Single-shot labeling of quantum observables

Nidhin Sudarsanan Ragini and Mario Ziman
RCQI, Institute of Physics, Slovak Academy of Sciences, Dúbravská cesta 9, 84511 Bratislava, Slovakia

(Received 16 January 2024; accepted 10 April 2024; published 9 May 2024)

We identify and study a particular class of distinguishability problems for quantum observables (positive-operator-valued measures), in which observables with permuted effects are involved, which we call the labeling problem. Consequently, we identify binary observables, which can be labeled perfectly. In this work, we study these problems in the single-shot regime.

DOI: 10.1103/PhysRevA.109.052415

I. INTRODUCTION

The lack of understanding of the individual action performed by a measurement apparatus is the essence of most quantum controversies. Quantum theory provides a way to describe its probabilistic aspects for all practical purposes. In particular, when one is interested in predicting the statistics of observed outcomes, each measurement outcome is fully determined by a positive operator known as an effect (see, for example, [1]). An effect itself does not explain the physics or form of the recorded event, but it captures the probabilities of its occurrence. In order to describe measurement statistics, we assign labels and effects for all outcomes. While the effects associated with the measurement apparatus can be verified experimentally, the choice of labels cannot be tested by the measurement apparatus itself.

In this paper we consider the situation when the mathematical description of effects is known but their connection to the labels is lost. The issue of outcome labeling we address is the task of assigning a mathematical description to the unlabeled outcomes, assuming the effects forming the observable are known. For example, consider a black box with green and red diodes. Each time the measurement is accomplished one of the diodes shines and the outcome is recorded. We know the measurement apparatus is designed to determine whether the measured atom is in its excited state or its ground state. It is not perfect and the associated effects are $E = \eta|e\rangle\langle e|$ and $G = (1-\eta)|e\rangle\langle e| + |g\rangle\langle g|$, respectively. The parameter η quantifies the imperfections in the registration of the excited state. Now is it possible to identify how the colors and effects are paired? How many measurement runs are needed? It is straightforward to realize that measuring the atom prepared in the ground state can lead only to an outcome described by the effect G. Therefore, the diode that shines in this situation is necessarily described by the effect G. The task of outcome labeling is accomplished and the labels (shining diodes) are identified with effects.

It turns out that this investigation is a special case of the discrimination problem whose origin stems from the seminal work of Helstrom [2], who investigated the discrimination problem for quantum states. In particular, in its simplest form we are given a black box device known to be either A or B. Our task is to design an experiment enabling us to conclude whether it is A or B in a single run of the experiment. The conclusions might be of various relevance or significance. One of the measures is their averaged error probability, known to justify operationally the choice of trace norm as the measure of statistical distance between A and B [2]. The outcome labeling problem of an n-valued measurement aims to distinguish among $n!$ permutations of different labelings and as such among $n!$ devices. Unfortunately, very little is known about decision problems beyond the binary case.

The performance of measurements is crucial for the successful development of quantum technologies (see, e.g., Ref. [3]). Therefore, the elementary as well as foundational question of discrimination of quantum measurements has attracted researchers, and various versions and aspects of the problem have been studied in Refs. [4–9]. In all of them, either the labeling is assumed to be known or the equivalence classes ignoring the labelings are considered as descriptions of unknown (and unlabeled) measurement devices. The question of assigning unknown labels was not addressed explicitly. The closest is the work in Ref. [10], where authors analyzed the minimum-error discrimination of two symmetric, informationally complete (SIC) positive-operator-valued measures (POVMs), one of them being an arbitrary but fixed permutation of the other. The question of labeling is about identification of the permutation. However, in [10] the permutation is fixed; thus, the problem is a binary decision problem that can be investigated in single-shot settings. However, the labeling problem for the SIC POVM aims to distinguish among $d^2!$ different permutations (d is the Hilbert space dimension of the measured system).

The paper is organized as follows. In Sec. II the general problem of labeling is formulated. Section III focuses on perfect labeling for single-shot settings. Minimum-error labeling and unambiguous labeling are studied in Secs. IV and V, respectively. Section VI introduces the concept of antilabeling. We summarize and discuss the main results in Sec. VII.

II. FORMULATION OF the LABELING PROBLEM

A. Mathematical tools

The mutually exclusive outcomes $x_1, \dots, x_n$ of n-valued quantum measurements are represented by effects, i.e.,

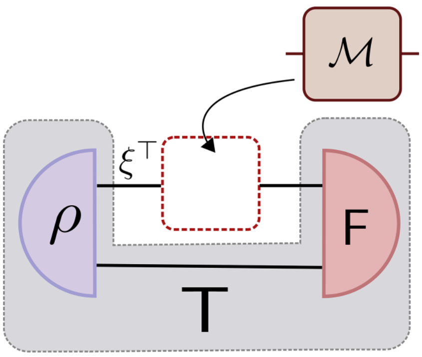

FIG. 1. To measure a channel $\mathcal{M}$, we send a test state ρ through an extended channel and perform a test observable F on the evolved state. We can identify that the pair of the state and observable constitutes the measurement procedure event and as such determines a quantum tester T. The reduced subsystem entering the unknown channel is in the state $\xi^{\top}$, where the transposition is used just to obtain simpler expressions in the other formulas.

