

Discrimination of symmetric states in operational
probabilistic theory

Kenji Nakahira

Quantum Information Science Research Center,
Quantum ICT Research Institute, Tamagawa University
6-1-1 Tamagawa-gakuen, Machida, Tokyo 194-8610, Japan

Tamagawa University Quantum ICT Research Institute Bulletin, Vol.9, No.1, 15-27, 2019

©Tamagawa University Quantum ICT Research Institute 2019

All rights reserved. No part of this publication may be reproduced in any form or by any means electrically, mechanically, by photocopying or otherwise, without prior permission of the copy right owner.

Discrimination of symmetric states in operational probabilistic theory

Kenji Nakahira

Quantum Information Science Research Center,
Quantum ICT Research Institute, Tamagawa University
6-1-1 Tamagawa-gakuen, Machida, Tokyo 194-8610 Japan
E-mail: nakahira@lab.tamagawa.ac.jp

Abstract—A state discrimination problem in an operational probabilistic theory (OPT) is investigated in diagrammatic terms. It is well-known that, in the case of quantum theory, if a state set has a certain symmetry, then there exists a minimum-error measurement having the same type of symmetry. However, to our knowledge, it is not yet clear whether this property also holds in a more general OPT. We show that it also holds in OPTs, i.e., for a symmetric state set, there exists a minimum-error measurement that has the same type of symmetry. It is also shown that this result can be utilized to optimize over a restricted class of measurements, such as sequential or separable measurements.

I. INTRODUCTION

Operational probabilistic theories (OPTs) and other similar theories, such as generalized probabilistic theories, provide a general operational framework that allows us to better understand the physical structure of quantum theory [1]–[5]. OPTs can be interpreted as a generalization of probability theory, including classical probability theory, quantum theory, and many others (such as the theory of Popescu-Rohrlich boxes [6]). One of the motivations for using OPTs is to investigate quantum processes from an operational point of view, which helps us to deeply understand quantum theory. Another motivation is that an OPT might be useful in developing new physical theories, such as a theory of quantum gravity.

One of the fundamental problems in probability theory is the state discrimination. In the case of quantum theory, a vast number of studies have been carried out to obtain an optimal measurement with respect to some criteria (e.g., [7]–[16]). Although obtaining a closed-form analytical solution for an optimal measurement is generally very difficult, it is known that if a state set has a certain symmetry, then there exists an optimal measurement having the same type of symmetry (e.g., [11]–[14], [17]–[25]). This property allows us to simplify finding an optimal measurement for a symmetric state set analytically and/or numerically. However, to our knowledge, this result has not been extended to a more general OPT. Note that it would not be surprising if this result does not hold in general since the space of states and that of effects are not symmetric in the case of a general system of OPTs, while

they are highly symmetric in the case of any quantum system.

In this paper, we investigate a state discrimination problem in an OPT; we consider the case in which a state set has a certain symmetry, in which case we show that there exists an optimal measurement that has the same type of symmetry. This result can be proved without reference to specific algebraic structures such as Hilbert spaces and operator algebras. We also show that this result can be applied to the discrimination problem over a restricted class of measurements. As examples, we discuss four classes of measurements: sequential, local operations and classical communication (LOCC), separable, and partially transformable (PT). It is worth noting that, in this paper, we restrict our attention to the minimum-error strategy to simplify the discussion. However, the results given in this paper can be easily applied to other various criteria (see [16], [26]) in OPTs.

II. BRIEF SUMMARY OF OPERATIONAL PROBABILISTIC THEORIES (OPTs)

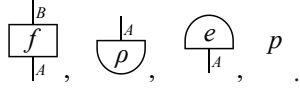
In this section, we briefly review the framework of OPTs. The framework can be explained in several ways, leading to essentially almost the same formalism. The proofs of some of the results are not presented in this paper, which can be found in, e.g., Refs. [1], [4], [27]–[29]. Note that we consider only fixed causal structure. We use diagrammatic representations that are used in Ref. [29] to represent formulae in an intuitive way, which is motivated by the work of Coecke, Abramsky, and others (see, e.g., [30]–[32]).

A. Systems and processes

An OPT consists of a collection of *systems* and a collection of *processes*. Systems and processes respectively represent a physical system (e.g., a photon) and a particular behavior of a physical process (e.g., a beam splitter). Each process has input and output systems. Let $\mathbf{Proc}_{A \rightarrow B}$ be the set of all processes having an input system A and an output system B , referred to as processes from A to B . A *trivial* (or empty) system, denoted by I , is a special system. A process from I to A , denoted like $|\rho\rangle$, is called a *state* of A . Similarly, a process from A to I , denoted like $\langle e|$, is called an *effect* of A . $\mathbf{St}_A := \mathbf{Proc}_{I \rightarrow A}$

and $\mathbf{Eff}_A := \mathbf{Proc}_{A \rightarrow I}$ are, respectively, called the *state space* and *effect space* of system A . A process from I to I is called a *scalar*. Let $\mathbf{Scalar} := \mathbf{Proc}_{I \rightarrow I}$.

In diagrammatic terms, a process $f \in \mathbf{Proc}_{A \rightarrow B}$, a state $|\rho\rangle \in \mathbf{St}_A$, an effect $\langle e| \in \mathbf{Eff}_A$, and a scalar $p \in \mathbf{Scalar}$ are depicted as

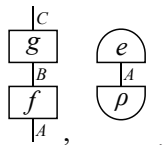


Labeled wires (labels are often omitted) represent systems, while boxes represent processes. Each process has an input wire at the bottom and an output wire at the top. I is represented by 'no wire'. For a scalar, the box will be omitted. Diagrammatic representations can be interpreted such as data flow diagrams, where time increases from the bottom to the top.

Example of (fully) quantum theory For simplicity, we consider only finite-dimensional systems in the examples of quantum theory. Let \mathbf{C} be the set of all complex numbers. Also, let $\mathcal{S}(\mathbf{C}^n)$ and $\mathcal{S}_+(\mathbf{C}^n)$ be, respectively, the sets of all complex Hermitian matrices and all complex positive semidefinite matrices of order n . \mathbf{St}_A and \mathbf{Eff}_A are isomorphic to $\mathcal{S}_+(\mathbf{C}^{N_A})$, where N_A is a natural number determined by a system A . Note that $\mathcal{S}_+(\mathbf{C}^{N_A})$ is a symmetric cone, whose shape is highly symmetric. In particular, $N_I = 1$ holds, i.e., $\mathbf{St}_I = \mathbf{Scalar} \cong \mathcal{S}_+(\mathbf{C}) \cong \mathbf{R}_+$, where \mathbf{R}_+ is the set of all nonnegative real numbers. $\mathbf{Proc}_{A \rightarrow B}$ is isomorphic to the space of all CP maps from $\mathcal{S}(\mathbf{C}^{N_A})$ to $\mathcal{S}(\mathbf{C}^{N_B})$. In the examples of quantum theory, we will identify a process with its corresponding CP map. Also, we will identify a state (or effect) with the corresponding positive semidefinite matrix.

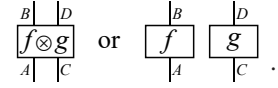
B. Sequential and parallel compositions

Two processes can be composed sequentially whenever the output system of one and the input system of the other are the same. The sequential composition of $f \in \mathbf{Proc}_{A \rightarrow B}$ and $g \in \mathbf{Proc}_{B \rightarrow C}$ is also a process, denoted as $g \circ f \in \mathbf{Proc}_{A \rightarrow C}$. When we write $g \circ f$, we always assume that the output system of f and the input system of g are equal. For any $|\rho\rangle \in \mathbf{St}_A$ and $\langle e| \in \mathbf{Eff}_A$, $\langle e| \circ |\rho\rangle \in \mathbf{Scalar}$ is denoted by $\langle e|\rho\rangle$. $g \circ f$ and $\langle e|\rho\rangle$ are respectively depicted as



Any two systems and processes can be composed in parallel. The parallel composition of two systems, A and B , is a system, denoted by $A \otimes B$. Assume that $I \otimes A = A = A \otimes I$ holds. The parallel composition of $f \in \mathbf{Proc}_{A \rightarrow B}$ and

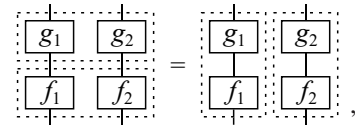
$g \in \mathbf{Proc}_{C \rightarrow D}$ is a process from $A \otimes C$ to $B \otimes D$, denoted as $f \otimes g$. $f \otimes g$ is diagrammatically depicted as



A collection of connected processes will be called a *diagram*. These sequential and parallel compositions are associative, e.g., $h \circ (g \circ f) = (h \circ g) \circ f$ holds for any $f \in \mathbf{Proc}_{A \rightarrow B}$, $g \in \mathbf{Proc}_{B \rightarrow C}$, and $h \in \mathbf{Proc}_{C \rightarrow D}$. Assume that

$$(g_1 \otimes g_2) \circ (f_1 \otimes f_2) = (g_1 \circ f_1) \otimes (g_2 \circ f_2),$$

or diagrammatically



holds for four any processes f_1, f_2, g_1 , and g_2 , where the auxiliary lines (dashed lines) are drawn to guide the eye. For any scalar a and process f , $a \otimes f$ is denoted by af or $a \cdot f$. One can see that $(ag) \circ f = a(g \circ f) = g \circ (af)$, $(af) \otimes h = a(f \otimes h) = f \otimes (ah)$, and $a \circ b = ab$ hold for any scalars a and b and any processes f, g , and h .

