1 - P) H_j^{(c)} (1 - P)$ , where we recall  $P = \sum_{k \in \omega} Q_k$ . By definition, operators  $H_j^{(c)}$  are positive semidefinite and they obey the following normalization:

$$\forall j \quad \sum_c H_j^{(c)} = \sum_k Q_k \rho Q_k + (1 - P) \rho (1 - P) \equiv \rho',$$

where we defined positive-semidefinite operator  $\rho'$ . Moreover,  $\text{tr}[\rho'] = 1$ , so we showed that operators  $H_j^{(c)}$  specify a valid test procedure.

Due to  $M_k \geq Q_k$  and  $\sum_k M_k = I$ , we have  $M_k = I - \sum_{l \neq k} M_l \leq I - \sum_{l \neq k} Q_l$ , which is equivalent to

$$M_k - Q_k \leq I - P. \tag{A1}$$

As a consequence,

$$Q_l M_k Q_l = \delta_{kl} Q_l \tag{A2}$$

because for  $k \neq l$  we get  $0 \leq Q_l M_k Q_l \leq 0$  and case  $k = l$  follows from the definition of  $Q_k$ . Finally, using Eq. (A1) we get  $(1 - P)(M_k - Q_k)(1 - P) = M_k - Q_k$ , which is useful to write as

$$M_k = Q_k M_k Q_k + (1 - P) M_k (1 - P), \tag{A3}$$

where we used the above identities and  $Q_k Q_l = \delta_{kl} Q_l$ .

Analogously, one can derive relations (A2) and (A3) for elements  $N_i$ . This enables us to show that the test procedure  $\mathcal{T}'$  leads to the same values of  $p_s$ ,  $p_e$ , and  $p_f$ , because the conditional probabilities  $p(c|\mathcal{M}, \mathcal{T})$ ,  $p(c|\mathcal{N}, \mathcal{T})$  do not change. Indeed, we have

$$\begin{aligned} p(c|\mathcal{M}, \mathcal{T}') &= \sum_j \sum_k \text{tr}[Q_k H_j^{(c)} Q_k M_j] \\ &+ \sum_j \text{tr}[(1 - P) H_j^{(c)} (1 - P) M_j] \end{aligned}$$

$$\begin{aligned}
&= \sum_j \text{tr}[H_j^{(c)} Q_j M_j Q_j] \\
&\quad + \sum_j \text{tr}[H_j^{(c)} (1-P) M_j (1-P)] \\
&= \sum_j \text{tr}[H_j^{(c)} M_j] = p(c|\mathcal{M}, \mathcal{T}), \quad (\text{A4})
\end{aligned}$$

where we used Eqs. (A2) and (A3). Analogously, one can show  $p(c|\mathcal{N}, \mathcal{T}') = p(c|\mathcal{N}, \mathcal{T})$ . Let us introduce Hilbert spaces  $\tilde{\mathcal{H}} = (1-P)\mathcal{H}$ ,  $\bar{\mathcal{H}} = P\mathcal{H}$  specified by the projector  $P$  and its complement. From the assumptions of the theorem we have that  $\tilde{\mathcal{M}}$  and  $\tilde{\mathcal{N}}$  form a measurement on the Hilbert space  $\tilde{\mathcal{H}}$ . If  $P\rho P = 0$ , then  $H_j^{(c)} = H_j^{(c)}$  and it already specifies a discrimination procedure for  $\tilde{\mathcal{M}}$  and  $\tilde{\mathcal{N}}$  in  $\tilde{\mathcal{H}}$ . Similarly, if  $(1-P)\rho(1-P) = 0$  then  $H_j^{(c)}$  specifies a discrimination procedure for  $\bar{\mathcal{M}} = \bar{\mathcal{N}} \leftrightarrow \{Q_j\}_{j=1}^n$  in  $\bar{\mathcal{H}}$ . In the rest of the cases, we define  $\lambda = \text{tr}[P\rho]$  and

$$\begin{aligned}
\tilde{\rho} &= \frac{(1-P)\rho(1-P)}{1-\lambda}, \quad \bar{\rho} = \frac{1}{\lambda} \sum_k Q_k \rho Q_k, \\
\tilde{H}_j^{(c)} &= \frac{(1-P)H_j^{(c)}(1-P)}{1-\lambda}, \quad \bar{H}_j^{(c)} = \frac{1}{\lambda} \sum_k Q_k H_j^{(c)} Q_k.
\end{aligned} \quad (\text{A5})$$