positive operators $M_1, \ldots, M_n$, respectively, satisfying the normalization $M_1 + \cdots + M_n = I$. For most of the situations we can safely say the measurements are a collection of effects; however, rigorously, the measurements are mappings assigning effects for particular outcomes. Let us denote by Ω the ordered set of outcome labels $\Omega := \{x_1, \ldots, x_n\}$ and by $\mathcal{E}(\mathcal{H})$ the set of all effects. Then a quantum measurement is described by an observable, as a normalized POVM $\mathsf{M} : \Omega \to \mathcal{E}(\mathcal{H})$ with $\mathsf{M}(x_j) = M_j$. Alternatively, we can think of observables as measure-and-prepare channels $\mathcal{M}$ transforming input states of the measured system ϱ into output states (effectively probability distributions) of n-dimensional systems

$$\mathcal{M}(\rho) = \sum_{x_k \in \Omega} \mathrm{Tr}(\rho M_k)|x_k\rangle\langle x_k|. \tag{1}$$

Once we have the channel form, the Choi-Jamiołkowski isomorphism [11,12] between completely positive maps and positive operators associates the observable M with a positive operator

$$\mathfrak{M} = (\mathcal{I} \otimes \mathcal{M})[\Psi_+] = \sum_k M_k^{\top} \otimes |x_k\rangle\langle x_k|, \tag{2}$$

where Ψ_+ is the (unnormalized) maximally entangled state on $\mathcal{H} \otimes \mathcal{H}$ defined by $\Psi_+ := |\psi_+\rangle\langle\psi_+|$ with $|\psi_+\rangle := \sum_{j=1}^{d} |j\rangle \otimes |j\rangle$ and $|1\rangle, \ldots, |d\rangle$ form an orthonormal basis of $\mathcal{H}$.

The mathematical framework describing the measurements of processes was introduced and developed in Refs. [13–15]. The most general single-shot experiment measuring the properties of quantum processes (channels) is illustrated in Fig. 1. In particular, the channel under consideration $\mathcal{M}$ acts on the subsystem of a bipartite test state ϱ and subsequently the output is measured by a bipartite test observable F leading to outcome α. Suppose the input subsystem of the channel is associated with the Hilbert space $\mathcal{H}$ and its output is associated with Hilbert space $\mathcal{K}$; thus, the channel $\mathcal{M}$ transforms operators $\mathcal{L}(\mathcal{H})$ into $\mathcal{L}(\mathcal{K})$. Further, let us denote by $\mathcal{H}_{\mathrm{anc}}$ the Hilbert space associated with the ancillary subsystem. Then the test state ϱ is a density operator from $\mathcal{L}(\mathcal{H}_{\mathrm{anc}} \otimes \mathcal{H})$ and $F_\alpha \in \mathcal{L}(\mathcal{H}_{\mathrm{anc}} \otimes \mathcal{K})$ is the effect associated with the outcome α of the test observable F. The outcome α occurs with probability $p_\alpha = \mathrm{tr}[(\mathcal{I} \otimes \mathcal{M})[\varrho]F_\alpha]$.

For any state ϱ the Choi-Jamiołkowski isomorphism defines a completely positive linear map $\mathcal{R}_\varrho : \mathcal{L}(\mathcal{H} \otimes \mathcal{H}) \to \mathcal{L}(\mathcal{H}_{\mathrm{anc}} \otimes \mathcal{H})$ such that $\varrho = (\mathcal{R}_\varrho \otimes \mathcal{I})[\Psi_+]$ (see Ref. [15]). Using its dual $\mathcal{R}_\varrho^*$, we obtain $p_\alpha = \mathrm{tr}((\mathcal{I} \otimes \mathcal{M})(\mathcal{R}_\varrho \otimes \mathcal{I})[\Psi_+]F_\alpha) = \mathrm{tr}((\mathcal{I} \otimes \mathcal{M})[\Psi_+](\mathcal{R}_\varrho^* \otimes \mathcal{I})[F_\alpha])$. In the last expression we may identify the Choi-Jamiołkowski representation $\mathfrak{M}$ of the tested channel $\mathcal{M}$ and introduce the so-called process effect $\mathsf{T}_\alpha = (\mathcal{R}_\varrho^* \otimes \mathcal{I})[F_\alpha]$ determining the probability of outcome α for any channel $\mathcal{M}$. As a result we obtain the description of the process measurement by the so-called tester being a collection of positive operators $\mathsf{T}_1, \ldots, \mathsf{T}_m$ $[\mathsf{T}_\alpha \in \mathcal{L}(\mathcal{H} \otimes \mathcal{K})]$ satisfying the normalization $\mathsf{T}_1 + \cdots + \mathsf{T}_m = \xi \otimes I$, where $\xi = [\mathrm{tr}_{\mathcal{K}}(\rho)]^{\top}$ is the transposition of the reduced state of the subsystem passing through the unknown channel $\mathcal{M}$. We recover the Born-like formula $p_\alpha = \mathrm{tr}(\mathfrak{M}\mathsf{T}_\alpha)$ for the outcome probabilities.