Example of quantum theory $N_{A \otimes B} = N_A N_B$ holds for any systems A and B . Since $\mathbf{St}_{I \otimes A} \cong \mathbf{St}_{A \otimes I} \cong \mathbf{St}_A \cong \mathcal{S}_+(\mathbf{C}^{N_A})$ holds from $N_I = 1$, $I \otimes A$ and $A \otimes I$ can be identified with A . For two processes f and g , $g \circ f$ is the CP map satisfying $(g \circ f)[|\rho\rangle] := g[f[|\rho\rangle]]$. In particular, $\langle e|\rho\rangle = \text{Tr}[\langle e| \cdot |\rho\rangle]$ holds (\cdot is the matrix product). For two states $|\rho\rangle$ and $|\sigma\rangle$, $|\rho\rangle \otimes |\sigma\rangle$ is the tensor product of the matrices $|\rho\rangle$ and $|\sigma\rangle$. $f \otimes g$ is the CP map defined as $(f \otimes g)[|\rho\rangle \otimes |\sigma\rangle] := f[|\rho\rangle] \otimes g[|\sigma\rangle]$.

C. Identity processes and discarding effects

An *identity process* on A , denoted by id_A or simply id , is the process satisfying (a) $f \circ \text{id}_A = f = \text{id}_B \circ f$, (b) $\text{id}_A \otimes \text{id}_B = \text{id}_{A \otimes B}$, and (c) $f \otimes \text{id}_I = f = \text{id}_I \otimes f$, where A and B are any systems and f is any process from A to B . Assume that there exists id_A for each system A . Diagrammatically, id_A is depicted as

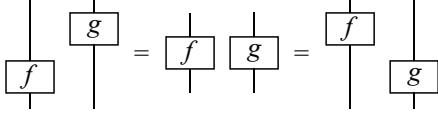


id_I is depicted as empty space. The above property (a) is depicted as:

The diagram shows three equivalent ways to represent the property (a). On the left, a box 'f' is connected to an identity process 'id_A' (represented by a vertical wire) on its input side. In the middle, the box 'f' is connected to an identity process 'id_B' on its output side. On the right, the box 'f' is connected to an identity process 'id_I' on both its input and output sides. The three representations are shown to be equal.

where the auxiliary boxes indicate the identity processes. This, intuitively, implies that the length of lines does not

change diagrams. It also follows from Eq. (1) that, for any processes f and g ,

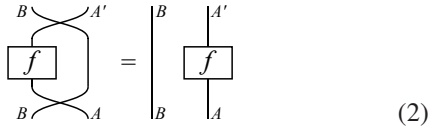


holds. Intuitively, this yields that the vertical shifts of processes do not affect diagrams.

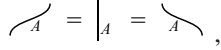
Assume that, for any systems A and B , there exists a process $\times_{A \otimes B} \in \mathbf{Proc}_{A \otimes B \rightarrow B \otimes A}$, called a *swap process* and diagrammatically depicted by



such that



(i.e., $\times_{A', B} \circ (f \otimes \text{id}_B) \circ \times_{B, A} = \text{id}_B \otimes f$) holds for any systems A, A' , and B and any process $f \in \mathbf{Proc}_{A \rightarrow A'}$. Also, assume that $\times_{A, I} = \text{id}_A = \times_{I, A}$, i.e.,



and $\times_{A, B \otimes C} = (\text{id}_B \otimes \times_{A, C}) \circ (\times_{A, B} \otimes \text{id}_C)$ hold for any systems A, B , and C .

Let us consider a trivial measurement of system A , which always gives the same outcome regardless of the input state. The effect representing an event associated to the outcome of a trivial measurement is called a *discarding effect* and denoted by $(\bar{\tau}|_A)$, or simply $(\bar{\tau}|)$, which is depicted as



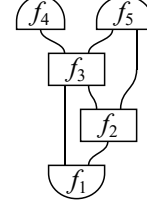
$(\bar{\tau}|_A)$ has a natural operational intuition: one performs any measurement on system A and then discards the results.

Example of quantum theory The identity process is the identity map. The discarding effect of A is the identity matrix of order N_A , denoted by $\mathbb{1}_{N_A}$.

D. Probabilistic behavior

Assume that $\mathbf{Scalar} = \mathbf{R}_+$ holds, i.e., each scalar is identified with a nonnegative real number. Let $\mathbf{Scalar}^F := \{p \in \mathbf{Scalar} : p \leq 1\}$. A state $|\rho\rangle \in \mathbf{St}_A$ is called *feasible* if $(\bar{\tau}|_A|\rho\rangle) \in \mathbf{Scalar}^F$ holds; in particular, $|\rho\rangle$ is called *normalized* (or *deterministic*) if $(\bar{\tau}|_A|\rho\rangle) = 1$ holds. Let \mathbf{St}_A^F and \mathbf{St}_A^N be, respectively, the sets of all feasible and normalized states of A . A process $f \in \mathbf{Proc}_{A \rightarrow B}$ is called *feasible* if $(f \otimes \text{id}_E) \circ |\sigma\rangle \in \mathbf{St}_{B \otimes E}^F$ holds for any system E and $|\sigma\rangle \in \mathbf{St}_{A \otimes E}^F$. Also, $f \in \mathbf{Proc}_{A \rightarrow B}$ is called *deterministic* if $(f \otimes \text{id}_E) \circ |\sigma\rangle \in \mathbf{St}_{B \otimes E}^N$ holds for any system E and $|\sigma\rangle \in \mathbf{St}_{A \otimes E}^N$. Let $\mathbf{Proc}_{A \rightarrow B}^F$ and $\mathbf{Proc}_{A \rightarrow B}^D$ be, respectively, the sets of all feasible and deterministic processes from A

to B . Also, let $\mathbf{Eff}_A^F := \mathbf{Proc}_{A \rightarrow I}^F$. Assume that each scalar consisting of the sequential and/or parallel compositions of k feasible processes f_1, \dots, f_k , for example, the scalar depicted by



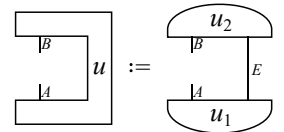
is the probability of the joint occurrence of f_1, \dots, f_5 ¹. In particular, for any $p, q \in \mathbf{Scalar}^F$, pq , i.e., the probability of the joint occurrence of p and q , is the product of real numbers p and q . Assume $(\bar{\tau}|_I) = 1$ and $(\bar{\tau}|_A) \in \mathbf{Proc}_{A \rightarrow I}^D$. 1 is the unique deterministic scalar and $\text{id}_I = 1$ holds. It follows that any process consisting of the sequential and/or parallel compositions of deterministic (resp. feasible) processes is deterministic (resp. feasible). One can easily see $\mathbf{St}_A^F = \mathbf{Proc}_{I \rightarrow A}^F$, $\mathbf{Scalar}^F = \mathbf{St}_I^F = \mathbf{Proc}_{I \rightarrow I}^F$, $\mathbf{St}_A^N = \mathbf{Proc}_{I \rightarrow A}^D$, and $\mathbf{Proc}_{A \rightarrow B}^D \subset \mathbf{Proc}_{A \rightarrow B}^F$. Any $|\rho\rangle \in \mathbf{St}_A^F$ is in the form $|\rho\rangle = p|\rho^N\rangle$ with $p \in \mathbf{Scalar}^F$ and $|\rho^N\rangle \in \mathbf{St}_A^N$ (note that $p = (\bar{\tau}|\rho\rangle)$ holds from $(\bar{\tau}|\rho^N\rangle) = 1$), which means that $|\rho\rangle$ can be identified with the process preparing the normalized state $|\rho^N\rangle$ with probability p .