It is now easy to see that operators  $\tilde{H}_j^{(c)}$  describe a valid discrimination procedure  $\tilde{\mathcal{T}}$  for  $\tilde{\mathcal{M}}$  and  $\tilde{\mathcal{N}}$ , while  $\bar{H}_j^{(c)}$  do the same for  $\bar{\mathcal{M}}$  and  $\bar{\mathcal{N}}$ . Using definitions (A5), the conditional probabilities can be rewritten [see also Eq. (A4)] as

$$\begin{aligned}
p(c|\mathcal{M}, \mathcal{T}') &= (1-\lambda) \sum_j \text{tr}[\tilde{H}_j^{(c)} M_j] + \lambda \sum_j \text{tr}[\bar{H}_j^{(c)} M_j] \\
&= (1-\lambda)p(c|\tilde{\mathcal{M}}, \tilde{\mathcal{T}}) + \lambda p(c|\bar{\mathcal{M}}, \bar{\mathcal{T}}). \quad (\text{A6})
\end{aligned}$$

As a consequence, we have

$$\begin{aligned}
p_s &= (1-\lambda)\tilde{p}_s + \lambda\bar{p}_s, \\
p_e &= (1-\lambda)\tilde{p}_e + \lambda\bar{p}_e, \\
p_f &= (1-\lambda)\tilde{p}_f + \lambda\bar{p}_f.
\end{aligned} \quad (\text{A7})$$

Thus, performance of any test  $\mathcal{T}$  can be also attained by a suitable test  $\mathcal{T}'$ , which naturally defines operators  $\tilde{H}_j^{(c)}$ ,  $\bar{H}_j^{(c)}$  for discrimination of measurements  $\tilde{\mathcal{M}}$ ,  $\tilde{\mathcal{N}}$  and  $\bar{\mathcal{M}}$ ,  $\bar{\mathcal{N}}$ , respectively. Moreover, also the opposite holds, i.e., every properly normalized set of operators  $\tilde{H}_j^{(c)}$ ,  $\bar{H}_j^{(c)}$  and a coefficient  $0 \leq \lambda \leq 1$  defines a valid test  $\mathcal{T}'$ .

Next, we want to show that in order to maximize probability of success  $p_s$  it suffices to consider tests  $\mathcal{T}''$  with  $\lambda = 0$ , i.e.,  $P\rho''P = 0$ . Such tests of the unknown measurement use input states that do not probe the subspace of the Hilbert space  $\mathcal{H}$  defined by projector  $P$ .

Since any pair of measurements can be discriminated at least as good as two indistinguishable measurements, we can find operators  $\hat{H}_j^{(c)}$  for discrimination of  $\tilde{\mathcal{M}}$ ,  $\tilde{\mathcal{N}}$  with failure probability  $\hat{p}_f = \bar{p}_f$  and  $\hat{p}_s \geq \bar{p}_s$ . A test  $\mathcal{T}''$  defined by operators  $H_j^{(c)} = (1-\lambda)\tilde{H}_j^{(c)} + \lambda\hat{H}_j^{(c)}$  has the same failure probability  $p_f$  as the test  $\mathcal{T}'$ , but it has a higher success probability  $(1-\lambda)\tilde{p}_s + \lambda\hat{p}_s \geq p_s$ . Moreover,  $P H_j^{(c)} P = 0$  implies  $P\rho''P = 0$ , so we showed that it suffices to consider only tests  $\mathcal{T}''$  or in other words it suffices to solve the discrimination problem for measurements  $\tilde{\mathcal{M}}$ ,  $\tilde{\mathcal{N}}$  instead of the original problem. The optimal discrimination procedure is the same in both cases except for being formally defined on a bigger Hilbert space.