Let us stress that for testing channels of measure-and-prepare form (2) the tester operators can be assumed, without loss of generality, to be of the form $\mathsf{T}_\alpha = \sum_k \mathsf{H}_k^{(\alpha)} \otimes |x_k\rangle\langle x_k|$ [7], where for all k the normalization $\sum_\alpha \mathsf{H}_k^{(\alpha)} = \xi$ is satisfied and all $\mathsf{H}_k^{(\alpha)} \in \mathcal{L}(\mathcal{H})$ are positive operators.

B. Labeling problem

The question of labeling takes place when the information on the mapping $x_k \mapsto M_k$ is lost but the collection of effects is known. If all the effects are different, then for an n-valued observable there are $n!$ different permutations of how the outcomes can be paired with the effects. Therefore, the goal of the labeling problem is to discriminate among $n!$ permutations. Let us fix the order of effects and denote by M_σ the observable obtained by σ permutation. The result of labeling is the identification of σ. In what follows we assume that all the permutations are equally likely, and thus the *a priori* information on σ is represented by the probability distribution $\pi(\sigma) = 1/n!$.

In the language of measure-and-prepare channels the labeling takes the form of discrimination among channels, whose actions are given by $\mathcal{M}_\sigma(A) = \sum_k \mathrm{tr}(A\mathsf{M}_k)|x_{\sigma(k)}\rangle\langle x_{\sigma(k)}| = P_\sigma \mathcal{M}_{\mathrm{id}}(A) P_\sigma^{\dagger}$, where $P_\sigma = \sum_k |x_{\sigma(k)}\rangle\langle x_k|$ is the permutation operator and id stands for the identity permutation. Consequently, for the associated Choi operators it holds that $\mathfrak{M}_\sigma = (I \otimes P_\sigma)\mathfrak{M}_{\mathrm{id}}(I \otimes P_\sigma)^{\dagger}$. Let us denote by $\mathsf{T}_1, \ldots, \mathsf{T}_m$ ($m = n!$) the operators forming the tester used to discriminate $\mathfrak{M}_1, \ldots, \mathfrak{M}_m$, respectively. The performance is characterized by conditional probabilities $p(\sigma|\sigma') = \mathrm{tr}(\mathfrak{M}_\sigma \mathsf{T}_{\sigma'})$. Intuitively, the closer this conditional probability is to $\delta_{\sigma\sigma'}$ the better, but there are various ways to quantify the performance. The most common one is the average error probability $p_{\mathrm{error}} = 1 - \frac{1}{n!}\sum_\sigma p(\sigma|\sigma)$, where $\frac{1}{n!}\sum_\sigma p(\sigma|\sigma)$ is the corresponding average success probability. Minimizing this function, one finds minimum-error labeling. Introducing an additional inconclusive outcome $\mathsf{T}_{\mathrm{inconclusive}}$ and requiring $p(\sigma|\sigma') = 0$ whenever $\sigma \neq \sigma'$ is known as unambiguous labeling, for which the conclusions are error-free, but the labeling fails with probability $p_{\mathrm{failure}} = \frac{1}{n!}\mathrm{tr}(\sum_\sigma \mathfrak{M}_\sigma \mathsf{T}_{\mathrm{inconclusive}})$. When $p_{\mathrm{error}} = p_{\mathrm{failure}} = 0$ we speak about perfect labeling.

III. PERFECT LABELING

Perfect discrimination of two quantum channels $\mathcal{M}_1$ and $\mathcal{M}_2$ in a single use corresponds to the scenario when there is no error in each of the conclusions associated with the discrimination task, that is, the average error $p_e = \mu\,\mathrm{tr}(\mathfrak{C}_1\mathsf{T}_2) + (1-\mu)\mathrm{tr}(\mathfrak{C}_2\mathsf{T}_1)$, where μ characterizes the *a priori* bias between the channels. In our case there is no bias and thus both labelings are equally likely and $\mu = \frac{1}{2}$. The condition itself (independently of μ) translates to the existence of testers $\{\mathsf{T}_1, \mathsf{T}_2\}$ with normalization $\xi \otimes I$ satisfying [15,16] the identity

$$\mathfrak{M}_1(\xi \otimes I)\mathfrak{M}_2 = O \tag{3}$$

for some normalization ξ. Let us recall that this condition is necessary and sufficient for the single-shot perfect distinguishability of two quantum channels and replaces the orthogonality condition $\varrho_1\varrho_2 = O$ for perfect discrimination of two quantum states. In this section we evaluate the validity of this condition for the labeling problem.