From the definition, a scalar larger than 1 is unfeasible. Unfeasible scalars cannot be interpreted as probabilities and thus are not intuitive. However, it is mathematically convenient to consider unfeasible scalars, so we assume $\mathbf{Scalar} = \mathbf{R}_+$. Similarly, $\mathbf{Proc}_{A \rightarrow B}$ is defined as

$$\mathbf{Proc}_{A \rightarrow B} := \{af : a \in \mathbf{Scalar}, f \in \mathbf{Proc}_{A \rightarrow B}^F\}. \quad (3)$$

Although unfeasible processes exist in each process space $\mathbf{Proc}_{A \rightarrow B}$ (i.e., $\mathbf{Proc}_{A \rightarrow B}^F \subsetneq \mathbf{Proc}_{A \rightarrow B}$), Eq. (3) implies that any unfeasible process is expressed as scalar multiplication of a feasible process.

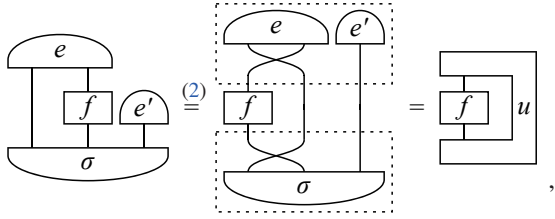
We consider the following diagram, denoted by u : $\mathbf{Proc}_{A \rightarrow B} \rightarrow \mathbf{Scalar}$, that maps a process $f \in \mathbf{Proc}_{A \rightarrow B}$ to a scalar $u(f) := (u_2| \circ (f \otimes \text{id}_E) \circ |u_1\rangle) \in \mathbf{Scalar}$, where $|u_1\rangle \in \mathbf{St}_{A \otimes E}$ and $\langle u_2| \in \mathbf{Eff}_{B \otimes E}$ hold. u can be interpreted as a set of a system E , a state $|u_1\rangle \in \mathbf{St}_{A \otimes E}$, and an effect $\langle u_2| \in \mathbf{Eff}_{B \otimes E}$, which is diagrammatically depicted as



Any scalar that includes $f \in \mathbf{Proc}_{A \rightarrow B}$ is expressed in the form $u(f)$ with some diagram $u : \mathbf{Proc}_{A \rightarrow B} \rightarrow \mathbf{Scalar}$; for

¹In OPTs, a certain set of scalars is associated with a probability distribution. See, e.g., Ref. [3] for details.

example,



where $|u_1\rangle$ and $|u_2\rangle$ are, respectively, the state and effect enclosed by the auxiliary boxes.

For two processes $f, f' \in \mathbf{Proc}_{A \rightarrow B}$, $f = f'$ is defined as

$$\boxed{f}_{|A}^{|B} = \boxed{f'}_{|A}^{|B} \Leftrightarrow \forall \boxed{u} \begin{array}{c} |B \\ \boxed{f} \\ |A \end{array} u = \boxed{f'} u = \boxed{f'} u.$$

This means that $f = f'$ holds if they are indistinguishable in a probabilistic sense. For $f, f' \in \mathbf{Proc}_{A \rightarrow B}$, $f \stackrel{\text{local}}{=} f'$ is defined as

$$\boxed{f}_{|A}^{|B} \stackrel{\text{local}}{=} \boxed{f'}_{|A}^{|B} \Leftrightarrow \forall \begin{array}{c} |A \\ \rho \end{array}, \begin{array}{c} e \\ |B \end{array}, \begin{array}{c} e \\ \rho \end{array} = \begin{array}{c} e \\ \rho \end{array}.$$

One can easily see that $f \stackrel{\text{local}}{=} f'$ holds if $f = f'$ holds, but the converse is not necessarily true. It follows that, in the case of $A = I$ or $B = I$, $f \stackrel{\text{local}}{=} f'$ and $f = f'$ are always the same, which means

$$\begin{array}{c} |A \\ \rho \end{array} = \begin{array}{c} |A \\ \rho' \end{array} \Leftrightarrow \forall \begin{array}{c} e \\ |A \end{array}, \begin{array}{c} e \\ \rho \end{array} = \begin{array}{c} e \\ \rho' \end{array}$$

and

$$\begin{array}{c} e \\ |A \end{array} = \begin{array}{c} e' \\ |A \end{array} \Leftrightarrow \forall \begin{array}{c} |A \\ \rho \end{array}, \begin{array}{c} e \\ \rho \end{array} = \begin{array}{c} e' \\ \rho \end{array}. \quad (4)$$

($\overline{\overline{\overline{|A}}}$) is the unique deterministic effect of A and that ($\overline{\overline{\overline{|A \otimes B}}}$) := ($\overline{\overline{\overline{|A}}}$) \otimes ($\overline{\overline{\overline{|B}}}$) holds, which is depicted as

$$\overline{\overline{\overline{|A \otimes B}}} = \overline{\overline{\overline{|A}}} \overline{\overline{\overline{|B}}}.$$

Example of quantum theory Since ($\overline{\overline{\overline{|A}}}$) $|\rho\rangle = \text{Tr} |\rho\rangle$ holds, $|\rho\rangle \in \mathbf{St}_A^F$ means $\text{Tr} |\rho\rangle \leq 1$. Also, $|\rho\rangle \in \mathbf{St}_A^N$ means $\text{Tr} |\rho\rangle = 1$. $f \in \mathbf{Proc}_{A \rightarrow B}^F$ holds if and only if f is a trace non-increasing CP map. Also, $f \in \mathbf{Proc}_{A \rightarrow B}^D$ means that f is a trace-preserving (TP) CP map. ($e \in \mathbf{Eff}_{A \rightarrow B}^F$) means that the maximal eigenvalue of the matrix (e) is not larger than 1. For any $f, f' \in \mathbf{Proc}_{A \rightarrow B}$, $f = f'$ and $f \stackrel{\text{local}}{=} f'$ are equivalent.

E. Process space spans vector space

Assume that, for any two feasible processes $g_1, g_2 \in \mathbf{Proc}_{A \rightarrow B}^F$ and any $p \in \mathbf{Scalar}^F$, there exists a feasible process $h \in \mathbf{Proc}_{A \rightarrow B}^F$ satisfying

$$\boxed{h} u = p \boxed{g_1} u + (1-p) \boxed{g_2} u$$

for any $u : \mathbf{Proc}_{A \rightarrow B} \rightarrow \mathbf{Scalar}$. Such a process h is denoted by $pg_1 + (1-p)g_2$. This can be interpreted as a probabilistic mixture of g_1 and g_2 with probabilities p and $1-p$. For any two processes $f_1, f_2 \in \mathbf{Proc}_{A \rightarrow B}$, the process h' satisfying

$$\boxed{h'} u = \boxed{f_1} u + \boxed{f_2} u$$

for any $u : \mathbf{Proc}_{A \rightarrow B} \rightarrow \mathbf{Scalar}$ is called the *sum* of f_1 and f_2 and denoted by $f_1 + f_2$. One can easily see that $f_1 + f_2 \in \mathbf{Proc}_{A \rightarrow B}$ holds for any $f_1, f_2 \in \mathbf{Proc}_{A \rightarrow B}$.

Let $\mathcal{I}_n := \{1, \dots, n\}$. A set of effects $\{(e_m | \in \mathbf{Eff}_A)\}_{m \in \mathcal{I}_M}$ is called a *measurement* if

$$\sum_{m=1}^M \begin{array}{c} e_m \\ \rho \end{array} = 1, \quad \forall \begin{array}{c} | \\ \rho \end{array} \in \mathbf{St}_A^N \quad (5)$$

holds. This means that the sum of probabilities over all possible outcomes is 1 whenever one performs a measurement on a normalized state. Equation (5) is equivalent to

$$\sum_{m=1}^M \begin{array}{c} e_m \\ \rho \end{array} = \overline{\overline{\overline{|}}}, \quad \forall \begin{array}{c} | \\ \rho \end{array} \in \mathbf{St}_A,$$

and thus, from Eq. (4), it follows that $\{(e_m | \in \mathbf{Eff}_A)\}_{m \in \mathcal{I}_M}$ is a measurement if and only if

$$\sum_{m=1}^M \begin{array}{c} e_m \\ | \end{array} = \overline{\overline{\overline{|}}} \quad (6)$$

holds.