- 
- [1] *Quantum State Estimation*, Lecture Notes in Physics, edited by M. Paris and J. Řeháček (Springer, Berlin, 2004), Vol. 649, pp. 417–465.
- [2] A. Chefles, *Contemp. Phys.* **41**, 401 (2000).
- [3] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, North-Holland Series in Statistics and Probability 1 (North Holland, Amsterdam, 1982).
- [4] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).
- [5] L. K. Grover, *Phys. Rev. Lett.* **79**, 325 (1997).
- [6] M. F. Pusey, J. Barrett, and T. Rudolph, *Nat. Phys.* **8**, 475 (2012).
- [7] J. Fiurášek and M. Ježek, *Phys. Rev. A* **67**, 012321 (2003).
- [8] A. Hayashi, T. Hashimoto, and M. Horibe, *Phys. Rev. A* **78**, 012333 (2008).
- [9] E. Bagan, R. Muñoz-Tapia, G. A. Olivares-Rentería, and J. A. Bergou, *Phys. Rev. A* **86**, 040303 (2012).
- [10] H. Sugimoto, T. Hashimoto, M. Horibe, and A. Hayashi, *Phys. Rev. A* **80**, 052322 (2009).
- [11] A. Acín, *Phys. Rev. Lett.* **87**, 177901 (2001).
- [12] G. M. D'Ariano, P. LoPresti, and M. G. A. Paris, *Phys. Rev. Lett.* **87**, 270404 (2001).
- [13] M. F. Sacchi, *Phys. Rev. A* **71**, 062340 (2005).
- [14] M. F. Sacchi, *J. Opt. B* **7**, S333 (2005).
- [15] G. Wang and M. Ying, *Phys. Rev. A* **73**, 042301 (2006).
- [16] R. Duan, Y. Feng, and M. Ying, *Phys. Rev. Lett.* **103**, 210501 (2009).
- [17] M. Piani and J. Watrous, *Phys. Rev. Lett.* **102**, 250501 (2009).
- [18] A. W. Harrow, A. Hassidim, D. W. Leung, and J. Watrous, *Phys. Rev. A* **81**, 032339 (2010).
- [19] M. Ziman and M. Sedláč, *J. Mod. Opt.* **57**, 253 (2010).
- [20] T. Hashimoto, A. Hayashi, M. Hayashi, and M. Horibe, *Phys. Rev. A* **81**, 062327 (2010).
- [21] G. Chiribella, G. M. D'Ariano, and M. Roetteler, *New J. Phys.* **15**, 103019 (2013).
- [22] Z. Ji, Y. Feng, R. Duan, and M. Ying, *Phys. Rev. Lett.* **96**, 200401 (2006).
- [23] M. Ziman and T. Heinosaari, *Phys. Rev. A* **77**, 042321 (2008).
- [24] M. Ziman, T. Heinosaari, and M. Sedláč, *Phys. Rev. A* **80**, 052102 (2009).
- [25] A. Laing, T. Rudolph, and J. L. O'Brien, *Phys. Rev. Lett.* **102**, 160502 (2009).

- [26] J. Fiurášek and M. Mičuda, *Phys. Rev. A* **80**, 042312 (2009).
- [27] M. Ziman, *Phys. Rev. A* **77**, 062112 (2008).
- [28] G. Chiribella, G. M. D'Ariano, and P. Perinotti, *Phys. Rev. Lett.* **101**, 180501 (2008).
- [29] G. Chiribella, G. M. D'Ariano, and P. Perinotti, *Phys. Rev. Lett.* **101**, 060401 (2008).
- [30] G. Gutoski and J. Watrous, *Proceedings of the 39th Annual ACM Symposium on Theory of Computation, June 2007* (ACM, New York, 2007).
- [31] M. Choi, *Linear Algebra Applicat.* **10**, 285 (1975).
- [32] A. Jamiolkowski, *Rep. Math. Phys.* **3**, 275 (1972).
- [33] C. H. Bennett and S. J. Wiesner, *Phys. Rev. Lett.* **69**, 2881 (1992).
- [34] E. Haapasalo, M. Sedlák, and M. Ziman, *Phys. Rev. A* **89**, 062303 (2014).
- [35] A. Jenčová, *J. Math. Phys.* **55**, 022201 (2014).
- [36] V. Paulsen, *Completely Bounded Maps and Operator Algebras* (Cambridge University Press, Cambridge, UK, 2003).
- [37] I. D. Ivanovic, *Phys. Lett. A* **123**, 257 (1987).
- [38] D. Dieks, *Phys. Lett. A* **126**, 303 (1988).
- [39] A. Peres, *Phys. Lett. A* **128**, 19 (1988).
- [40] G. Jaeger and A. Shimony, *Phys. Lett. A* **197**, 83–87 (1995).
- [41] T. Rudolph, R. W. Spekkens, and S. Turner, *Phys. Rev. A* **68**, 010301(R) (2003).
- [42] H. Sugimoto, Y. Taninaka, and A. Hayashi *Phys. Rev. A* **86**, 042311 (2012).
- [43] J. Bae, *New J. Phys.* **15**, 073037 (2013).
- [44] R. B. M. Clarke, V. M. Kendon, A. Chefles, S. M. Barnett, E. Riis, and M. Sasaki *Phys. Rev. A* **64**, 012303 (2001).
- [45] M. Miková, M. Sedlák, I. Straka, M. Mičuda, M. Ziman, M. Ježek, M. Dušek, and J. Fiurášek, *Phys. Rev. A* **90**, 022317 (2014).
- [46] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, *Rev. Mod. Phys.* **81**, 1301 (2009).