A. Binary measurements

Let us start with the labeling of so-called binary observables characterized by a pair of effects M_1 and $M_2 = I - M_1$. There are only two permutations of outcomes determining observables M_{12} and M_{21} associated with the Choi operators

$$\mathfrak{M}_{12} = M_1^\top \otimes |1\rangle\langle 1| + M_2^\top \otimes |2\rangle\langle 2|, \tag{4}$$

$$\mathfrak{M}_{21} = M_1^\top \otimes |2\rangle\langle 2| + M_2^\top \otimes |1\rangle\langle 1|. \tag{5}$$

Theorem 1. A binary observable associated with effects M_1 and M_2 ($M_1 \neq M_2$) can be perfectly labeled in a single shot if and only if at least one of the effect operators is rank deficient.

Proof. If $M_1 = M_2$, then the observable is labeled (in fact, $M_1 = M_2 = 1/2I$, where I is the identity on a d-dimensional system). If $M_1 \neq M_2$, the perfect discrimination condition $\mathfrak{M}_{12}(\xi \otimes I)\mathfrak{M}_{21} = O$ implies that

$$M_1^\top\, \xi\, M_2^\top = O. \tag{6}$$

Suppose both operators M_1 and M_2 are invertible. Then also their transpositions are invertible and we may apply inverse operators to obtain $\xi = O$, contradicting the assumption that ξ is a density operator. In other words, no binary observable with full rank effects can satisfy the perfect discrimination condition. In contrast, suppose that there exists at least one state vector $|\varphi\rangle$ such that $M_1|\varphi\rangle = 0$ or $M_2|\varphi\rangle = 0$. Then $M_1^\top|\varphi\rangle\langle\varphi|M_2^\top = 0$ and thus the perfect discrimination condition holds. ■

This result is clear. If the measurement device is used only once, at most one of the outcomes is recorded. The purity of successful probe state $\xi = |\varphi\rangle\langle\varphi|$ implies the potential auxiliary system is uncorrelated, rendering it irrelevant for labeling. Consequently, a single click can only reveal some information about the recorded outcome. This information is "negative" in the following sense. Recording an outcome implies the effect is not the one having the vector φ in its kernel. For example, if $M_1|\varphi\rangle = 0$, then the recorded outcome is not described by the effect M_1. In the considered binary case this implies the unrecorded outcome can be associated with M_1 and the recorded one with M_2.

B. Nonbinary measurements

We refer to an observable with more than two outcomes as nonbinary. The negative logic used to label binary observables cannot be extended to nonbinary ones. Knowing the recorded outcome is not described by an effect M_1 does not in general help one identify either the recorded or the unrecorded outcomes. It is straightforward to see that the recording of a single outcome of the unlabeled device cannot reveal information about all the outcomes. As such, we can formulate the following theorem.

Theorem 2. Perfect labeling of a nonbinary observable with effects $M_1, \ldots, M_n$ (all different, i.e., $M_j \neq M_k$ for all $j \neq k$) requires at least $n-1$ uses of the measurement device.

Proof. Using the device once, we record exactly one outcome that can be the subject of labeling. In general, only recorded outcomes can be labeled; thus the n-valued observable demands at least $n-1$ uses. ■

Let us note that by identifying $n-1$ outcomes of the considered observable we obtain a binary observable described by the effects $X_1 = M_1$ and $X_2 = M_2 + \cdots + M_n$. This binarization is a purely abstract construction, because for the unlabeled device itself we do not know which outcomes are identified. However, if it happens that the conditions of Theorem 1 are met and X_2 is rank deficient, then we can label the outcome of a nonbinary observable associated with M_1. In the second use of the device we may choose a different binarization and potentially label a different outcome. Altogether there are n different binarizations of this form, but exploiting $n-1$ of them is sufficient to accomplish the goal of perfect labeling.

Let us stress that the condition that all the effects are different is important. The situation might be different if some of the outcomes are described by the same effects. In what follows we show an example of nonbinary observable for which all outcomes can be labeled in a single shot. Consider an observable with the effects $M_1 = |\varphi\rangle\langle\varphi|$ and $M_2 = M_3 = \cdots = M_n = \frac{1}{n-1}(I - |\varphi\rangle\langle\varphi|) \equiv M_1^\perp$. Using the probe state $|\varphi\rangle$, the recorded outcome is necessarily associated with the effect M_1, but simultaneously the remaining effects are described by the effect $M_1^\perp$. One use of the observable is sufficient to accomplish perfect labeling. It is straightforward to design similar pathological examples of n-valued observables that can be perfectly labeled in m uses for all $m \leqslant n$. However, for generic observables the number of outcomes determines the number of uses required to completely label all the outcomes.

It also follows that the single-shot setting allows for partial labeling of the recorded outcome. For example, consider an unlabeled observable with effects $M_1, \ldots, M_n$ with m identical effects, i.e., $M_1 = M_2 = \cdots = M_m$ for $m < n$ (in the case $m = n$ the outcomes are labeled). Assuming $\tilde{M} = M_1 + \cdots + M_m = mM_1$ has eigenvalue one, then there exists a vector state $|\varphi\rangle$ such that $\tilde{M}|\varphi\rangle = |\varphi\rangle$. In other words, if the state $|\varphi\rangle$ is measured, then necessarily the recorded outcome is associated with effect M_1 and the task of partial labeling is accomplished. Let us note that using the same state $|\varphi\rangle$ also in the second use does not guarantee labeling of another unlabeled outcome.