We can consider the real vector space $\mathbf{V}_{A \rightarrow B}$ spanned by $\mathbf{Proc}_{A \rightarrow B}$, whose elements are formal sums of the form $\sum_i a_i f_i$ with $a_i \in \mathbf{R}$ and $f_i \in \mathbf{Proc}_{A \rightarrow B}$, where the element $\overline{f} := \sum_i a_i f_i$ satisfies

$$\boxed{\overline{f}} u = \boxed{\sum_j a_j \boxed{f_j}} u = \sum_j a_j \boxed{f_j} u \quad (7)$$

for any $u : \mathbf{Proc}_{A \rightarrow B} \rightarrow \mathbf{Scalar}$. We will call an element of $\mathbf{V}_{A \rightarrow B}$ an *extended process*, which is denoted with an overline such as \overline{f} (unless it is clearly a process). Equation (7) implies that the diagram u distributes over addition.

$\mathbf{Proc}_{A \rightarrow B} \subseteq \mathbf{V}_{A \rightarrow B}$ obviously holds. Any $\bar{f} \in \mathbf{V}_{A \rightarrow B}$ is expressed by $\bar{f} = f_+ - f_-$ with some $f_+, f_- \in \mathbf{Proc}_{A \rightarrow B}$. As well as processes, extended processes can be composed sequentially and in parallel. Specifically, for any $\bar{f} := \sum_i a_i f_i \in \mathbf{V}_{A \rightarrow B}$, $\bar{g} := \sum_j b_j g_j \in \mathbf{V}_{B \rightarrow C}$, and $\bar{h} := \sum_k c_k h_k \in \mathbf{V}_{C \rightarrow D}$ with $a_i, b_j, c_k \in \mathbf{R}$, $f_i \in \mathbf{Proc}_{A \rightarrow B}$, $g_j \in \mathbf{Proc}_{B \rightarrow C}$, and $h_k \in \mathbf{Proc}_{C \rightarrow D}$, $\bar{g} \circ \bar{f} = \sum_i \sum_j a_i b_j (g_j \circ f_i)$ and $\bar{f} \otimes \bar{h} = \sum_i \sum_k a_i c_k (f_i \otimes h_k)$ hold. This is diagrammatically depicted as

$$\begin{array}{c} \boxed{\sum_j b_j g_j} \\ \boxed{\sum_i a_i f_i} \end{array} = \sum_i \sum_j a_i b_j \begin{array}{c} \boxed{g_j} \\ \boxed{f_i} \end{array}$$

and

$$\boxed{\sum_i a_i f_i} \otimes \boxed{\sum_k c_k h_k} = \sum_i \sum_k a_i c_k \boxed{f_i} \otimes \boxed{h_k}.$$

Let $\mathbf{V}_A := \mathbf{V}_{I \rightarrow A}$ and $\mathbf{V}_A^* := \mathbf{V}_{A \rightarrow I}$; then, \mathbf{V}_A^* can be regarded as the dual vector space of \mathbf{V}_A . We can easily verify that, for any systems A and B , $\mathbf{Proc}_{A \rightarrow B}$ is a convex cone. In particular, \mathbf{St}_A is a convex cone in \mathbf{V}_A . The dimension of the real vector space \mathbf{V}_A is called the *dimension* of A . Assume that $\mathbf{Proc}_{A \rightarrow B}$ is closed. $f \in \mathbf{Proc}_{A \rightarrow B}$ is called *atomic* if $f_1 \propto f$ holds for any $f_1, f_2 \in \mathbf{Proc}_{A \rightarrow B}$ satisfying $f = f_1 + f_2$, where we denote $f_1 \propto f$ if there exists a scalar a satisfying either $f_1 = af$ or $af_1 = f$. In particular, states and effects are also called *pure* if they are atomic. Let \mathbf{St}_A^{P} be the set of all pure states of A . Also, let $\mathbf{St}_A^{\text{NP}}$ be the set of all normalized pure states, i.e., $\mathbf{St}_A^{\text{NP}} := \mathbf{St}_A^{\text{N}} \cap \mathbf{St}_A^{\text{P}}$. Any $|\rho\rangle \in \mathbf{St}_A$ can be expressed in the form

$$\boxed{\rho} = \sum_{i=1}^k p_i \boxed{\psi_i},$$

where $|\psi_1\rangle, \dots, |\psi_k\rangle \in \mathbf{St}_A^{\text{NP}}$ and $p_1, \dots, p_k \in \mathbf{R}_+$.

One can show that, for any $f \in \mathbf{Proc}_{A \rightarrow B}$, we have

$$f \in \mathbf{Proc}_{A \rightarrow B}^{\text{D}} \Leftrightarrow \boxed{f} = \boxed{\text{---}}. \quad (8)$$

Similarly to Eq. (8), an extended process $\bar{f} \in \mathbf{V}_{A \rightarrow B}$ is called *deterministic* if $\boxed{\bar{f}} \circ \boxed{\text{---}} = \boxed{\text{---}}$ holds.

An extended process $\bar{f} \in \mathbf{V}_{A \rightarrow B}$ is called *reversible* if there exists $\bar{g} \in \mathbf{V}_{B \rightarrow A}$, called an *inverse* of \bar{f} , such that $\bar{g} \circ \bar{f} \stackrel{\text{local}}{=} \text{id}_A$ and $\bar{f} \circ \bar{g} \stackrel{\text{local}}{=} \text{id}_B$. Such \bar{g} is also reversible. We denote $A \cong B$ if there exists a reversible extended process from A to B .

A process $f \in \mathbf{Proc}_{A \otimes B \rightarrow C \otimes D}$ is called *separable* if it can be expressed in the form

$$\boxed{f} = \sum_{i=1}^I \boxed{g_i} \otimes \boxed{h_i}$$

with $g_i \in \mathbf{Proc}_{A \rightarrow C}$ and $h_i \in \mathbf{Proc}_{B \rightarrow D}$.

Example of quantum theory The sum of processes is equal to the sum of CP maps. In particular, the sum of states (or effects) is the sum of matrices. $\Pi := \{(e_m) \in \mathbf{Eff}_A\}_{m \in \mathcal{I}_M}$ is a measurement if and only if $\sum_{m=1}^M (e_m) = \mathbb{1}_{N_A}$ holds, i.e., Π is a positive operator-valued measure (POVM) (note that each effect (e_m) is a positive semidefinite matrix). In fully quantum theory, $\mathbf{V}_{A \rightarrow B}$ is isomorphic to the space of all linear maps from $\mathcal{S}(\mathbf{C}^{N_A})$ to $\mathcal{S}(\mathbf{C}^{N_B})$ (which are also called Hermitian-preserving maps); in particular, \mathbf{V}_A and \mathbf{V}_A^* are isomorphic to $\mathcal{S}(\mathbf{C}^{N_A})$. $|\psi\rangle \in \mathbf{St}_A$ is pure if and only if $|\psi\rangle = |\psi\rangle \langle \psi|$ holds for some vector $|\psi\rangle$.

III. OPTs WITH CLASSICAL SYSTEMS

A. Classical systems

We will call an M -dimensional system C *classical* if there exist M normalized pure states $|1\rangle, \dots, |M\rangle \in \mathbf{St}_C^{\text{NP}}$ and a measurement $\{(m) \in \mathbf{Eff}_C\}_{m \in \mathcal{I}_M}$ satisfying

$$\begin{array}{c} \boxed{m} \\ \vdots \\ \boxed{n} \end{array} = \delta_{m,n} \quad (9)$$

and

$$\sum_{m=1}^M \begin{array}{c} \boxed{m} \\ \boxed{m} \end{array} = \boxed{\text{---}}, \quad (10)$$

where $\delta_{m,n}$ is the Kronecker delta. A classical system is depicted as the dotted line. One can easily see that any $|\rho\rangle \in \mathbf{St}_C$ is expressed in the form

$$\boxed{\rho} \stackrel{(10)}{=} \sum_{m=1}^M \begin{array}{c} \boxed{m} \\ \boxed{m} \end{array} = \sum_{m=1}^M p_m \boxed{m}, \quad (11)$$

where $p_m := (m|\rho) \in \mathbf{Scalar}$. This immediately gives that $\mathbf{St}_C^{\text{NP}} = \{|m\rangle\}_{m \in \mathcal{I}_M}$ holds. Indeed, from Eq. (11), $|\rho\rangle \in \mathbf{St}_C^{\text{NP}}$ holds if and only if $p_m = \delta_{i,m}$ holds for some $i \in \mathcal{I}_M$. In what follows, we consider an OPT that has an M -dimensional classical system C . Note that a classical system is not intrinsically necessary for investigating a state discrimination problem in an OPT, but it helps us to express this problem in straightforward diagrammatic terms.