Proposition 1. An effect $E \in \{M_1, \dots, M_n\}$ can be partially labeled in a single use of the observable if and only if there are $1 \leqslant m \leqslant n$ outcomes j such that $M_j = E$ and the largest eigenvalue of the effect E equals $1/m$.

Proof. Let $|\varphi\rangle$ be an eigenvector associated with the eigenvalue $1/m$. Then the probability of observing any of the outcomes $M_j \neq E$ equals $\langle\varphi|(\sum_{j:M_j\neq E} M_j)|\varphi\rangle = \langle\varphi|(I - mE)|\varphi\rangle = 0$. In other words, only outcomes associated with E can be recorded. Therefore, the observed outcome is labeled to the effect E. ■

IV. MINIMUM-ERROR LABELING

In this section we study the limitations of single-shot labeling for general observables. More specifically, we quantify the optimal success probability for labeling given the observable is binary and used only once, which is encapsulated in the following theorem.

Theorem 3. The optimal minimum-error probability for labeling a binary observable associated with the effects $\{M_1, M_2\}$ is given by

$$p_{\mathrm{e}} = \tfrac{1}{2}(1 - \|M_1 - M_2\|_2). \tag{7}$$

Proof. Consider a binary observable with effects M_1 and M_2. Following Eq. (4), we denote by $\mathfrak{M}_{12}$ and $\mathfrak{M}_{21}$ the Choi operators associated with two possible labelings. Let us denote by T_{12} and T_{21} the elements of the tester associated with the labelings $\{M_1, M_2\}$ and $\{M_2, M_1\}$, respectively, and satisfying the normalization condition $\mathsf{T}_{12} + \mathsf{T}_{21} = \xi \otimes I$, where ξ is a density operator. The average error probability reads

$$p_{\mathrm{e}} = \tfrac{1}{2}\mathrm{tr}(\mathsf{T}_{12}\mathfrak{M}_{21}) + \tfrac{1}{2}\mathrm{tr}(\mathsf{T}_{21}\mathfrak{M}_{12}). \tag{8}$$

Using the definitions and the tester's normalization, we obtain

$$\begin{aligned} p_{\mathrm{e}} &= \tfrac{1}{2}\mathrm{tr}[\mathsf{T}_{12}(\mathfrak{M}_{21} - \mathfrak{M}_{12}) + (\xi \otimes I)\mathfrak{M}_{12}] \\ &= \tfrac{1}{2}\mathrm{tr}\big[\mathsf{T}_{12}\big[\big(M_1^T - M_2^T\big) \otimes (|1\rangle\langle 1| - |2\rangle\langle 2|)\big]\big] + \tfrac{1}{2}. \end{aligned}$$

In the spectral form $M_1^T - M_2^T = \sum_x \mu_x |\omega_x\rangle\langle\omega_x|$, where μ_x are real eigenvalues and $|\omega_x\rangle$ are the associated eigenvectors. It follows that the operator $(M_1^T - M_2^T) \otimes (|1\rangle\langle 1| - |2\rangle\langle 2|)$ has $2d$ eigenvalues $\pm\mu_x$ (including multiplicities) and the error probability equals

$$\begin{aligned} p_{\mathrm{e}} = \frac{1}{2} &+ \frac{1}{2}\sum_x \mu_x \langle\omega_x \otimes 1|\mathsf{T}_{12}|\omega_x \otimes 1\rangle \\ &- \frac{1}{2}\sum_x \mu_x \langle\omega_x \otimes 2|\mathsf{T}_{12}|\omega_x \otimes 2\rangle. \end{aligned}$$

In order to minimize the error probability we need to suppress positive terms and maximize the negative one. The negativity or positivity of the terms depends solely on the value of μ_x, because the operator T_{12} is positive. We know there is one positive and one negative term for each nonzero $|\mu_x|$. We set T_{12} to vanish on all positive terms, i.e., $\mathsf{T}_{12}|\omega_x \otimes j\rangle = 0$ if $(-1)^{(j-1)}\mu_x$ is positive. Then

$$p_{\mathrm{e}} = \frac{1}{2} - \frac{1}{2}\sum_x |\mu_x|\langle\omega_x \otimes j_x|\mathsf{T}_{12}|\omega_x \otimes j_x\rangle,$$