It follows from Eq. (10) that any state of $C \otimes A$ (where A is an arbitrary system that is not classical in general) is separable. Indeed, one can easily obtain

$$\begin{array}{c} \boxed{\sigma} \\ \boxed{\sigma} \end{array} \stackrel{(10)}{=} \sum_{m=1}^M \begin{array}{c} \boxed{m} \\ \boxed{m} \end{array} \otimes \boxed{\rho'_m}, \quad (12)$$

where the state enclosed by the auxiliary box is denoted by $|\rho'_m\rangle$. Similarly, any state of $A \otimes C$, effect of $C \otimes A$, and effect of $A \otimes C$ are separable. Moreover, for any $f \in \mathbf{Proc}_{C \rightarrow A}$ with a classical system C , $|\sigma\rangle \in \mathbf{St}_{C \otimes E}$, and $(e| \in \mathbf{Eff}_{A \otimes E}$, Eq. (12) yields

$$\begin{aligned}
 \begin{array}{c} \text{---} \\ | \\ \text{---} \\ e \\ | \\ \text{---} \\ f \\ | \\ \text{---} \\ C \\ | \\ \text{---} \\ \sigma \end{array} &= \sum_{m=1}^M \begin{array}{c} \text{---} \\ | \\ \text{---} \\ e \\ | \\ \text{---} \\ f \\ | \\ \text{---} \\ m \end{array} \begin{array}{c} \text{---} \\ | \\ \text{---} \\ \rho'_m \end{array} \\
 &= \sum_{m=1}^M \begin{array}{c} \text{---} \\ | \\ \text{---} \\ e \\ | \\ \text{---} \\ f \\ | \\ \text{---} \\ m \end{array} \begin{array}{c} \text{---} \\ | \\ \text{---} \\ e_m \\ | \\ \text{---} \\ f \\ | \\ \text{---} \\ m \end{array} \\
 &= \sum_{m=1}^M \begin{array}{c} \text{---} \\ | \\ \text{---} \\ e_m \\ | \\ \text{---} \\ f \\ | \\ \text{---} \\ m \end{array}, \quad (13)
 \end{aligned}$$

where the effect enclosed by the auxiliary box is denoted by $(e_m|$. One can see from Eq. (13) that, for any $f, f' \in \mathbf{Proc}_{C \rightarrow A}$,

$$f = f' \Leftrightarrow f \stackrel{\text{local}}{=} f' \quad (14)$$

holds. Similarly, Eq. (14) holds for any $f, f' \in \mathbf{Proc}_{A \rightarrow C}$ with a classical system C .

B. $|\cup\rangle$ and $(\cap|$

We will introduce the state $|\cup\rangle$ and the effect $(\cap|$ defined as

$$\begin{array}{c} \text{---} \\ | \\ \text{---} \\ \cup \\ | \\ \text{---} \end{array} := \sum_{l=1}^M \begin{array}{c} \text{---} \\ | \\ \text{---} \\ l \\ | \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ | \\ \text{---} \\ l \\ | \\ \text{---} \end{array}, \quad \begin{array}{c} \text{---} \\ | \\ \text{---} \\ \cap \\ | \\ \text{---} \end{array} := \sum_{l=1}^M \begin{array}{c} \text{---} \\ | \\ \text{---} \\ l \\ | \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ | \\ \text{---} \\ l \\ | \\ \text{---} \end{array}. \quad (15)$$

Clearly, we have

$$\begin{array}{c} \text{---} \\ | \\ \text{---} \\ \cup \\ | \\ \text{---} \\ n \end{array} \stackrel{(15)}{=} \sum_{l=1}^M \begin{array}{c} \text{---} \\ | \\ \text{---} \\ l \\ | \\ \text{---} \\ n \end{array} \begin{array}{c} \text{---} \\ | \\ \text{---} \\ l \\ | \\ \text{---} \\ n \end{array} \stackrel{(9)}{=} \sum_{l=1}^M \begin{array}{c} \text{---} \\ | \\ \text{---} \\ n \\ | \\ \text{---} \\ n \end{array} \stackrel{(15)}{=} \begin{array}{c} \text{---} \\ | \\ \text{---} \\ \cup \\ | \\ \text{---} \\ n \end{array}, \quad (16)$$

$$\begin{array}{c} \text{---} \\ | \\ \text{---} \\ \cap \\ | \\ \text{---} \\ n \end{array} \stackrel{(15)}{=} \sum_{l=1}^M \begin{array}{c} \text{---} \\ | \\ \text{---} \\ n \\ | \\ \text{---} \\ l \end{array} \begin{array}{c} \text{---} \\ | \\ \text{---} \\ l \\ | \\ \text{---} \\ n \end{array} \stackrel{(9)}{=} \sum_{l=1}^M \begin{array}{c} \text{---} \\ | \\ \text{---} \\ n \\ | \\ \text{---} \\ n \end{array} \stackrel{(15)}{=} \begin{array}{c} \text{---} \\ | \\ \text{---} \\ \cap \\ | \\ \text{---} \\ n \end{array}, \quad (17)$$

and

$$\begin{array}{c} \text{---} \\ | \\ \text{---} \\ \cup \\ | \\ \text{---} \\ m \end{array} \begin{array}{c} \text{---} \\ | \\ \text{---} \\ \cap \\ | \\ \text{---} \\ n \end{array} \stackrel{(16)}{=} \begin{array}{c} \text{---} \\ | \\ \text{---} \\ n \\ | \\ \text{---} \\ m \end{array} \stackrel{(9)}{=} \delta_{m,n}.$$

One can also show

$$\begin{array}{c} \text{---} \\ | \\ \text{---} \\ \cup \\ | \\ \text{---} \\ \cap \\ | \\ \text{---} \end{array} = \begin{array}{c} \text{---} \\ | \\ \text{---} \\ \cap \\ | \\ \text{---} \\ \cup \\ | \\ \text{---} \end{array}. \quad (18)$$

Intuitively, this means that a curved line consisting of ‘ \cup ’ and ‘ \cap ’ can be yanked. The first equality of Eq. (18) is obtained from

$$\begin{array}{c} \text{---} \\ | \\ \text{---} \\ \cup \\ | \\ \text{---} \\ \cap \\ | \\ \text{---} \end{array} \stackrel{(15)}{=} \sum_{l=1}^M \begin{array}{c} \text{---} \\ | \\ \text{---} \\ l \\ | \\ \text{---} \\ l \\ | \\ \text{---} \end{array} \stackrel{(9)}{=} \sum_{m=1}^M \begin{array}{c} \text{---} \\ | \\ \text{---} \\ m \\ | \\ \text{---} \\ m \\ | \\ \text{---} \end{array} \stackrel{(10)}{=} \begin{array}{c} \text{---} \\ | \\ \text{---} \\ \cap \\ | \\ \text{---} \\ \cup \\ | \\ \text{---} \end{array}.$$

The same is true for the second equality of Eq. (18). $|\underline{\cup}\rangle \in \mathbf{St}_C$ is defined as

$$|\underline{\cup}\rangle := \sum_{m=1}^M \begin{array}{c} \text{---} \\ | \\ \text{---} \\ m \\ | \\ \text{---} \end{array}; \quad (19)$$

then, we have

$$\begin{array}{c} \text{---} \\ | \\ \text{---} \\ \underline{\cup} \\ | \\ \text{---} \end{array} \stackrel{(6)}{=} \begin{array}{c} \text{---} \\ | \\ \text{---} \\ \underline{\cup} \\ | \\ \text{---} \end{array} \stackrel{(17)}{=} \begin{array}{c} \text{---} \\ | \\ \text{---} \\ \underline{\cup} \\ | \\ \text{---} \end{array}. \quad (20)$$

C. Expression of measurements

In an OPT with a classical system C , any measurement $\{(e_m|)_{m \in \mathcal{I}_M}\}$ can be expressed by the process $e := \sum_{n=1}^M |n\rangle(e_n| \in \mathbf{Proc}_{A \rightarrow C}$. Diagrammatically,

$$\begin{array}{c} \text{---} \\ | \\ \text{---} \\ e \\ | \\ \text{---} \end{array} := \sum_{n=1}^M \begin{array}{c} \text{---} \\ | \\ \text{---} \\ n \\ | \\ \text{---} \\ e_n \\ | \\ \text{---} \end{array}. \quad (21)$$

Each effect $(e_n|$ with $n \in \mathcal{I}_M$ is obtained from

$$\begin{array}{c} \text{---} \\ | \\ \text{---} \\ e_n \\ | \\ \text{---} \end{array} \stackrel{(9)}{=} \begin{array}{c} \text{---} \\ | \\ \text{---} \\ n \\ | \\ \text{---} \\ e \\ | \\ \text{---} \end{array}.$$

One can easily verify that $(\bar{\cap}_C| \circ e = (\bar{\cap}_A|$, i.e., e is deterministic from Eq. (8). Conversely, any deterministic process from A to C can be interpreted as a measurement since $e' \in \mathbf{Proc}_{A \rightarrow C}^D$ can be depicted as

$$\begin{array}{c} \text{---} \\ | \\ \text{---} \\ e' \\ | \\ \text{---} \end{array} \stackrel{(10)}{=} \sum_{n=1}^M \begin{array}{c} \text{---} \\ | \\ \text{---} \\ n \\ | \\ \text{---} \\ n \\ | \\ \text{---} \\ e' \\ | \\ \text{---} \end{array} = \sum_{n=1}^M \begin{array}{c} \text{---} \\ | \\ \text{---} \\ n \\ | \\ \text{---} \\ e'_n \\ | \\ \text{---} \end{array},$$

where $(e'_n| := (n| \circ e'$. For each system A , $\mathbf{Proc}_{A \rightarrow C}^D$ is denoted by \mathbf{Meas}_A^C or simply \mathbf{Meas}_A .

D. State preparation

We assume that one of M normalized states of system A , $|\rho_1^N\rangle, \dots, |\rho_M^N\rangle \in \mathbf{St}_A^N$, is prepared with prior probabilities ξ_1, \dots, ξ_M with $\sum_{m=1}^M \xi_m = 1$. Let $|\rho_m\rangle := \xi_m |\rho_m^N\rangle$. We consider the process $\rho \in \mathbf{Proc}_{C \rightarrow A}$ defined as

$$\begin{array}{c} \text{---} \\ | \\ \text{---} \\ \rho \\ | \\ \text{---} \end{array} := \sum_{n=1}^M \begin{array}{c} \text{---} \\ | \\ \text{---} \\ \rho_n \\ | \\ \text{---} \\ n \\ | \\ \text{---} \end{array} = \sum_{n=1}^M \xi_n \begin{array}{c} \text{---} \\ | \\ \text{---} \\ \rho_n^N \\ | \\ \text{---} \\ n \\ | \\ \text{---} \end{array}. \quad (22)$$

Each $|\rho_n\rangle$ is obtained from

$$\begin{array}{c} \text{---} \\ | \\ \rho_n \end{array} \stackrel{(9)}{=} \begin{array}{c} \text{---} \\ | \\ \rho \\ \text{---} \\ | \\ n \end{array}.$$

Note that ρ is not deterministic unless $M = 1$. We have

$$\begin{array}{c} \text{---} \\ | \\ \rho \\ \text{---} \\ | \\ \text{---} \\ \text{---} \end{array} \stackrel{(22)}{=} \sum_{m=1}^M \begin{array}{c} \text{---} \\ | \\ \rho_m \\ \text{---} \\ | \\ \text{---} \\ \text{---} \end{array} \stackrel{(19)}{=} \sum_{m=1}^M \begin{array}{c} \text{---} \\ | \\ \rho_m \end{array}.$$

$\sum_{m=1}^M |\rho_m\rangle$ is obviously normalized. We will call a process $\rho \in \mathbf{Proc}_{C \rightarrow A}$ with $\rho \circ |\underline{\text{---}}\rangle \in \mathbf{St}_C^N$ a *state preparation*. Let \mathbf{Prep}_A be the set of all state preparations.

IV. MINIMUM-ERROR MEASUREMENT FOR GROUP COVARIANT STATES

A. Discrimination problem

Let us review the problem of discriminating a given set of M known normalized states with given prior probabilities [33]–[35]. Here, we consider the following scenario: One party (Charlie) randomly chooses one of the M states $|\rho_1^N\rangle, \dots, |\rho_M^N\rangle \in \mathbf{St}_A^N$ with prior probabilities ξ_1, \dots, ξ_M . Such a process is expressed by the state preparation $\rho \in \mathbf{Prep}_A$ depicted by Eq. (22). Since she knows which state she has, we can interpret that she has the following state

$$\sum_{m=1}^M \begin{array}{c} \text{---} \\ | \\ \rho \\ \text{---} \\ | \\ m \end{array} \stackrel{(15)}{=} \begin{array}{c} \text{---} \\ | \\ \rho \\ \text{---} \\ | \\ \text{---} \\ \text{---} \end{array}.$$

Indeed, by performing a measurement $\{(m)\}_{m \in I_M}$ on the classical system C , she can always determine which state she has. Charlie sends the state to the other party (Alice). Alice knows the possible states $|\rho_1^N\rangle, \dots, |\rho_M^N\rangle$ and their prior probabilities but does not know which state Charlie sent. We can interpret that Alice gets the following normalized state

$$\begin{array}{c} \text{---} \\ | \\ \rho \\ \text{---} \\ | \\ \text{---} \\ \text{---} \end{array} \stackrel{(20)}{=} \begin{array}{c} \text{---} \\ | \\ \rho \\ \text{---} \\ | \\ \text{---} \\ \text{---} \end{array}.$$

What Alice has to do is to perform a measurement that will correctly discriminate between the states $|\rho_1^N\rangle, \dots, |\rho_M^N\rangle$ with high probability. Alice performs a measurement $e := \sum_{m=1}^M |m\rangle \circ (e_m) \in \mathbf{Meas}_A$ with $(e_m) \in \mathbf{Eff}_A$ to discriminate between the states as accurately as possible. After that, Charlie and Alice check whether Alice correctly determines the state. In this paper, we apply the strategy that maximizes the average success

probability. This probability is given as a function of the measurement e , denoted by $P(e)$, which is expressed by

$$P(e) := \begin{array}{c} \text{Charlie} \quad \text{Alice} \\ \text{---} \quad \text{---} \\ | \quad | \\ \rho \\ \text{---} \\ | \\ \text{---} \\ \text{---} \end{array} \stackrel{(23)}{=} \begin{array}{c} \text{---} \\ | \\ e \\ \text{---} \\ | \\ \rho \\ \text{---} \\ | \\ \text{---} \\ \text{---} \end{array}.$$

Note that $P(e)$ is also expressed by

$$P(e) = \sum_{m=1}^M (e_m | \rho_m) = \sum_{m=1}^M (m | \circ e \circ \rho \circ |m).$$

A measurement that maximizes the average success probability is called a *minimum-error measurement*. The problem of finding a minimum-error measurement is formulated as

$$\begin{array}{ll} \text{maximize} & P(e) \\ \text{subject to} & e \in \mathbf{Meas}_A \end{array} \quad (24)$$

with variable e . If we want to optimize over a restricted class of measurements, denoted by \mathcal{M}_A with $\mathcal{M}_A \subseteq \mathbf{Meas}_A$, then we consider the following problem:

$$\begin{array}{ll} \text{maximize} & P(e) \\ \text{subject to} & e \in \mathcal{M}_A. \end{array} \quad (25)$$

B. Group action

We use group theory to represent the symmetric properties of a state preparation. Let \mathcal{G} be a group and $1 \in \mathcal{G}$ be its identity element. Let $\text{Aut}(\mathcal{Z})$ be the group of all automorphisms of an object \mathcal{Z} . A map $\psi : \mathcal{G} \ni g \mapsto \psi_g \in \text{Aut}(\mathcal{Z})$ is called a (right) *group action* of \mathcal{G} on \mathcal{Z} if $\psi_{gh}(z) = \psi_h[\psi_g(z)]$ ($\forall g, h \in \mathcal{G}$) and $\psi_1(z) = z$ hold for any $z \in \mathcal{Z}$. In the following, we give two examples of group actions.