where $j_x = 1$ if $\mu_x < 0$ and $j_x = 2$ if $\mu_x > 0$. Due to normalization $\mathsf{T}_{12} \leqslant \xi \otimes I$, it follows that $\langle\omega_x \otimes j|\mathsf{T}_{12}|\omega_x \otimes j\rangle \leqslant \langle\omega_x|\xi|\omega_x\rangle$. The numbers $\langle\omega_x|\xi|\omega_x\rangle$ define a probability; thus the values $q_x = \langle\omega_x \otimes j_x|\mathsf{T}_{12}|\omega_x \otimes j_x\rangle$ form a subnormalized probability distribution. Our goal is to maximize the expression $\sum_x |\mu_x| q_x$ over subnormalized probability distributions. It turns out that the maximum is achieved when $\mathsf{T}_{12} = |\omega_x \otimes j_x\rangle\langle\omega_x \otimes j_x|$ is associated with the largest eigenvalue $|\mu_x|$. More specifically, the minimum-error probability equals

$$p_{\mathrm{e}} = \tfrac{1}{2}(1 - \max_x |\mu_x|) = \tfrac{1}{2}(1 - \|M_1 - M_2\|_2), \tag{9}$$

where $\|\cdot\|_2$ is the operator 2-norm defined as $\|X\|_2 = \sup_{\varphi\neq 0}\langle X\varphi|X\varphi\rangle/\langle\varphi|\varphi\rangle = \lambda_{\max}$, where $\lambda_{\max}$ denotes the largest singular value of X. ■

Because of the normalization $M_1 + M_2 = I$, the identity $M_1 - M_2 = 2M_1 - I$ holds. It follows that the eigenvalues and eigenvectors of M_1, M_2, and $M_1 - M_2$ are related. Indeed, all the operators mutually commute, because $[M_1, M_2] = 0$; hence they share the same system of eigenvectors. Let us denote by λ_x the eigenvalues of M_1 and by κ_x the eigenvalues of M_2. Then $\lambda_x + \kappa_x = 1$ and $\mu_x = \lambda_x - \kappa_x = 2\lambda_x - 1$. Consequently, the maximal value of $|\mu_x|$ is achieved for either the maximal or the minimal values of λ_x. Let us denote by $\lambda_{\max}$ and $\lambda_{\min}$ the maximal and the minimal eigenvalue of M_1. Let us stress that the minimal eigenvalue of M_1 is the maximal one for M_2 and vice versa. It follows that the largest of the distances of the minimal and the maximal eigenvalue from $\frac{1}{2}$ (the value for which $\mu = 0$) determines the value of the minimum-error probability. In particular, $|\mu_x| = 1$ (perfect labeling) only if either $\lambda_{\min} = 0$ or $\lambda_{\max} = 1$. The binary observables with all the eigenvalues of M_1 close to $\frac{1}{2}$ are the ones for which the minimum error is the largest, and thus the labeling is least reliable.

The optimal labeling procedure consists of preparation of the state $|\omega_x\rangle$ being the eigenvector of M_1 associated with the maximal or minimal eigenvalue λ_x maximizing the value of $|\mu_x|$. The effect M_1 is recorded with probability $\lambda_x = \langle\omega_x|M_1|\omega_x\rangle$ and the effect M_2 with probability $\kappa_x = \langle\omega_x|M_2|\omega_x\rangle$. If $\lambda_x > \kappa_x$, then we conclude that the recorded outcome is M_1 and the unrecorded one is M_2. In this case the probability κ_x characterizes the error, and indeed introducing $|\mu_x| = \lambda_x - \kappa_x$ into Eq. (9), we obtain $p_{\mathrm{e}} = [1 - (\lambda_x - \kappa_x)]/2 = \kappa_x$. Similarly, if $\lambda_x < \kappa_x$ we have $|\mu_x| = \kappa_x - \lambda_x$ and $p_{\mathrm{e}} = \lambda_x$, and thus the recorded outcome is labeled by the effect M_2 and the unrecorded by M_1. In other words, in the optimal minimum-error labeling procedure we label the effect with the largest eigenvalue by performing the unlabeled measurement on the associated eigenvector state $|\omega_x\rangle$.

For nonbinary observables we are met with the same basic limitation. A single shot can reveal the identity of unrecorded outcomes only if they are described by an identical effect. In general, we can talk about error probability for the question of partial labeling. More specifically, we use the unlabeled measurement device once and label the recorded outcome by one of the effects $M_1, \dots, M_n$ forming the n-outcome observable. Suppose $M_x \in \{M_1, \dots, M_n\}$ is chosen to be the label for the recorded outcome. Then $p_{\mathrm{e}} = \sum_{j\neq x}\mathrm{tr}(\omega M_j) = \mathrm{tr}(\omega\sum_{j\neq x} M_j) = \mathrm{tr}(\omega(I - M_x)) = 1 - \mathrm{tr}(\omega M_x)$ is the probability of error for labeling the recorded outcome to M_x.