The first example is a group action, $\tau : \mathcal{G} \rightarrow \text{Aut}(I_M)$, of \mathcal{G} on I_M . $\text{Aut}(I_M)$ is the group of all permutations of I_M . We will identify I_M with $\{(m) \in \mathbf{Eff}_C\}_{m \in I_M}$; then, we have

$$\begin{array}{c} \text{---} \\ | \\ \tau_1 \\ \text{---} \\ | \\ \text{---} \\ \text{---} \end{array} = \begin{array}{c} \text{---} \\ | \\ \text{---} \\ \text{---} \\ | \\ \text{---} \\ \text{---} \end{array}.$$

Also, we have that, for any $g, h \in \mathcal{G}$,

$$\begin{array}{c} \text{---} \\ | \\ m \\ \text{---} \\ | \\ \tau_g \\ \text{---} \\ | \\ \tau_h \\ \text{---} \\ | \\ \text{---} \\ \text{---} \end{array} = \begin{array}{c} \text{---} \\ | \\ m \\ \text{---} \\ | \\ \tau_{gh} \\ \text{---} \\ | \\ \text{---} \\ \text{---} \end{array},$$

which means $\tau_g \circ \tau_h \stackrel{\text{local}}{=} \tau_{gh}$. Thus, from Eq. (14) with $A = C$, we have

$$\begin{array}{c} \vdots \\ \boxed{\tau_g} \\ \vdots \\ \boxed{\tau_h} \\ \vdots \end{array} = \begin{array}{c} \vdots \\ \boxed{\tau_{gh}} \\ \vdots \end{array} \quad (26)$$

Clearly, for any $g \in \mathcal{G}$, τ_g is deterministic and $\tau_{g^{-1}}$ is the inverse of τ_g . Note that we also have²

$$\forall \begin{array}{c} \vdots \\ \textcircled{m} \\ \vdots \end{array}, \begin{array}{c} \vdots \\ \boxed{\tau_g} \\ \vdots \\ \boxed{\tau_h} \\ \vdots \\ \textcircled{m} \end{array} = \begin{array}{c} \vdots \\ \boxed{\tau_{gh}} \\ \vdots \\ \textcircled{m} \end{array} \quad (27)$$

We can easily verify

$$\begin{array}{c} \vdots \\ \boxed{\tau_g} \\ \vdots \end{array} = \begin{array}{c} \vdots \\ \boxed{\tau_{g^{-1}}} \\ \vdots \end{array} \quad (28)$$

Indeed, from Eq. (15), we have that, for any $m, n \in \mathcal{I}_M$,

$$\begin{array}{c} \vdots \\ \boxed{\tau_g} \\ \vdots \\ \textcircled{m} \quad \textcircled{n} \end{array} = \delta_{\tau_g(m), n} = \delta_{m, \tau_{g^{-1}}(n)} = \begin{array}{c} \vdots \\ \boxed{\tau_{g^{-1}}} \\ \vdots \\ \textcircled{m} \quad \textcircled{n} \end{array} \quad (29)$$

Since any state in $C \otimes C$ is in the form $\sum_{m=1}^M \sum_{n=1}^M c_{m,n} |m\rangle \otimes |n\rangle$ with $c_{m,n} \in \mathbf{R}_+$, Eq. (29) yields Eq. (28). Similarly, we have

$$\begin{array}{c} \vdots \\ \boxed{\tau_g} \\ \vdots \end{array} = \begin{array}{c} \vdots \\ \boxed{\tau_{g^{-1}}} \\ \vdots \end{array} \quad (30)$$

The second example is a group action, $\bar{\pi} : \mathcal{G} \rightarrow \text{Aut}(\mathcal{M}_A)$, of \mathcal{G} on \mathcal{M}_A , where \mathcal{M}_A is a convex subset of \mathbf{Meas}_A . Each automorphism of \mathcal{M}_A is in $\mathbf{V}_{A \rightarrow A}$, i.e., each element of $\text{Aut}(\mathcal{M}_A)$ is linear. We will say that $\mathcal{M}_A \subseteq \mathbf{Meas}_A$ is *symmetric under permutations of the measurement outcomes* if, for any $h \in \text{Aut}(\mathcal{I}_M)$ and $e \in \mathcal{M}_A$, $h \circ e \in \mathcal{M}_A$ holds. This means that a process that first performs the measurement $e \in \mathcal{M}_A$ and then makes permutations among the measurement results is also in \mathcal{M}_A . $\bar{\pi}$ satisfies

$$\boxed{\bar{\pi}_1} \stackrel{\text{local}}{=} \text{---}$$

²We can also identify \mathcal{I}_M with $\mathbf{St}_C^{\text{NP}} = \{|m\rangle\}_{m \in \mathcal{I}_M}$; in this case, Eq. (27) implies that τ can be interpreted as a left group action of \mathcal{G} on \mathcal{I}_M . A map $\psi : \mathcal{G} \ni g \mapsto \psi_g \in \text{Aut}(\mathcal{Z})$ is called a *left group action* of \mathcal{G} on \mathcal{Z} if $\psi_{gh}(z) = \psi_g[\psi_h(z)]$ ($\forall g, h \in \mathcal{G}$) and $\psi_1(z) = z$ hold for any $z \in \mathcal{Z}$.

and

$$\forall \begin{array}{c} \vdots \\ \textcircled{e} \\ \vdots \end{array} \in \mathcal{M}_A, \begin{array}{c} \vdots \\ \textcircled{e} \\ \vdots \\ \boxed{\bar{\pi}_g} \\ \vdots \\ \boxed{\bar{\pi}_h} \\ \vdots \end{array} = \begin{array}{c} \vdots \\ \textcircled{e} \\ \vdots \\ \boxed{\bar{\pi}_{gh}} \\ \vdots \end{array} \quad (31)$$

for any $g, h \in \mathcal{G}$. Clearly, $\bar{\pi}_{g^{-1}}$ is the inverse of $\bar{\pi}_g$. Each $\bar{f} \in \text{Aut}(\mathcal{M}_A)$ is reversible.

An example of \mathcal{M}_A is \mathbf{Meas}_A itself; \mathbf{Meas}_A is obviously symmetric under permutations of the measurement outcomes. It is also easily seen that the followings are equivalent:

- 1) $\bar{f} \in \text{Aut}(\mathbf{Meas}_A)$.
- 2) $\bar{f} \in \mathbf{V}_{A \rightarrow A}$ is reversible and satisfies $e \circ \bar{f} \in \mathbf{Meas}_A$ for any $e \in \mathbf{Meas}_A$.
- 3) $\bar{f} \in \mathbf{V}_{A \rightarrow A}$ is reversible, deterministic, and positive for effects, where we will call $\bar{g} \in \mathbf{V}_{A \rightarrow B}$ *positive for effects* if $|e| \circ \bar{g} \in \mathbf{Eff}_A$ holds for any $|e| \in \mathbf{Eff}_B$.

Other examples of \mathcal{M}_A will be shown in Subsec. IV-D. Note that $\text{Aut}(\mathcal{M}_A) \subseteq \text{Aut}(\mathbf{Meas}_A)$ does not hold in general; an example will be given in the example of quantum theory stated in Subsec. IV-D.

C. Symmetric properties

We consider a set $(\mathcal{G}, \tau, \bar{\pi})$, where $\tau : \mathcal{G} \rightarrow \text{Aut}(\mathcal{I}_M)$ and $\bar{\pi} : \mathcal{G} \rightarrow \text{Aut}(\mathcal{M}_A)$ are group actions with a group \mathcal{G} . We will say that a state preparation $\rho := \sum_{n=1}^M |\rho_n\rangle \circ |n\rangle \in \mathbf{Prep}_A$ is $(\mathcal{G}, \tau, \bar{\pi})$ -*covariant* (or simply \mathcal{G} -*covariant*) if

$$\begin{array}{c} \vdots \\ \boxed{\tau_g} \\ \vdots \\ \boxed{\bar{\pi}_g} \\ \vdots \\ \rho \end{array} = \begin{array}{c} \vdots \\ \rho \end{array} \quad (32)$$

holds for any $g \in \mathcal{G}$. We will also say that a measurement $e^\diamond := \sum_{n=1}^M |n\rangle \circ (e_n^\diamond) \in \mathcal{M}_A$ is $(\mathcal{G}, \tau, \bar{\pi})$ -*covariant* if

$$\begin{array}{c} \vdots \\ \textcircled{e^\diamond} \\ \vdots \\ \boxed{\tau_g} \\ \vdots \\ \boxed{\bar{\pi}_g} \\ \vdots \end{array} = \begin{array}{c} \vdots \\ \textcircled{e^\diamond} \\ \vdots \end{array} \quad (33)$$

holds for any $g \in \mathcal{G}$. Equation (32) is the same as

$$\begin{array}{c} \vdots \\ \boxed{\bar{\pi}_g} \\ \vdots \\ \rho_m \end{array} = \begin{array}{c} \vdots \\ \rho_{\tau_g(m)} \end{array}$$

since

and

Note that Eq. (32) is also the same as

Similar results hold for e ; for example, Eq. (33) is the same as

since

and

Theorem 1: Let \mathcal{M}_A be a convex subset of \mathbf{Meas}_A that is symmetric under permutations of the measurement outcomes. Let $\tau : \mathcal{G} \rightarrow \text{Aut}(\mathcal{I}_M)$ and $\bar{\pi} : \mathcal{G} \rightarrow \text{Aut}(\mathcal{M}_A)$ be group actions with a group \mathcal{G} . If a state preparation $\rho \in \mathbf{Prep}_A$ is $(\mathcal{G}, \tau, \bar{\pi})$ -covariant, then, for any measurement $e \in \mathcal{M}_A$, there exists a $(\mathcal{G}, \tau, \bar{\pi})$ -covariant measurement $e^\diamond \in \mathcal{M}_A$ satisfying $P(e^\diamond) = P(e)$.