It is minimized if $\mathrm{tr}(\omega M_x)$ is maximal. Finding $\lambda_{\max} = \max_j \|M_j\|$, where $\|\cdot\|$ is the operator norm that equals the largest eigenvalue of M_j, we obtain the minimum-error probability, which reads $p_e = 1 - \lambda_{\max}$ and $\omega = |\omega_{\max}\rangle\langle\omega_{\max}|$, where $|\omega_{\max}\rangle$ is the eigenvector associated with the eigenvalue $\lambda_{\max}$. Let us denote by $M_{\max}$ the effect with eigenvalue $\lambda_{\max}$. In the optimal partial labeling procedure the outcome recorded in the measurement of $|\omega_{\max}\rangle$ is labeled $M_{\max}$.

Example: Biased coin-tossing binary observables

This family is formed by trivial effects $M_1 = qI$ and $M_2 = (1-q)I$. The formula for minimum-error probability implies that $p_e = 1 - \lambda_{\max}$, where $\lambda_{\max}$ is the largest eigenvalue of M_1 and M_2. Assuming $\frac{1}{2} \leqslant q \leqslant 1$, we have $\lambda_{\max} = q$. Interestingly, the minimum error for this trivial observable can achieve any value including perfect discrimination between O and I when $q = 1$. For an unbiased coin-tossing observable ($q = \frac{1}{2}$), the minimum error equals $\frac{1}{2}$. Nevertheless, this is not completely correct. In particular, for the unbiased situation no verification is needed as both outcomes are described by the same effect and there is nothing unknown about the labels. Contrary to the minimum-error formula, but in accordance with Theorem 1, the labeling in this case is trivially free of any error.

V. UNAMBIGUOUS LABELING

The discrimination procedure is unambiguous if so-called no-error conditions are satisfied and an inconclusive result is allowed. For the case of labeling of binary observables this means that

$$\mathrm{tr}(\mathfrak{M}_{12}\mathsf{T}_{21}) = \mathrm{tr}(\mathfrak{M}_{21}\mathsf{T}_{21}) = 0, \tag{10}$$

with the normalization $\mathsf{T}_{12} + \mathsf{T}_{21} + \mathsf{T}_? = \xi \otimes I$. The operator $\mathsf{T}_?$ is identified with the inconclusive result when no labeling is made. The performance of unambiguous labeling is evaluated by the so-called failure probability

$$p_f = \tfrac{1}{2}\mathrm{tr}[\mathsf{T}_?(\mathfrak{M}_{12} + \mathfrak{M}_{21})] = \tfrac{1}{2}\mathrm{tr}[\mathsf{T}_?] \tag{11}$$

associated with the occurrence of an inconclusive result. The goal is to design a tester T_{12}, T_{21}, or $\mathsf{T}_?$ such that the failure probability is minimized. A nontrivial solution is possible if the no-error conditions can be satisfied for nonzero operators T_{12} and T_{21}. It follows that $\mathfrak{M}_{12}$ or $\mathfrak{M}_{21}$ must have nonempty kernels. This is possible only if at least one of the effects M_1 and M_2 is rank deficient. However, Theorem 1 implies that such an observable can be perfectly labeled, and thus $\mathsf{T}_? = O$ ($p_f = 0$) and for binary observables the unambiguous labeling reduces to perfect labeling.

For nonbinary measurements the no-error conditions also require that at least one of the effects is rank deficient. However, the single use of the unlabeled measurement device does not provide sufficient information for the unambiguous labeling of all effects. We may be able to label partially a particular effect E, but all the situations are covered by Theorem 1. It follows that the unambiguous effect labeling is always perfect in the sense that the recorded results are always conclusive.

VI. ANTILABELING

Analyzing the unambiguous labeling for nonbinary measurements, we observe the following. Although the no-error conditions do not allow us to label the rank deficient effects E, we can still use the corresponding state $|\varphi\rangle$, for which $E|\varphi\rangle = 0$ as the probe state. It is straightforward to see that recorded outcome cannot be labeled by the effect E, because its probability in this case vanishes. Instead of assigning a label for the recorded outcome we may exclude all the labels $E \in \{M_1, \dots, M_n\}$, for which $M_j|\varphi\rangle = 0$.

This type of discrimination tasks was named antidistinguishability [17] and following this we will call this process antilabeling. For the binary case, it trivially coincides with perfect labeling; however, for the nonbinary case the question is open. Since in this study we focus on only single-shot protocols, the antilabeling also can be achieved only partially, that is, not all effects can be excluded in a single run of the experiment. The question of partial antilabeling reduces to antilabeling of some fixed effect $E \in \{M_1, \dots, M_n\}$. Rank deficiency is necessary and sufficient for partial antilabeling of such effects. Let us stress that this is not the case for the question of partial labeling.

Consider rank-1 nonbinary measurements, i.e., $M_j = q_j|\varphi_j\rangle\langle\varphi_j|$, such that $|\varphi_j\rangle \neq |\varphi_k\rangle$ if $j \neq k$. In general, these effects cannot be partially labeled unless $q_j = 1$ for some j. However, all of them can be partially antilabeled. It is sufficient to choose a probe state $|\varphi_j^{\perp}\rangle$ and the recorded outcome cannot be described by the effect $M_j = q_j|\varphi_j\rangle\langle\varphi_j|$. This means the recorded outcome is antilabeled not to be M_j. If there are more M_k such that $M_k|\varphi_j^{\perp}\rangle = 0$, then all such effects are also antilabeled and we may conclude that the recorded outcome is not labeled by any of those effects M_k.

VII. CONCLUSION

In this work we introduced the problem of measurement labeling as a special instance of the measurement discrimination problem. We completely characterized the various forms of labeling questions given the unlabeled measurement device is used only once (single shot). In particular, we showed that perfect labeling exists for binary observables if the effects are rank deficient (Theorem 1). For nonbinary measurements the perfect labeling generally requires more uses of the measurement device (Theorem 2); however, we characterized the measurements for which a single shot is sufficient. Proposition 1 specifies conditions when a single shot reveals partial information about the observed outcomes and thus perfectly labels particular outcomes. Further, we evaluated the minimum-error probability for the case of binary measurements. We found [Eq. (9)] it is proportional to the 2-norm of the difference between effects forming the observable and thus by the largest eigenvalue of the effects. Unambiguous labeling for binary measurements reduces to the perfect one. The case of nonbinary measurements motivated us to introduce the concept of antilabeling, which excludes labelings of the recorded outcomes instead of identifying them.

The problem of labeling is a specific question that is not only of foundational interest. Naturally, it refers to situations when the identity of labels is not available, but how can one lose this type of information? It is unlikely someone will question the description of detectors, for which clicks simply mean the particle is detected. However, detectors are used in more sophisticated experiments to measure quantum properties of systems. In fact, we typically combine many devices and clicks of different detectors or their combinations constitute different (more abstract) outcomes. The risk of forgetting the labels does not originate in the physics itself, but in its users. The labeling serves as a relatively simple tool to operationally retrieve the identity of outcomes without the need to access (sometimes invasively) the details of experimental setup.

ACKNOWLEDGMENTS

The authors acknowledge support through Projects No. APVV-22-0570 (DeQHOST) and No. VEGA 2/0183/21 (DESCOM). N.S.R. acknowledges Sk Sazim and Seyed Arash Ghoreishi for discussions and providing insights into selected calculations. N.S.R. acknowledges also Nana Siddhartha Yenamandala, and Leevi Leppäjärvi and Oskari Kerppo for discussions regarding antidistinguishability.

[1] T. Heinosaari and M. Ziman, *The Mathematical Language of Quantum Theory: From Uncertainty to Entanglement* (Cambridge University Press, Cambridge, 2011).
[2] C. W. Helstrom, Quantum detection and estimation theory, J. Stat. Phys. **1**, 231 (1969).
[3] T. Heinosaari, D. Reitzner, and A. Toigo, Anticipative measurements in hybrid quantum-classical computation, Phys. Rev. A **107**, 032612 (2023).
[4] Z. Ji, Y. Feng, R. Duan, and M. Ying, Identification and distance measures of measurement apparatus, Phys. Rev. Lett. **96**, 200401 (2006).
[5] M. Ziman and T. Heinosaari, Discrimination of quantum observables using limited resources, Phys. Rev. A **77**, 042321 (2008).
[6] M. Ziman, T. Heinosaari, and M. Sedlák, Unambiguous comparison of quantum measurements, Phys. Rev. A **80**, 052102 (2009).
[7] M. Sedlák and M. Ziman, Optimal single-shot strategies for discrimination of quantum measurements, Phys. Rev. A **90**, 052312 (2014).
[8] Z. Puchała, Ł. Pawela, A. Krawiec, and R. Kukulski, Strategies for optimal single-shot discrimination of quantum measurements, Phys. Rev. A **98**, 042103 (2018).
[9] Z. Puchała, Ł. Pawela, A. Krawiec, R. Kukulski, and M. Oszmaniec, Multiple-shot and unambiguous discrimination of von Neumann measurements, Quantum **5**, 425 (2021).
[10] A. Krawiec, Ł. Pawela, and Z. Puchała, Discrimination of POVMs with rank-one effects, Quantum Inf. Process. **19**, 428 (2020)
[11] M.-D. Choi, Completely positive linear maps on complex matrices, Linear Algebra Appl. **10**, 285 (1975).
[12] A. Jamiołkowski, Linear transformations which preserve trace and positive semidefiniteness of operators, Rep. Math. Phys. **3**, 275 (1972).
[13] G. Gutoski and J. Watrous, Toward a general theory of quantum games, *Proceedings of the 39th Annual ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, 2007), pp. 565–574.
[14] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Quantum circuit architecture, Phys. Rev. Lett. **101**, 060401 (2008).
[15] M. Ziman, Process positive-operator-valued measure: A mathematical framework for the description of process tomography experiments, Phys. Rev. A **77**, 062112 (2008).
[16] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Memory effects in quantum channel discrimination, Phys. Rev. Lett. **101**, 180501 (2008).
[17] T. Heinosaari and O. Kerppo, Antidistinguishability of pure quantum states, J. Phys. A: Math. Theor. **51**, 365303 (2018).