Proof. Let

From the definition of \mathcal{M}_A , we can easily verify $e^\diamond \in \mathcal{M}_A$. Equation (34) yields

and thus

holds for any $g \in \mathcal{G}$, where the last equality follows from $\{h : h \in \mathcal{G}\} = \mathcal{G} = \{hg : h \in \mathcal{G}\}$. Thus, e^\diamond is $(\mathcal{G}, \tau, \bar{\pi})$ -covariant. Also, we obtain

which gives $P(e^\diamond) = P(e)$. \square

Theorem 1 immediately yields the following corollary.

Corollary 2: Let \mathcal{M}_A be a convex subset of \mathbf{Meas}_A that is symmetric under permutations of the measurement outcomes. Let $\tau : \mathcal{G} \rightarrow \text{Aut}(\mathcal{I}_M)$ and $\bar{\pi} : \mathcal{G} \rightarrow \text{Aut}(\mathcal{M}_A)$ be group actions with a group \mathcal{G} . If a state preparation $\rho \in \mathbf{Prep}_A$ is $(\mathcal{G}, \tau, \bar{\pi})$ -covariant, then there exists a $(\mathcal{G}, \tau, \bar{\pi})$ -covariant measurement that is optimal for Problem (25).

Example of quantum theory We consider the special case of $\mathcal{M}_A = \mathbf{Meas}_A$. Any $\bar{f} \in \text{Aut}(\mathbf{Meas}_A)$ is expressed in the form

$$\bar{f} \circ |\rho\rangle = U_{\bar{f}} \cdot |\rho\rangle \cdot U_{\bar{f}}^\dagger, \quad (36)$$

where $U_{\bar{f}}$ is a unitary or anti-unitary matrix of order N_A and \dagger denotes the conjugate transpose³. Thus, $\bar{\pi}_g \in \text{Aut}(\mathbf{Meas}_A)$ must be in the form

$$\bar{\pi}_g \circ |\rho_m\rangle = U_g \cdot |\rho_m\rangle \cdot U_g^\dagger,$$

where U_g is a unitary or anti-unitary matrix of order N_A . Since $\bar{\pi} : \mathcal{G} \rightarrow \text{Aut}(\mathbf{Meas}_A)$ is a group action, $U_1 \cdot |\rho_m\rangle \cdot U_1^\dagger = \mathbb{1}_{N_A}$ and $U_g[U_h \cdot |\rho_m\rangle \cdot U_h^\dagger]U_g^\dagger = U_{gh} \cdot |\rho_m\rangle \cdot U_{gh}^\dagger$ must hold for any $g, h \in \mathcal{G}$. This type of symmetry has been discussed in Ref. [25].

D. Optimization over a restricted class of measurements

Theorem 1 can be utilized to optimize over a restricted class of measurements, as we will see in this subsection. We here consider state discrimination problems in a bipartite system.

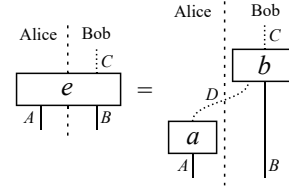
Let us introduce a three-party: Alice, Bob, and Charlie. We will consider the following scenario. Charlie randomly chooses one of the M states $\rho_1^N, \dots, \rho_M^N \in \mathbf{St}_{A \otimes B}^N$ with prior probabilities ξ_1, \dots, ξ_M , which is expressed by the state preparation $\rho := \sum_{n=1}^M |\rho_n\rangle \circ (n) \in \mathbf{Prep}_{A \otimes B}$ with $|\rho_n\rangle := \xi_n |\rho_n^N\rangle$ and $(n) \in \mathbf{Eff}_C$. A and B respectively refer to the systems of Alice and Bob. Also, C is a classical system with $N_C = M$. Charlie sends the state to Alice and Bob, who perform a measurement $e := \sum_{m=1}^M |m\rangle \circ (e_m) \in \mathbf{Meas}_{A \otimes B}$ with $(e_m) \in \mathbf{Eff}_{A \otimes B}$. Then, Alice, Bob, and Charlie check whether Alice and Bob correctly determine the state. Similarly to Eq. (23), the average success probability of a measurement e is depicted as

$$P(e) := \begin{array}{c} \text{Charlie} \quad \text{Alice} \quad \text{Bob} \\ \vdots \quad \vdots \quad \vdots \\ C \quad A \quad B \\ \rho \quad e \\ \vdots \quad \vdots \quad \vdots \\ C \quad A \quad B \end{array} = \begin{array}{c} \rho \\ e \end{array}$$

If Alice and Bob can perform any measurement, this problem is expressed by Eq. (24) with $A \otimes B$ instead of A . In this subsection, assume that they can only perform restricted measurements. We consider four classes of measurements: sequential, LOCC, separable, and PT. Let D and D' be classical systems, which can be infinite-dimensional.

³Easy proof: One can easily verify that any $\bar{f} \in \text{Aut}(\mathbf{Meas}_A)$ maps pure effects to pure effects and thus maps normalized pure states to normalized pure states. Therefore, according to Lemma 4 of Ref. [36], \bar{f} is expressed in the form of Eq. (36) or in the form $\bar{f} \circ |\rho\rangle = [\text{Tr} |\rho\rangle] \cdot |\phi\rangle$ with a fixed $|\phi\rangle \in \mathbf{St}_A^{\text{NP}}$. The latter case is ruled out since \bar{f} is reversible.

A measurement $e \in \mathbf{Meas}_{A \otimes B}$ is referred to as *sequential* if it can be expressed in the form



with $a \in \mathbf{Meas}_A^D$ and $b \in \mathbf{Meas}_{D \otimes B}$. $b \in \mathbf{Meas}_{D \otimes B}$ can be interpreted as a classical controlled measurement. Indeed, let us define b^i as a measurement of B depicted by

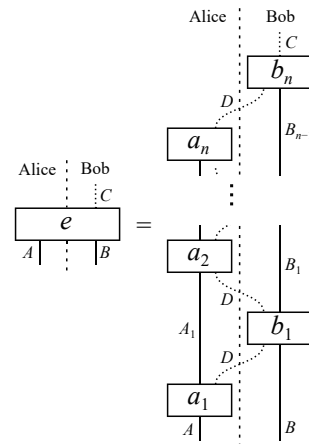
$$b^i := \begin{array}{c} \vdots \\ b \\ \vdots \end{array} \quad ; \quad \begin{array}{c} \vdots \\ i \\ \vdots \end{array} \quad ; \quad \begin{array}{c} \vdots \\ i \\ \vdots \end{array} \quad ; \quad \begin{array}{c} \vdots \\ b^i \\ \vdots \end{array} \quad ; \quad (37)$$

then, we have

$$\begin{array}{c} \vdots \\ b \\ \vdots \end{array} \stackrel{(10)}{=} \sum_i \begin{array}{c} \vdots \\ i \\ \vdots \end{array} \begin{array}{c} \vdots \\ b^i \\ \vdots \end{array} \stackrel{(37)}{=} \sum_i \begin{array}{c} \vdots \\ i \\ \vdots \end{array} \begin{array}{c} \vdots \\ b^i \\ \vdots \end{array}$$

b can be interpreted as a process which performs a measurement $b^i \in \mathbf{Meas}_B$ when the state $|i\rangle$ is inputted to the system D .

A measurement $e \in \mathbf{Meas}_{A \otimes B}$ is referred to as *LOCC* if it can be expressed in the form



with $a_1 \in \mathbf{Proc}_{A \rightarrow A_1 \otimes D}^D$, $a_k \in \mathbf{Proc}_{A_{k-1} \otimes D \rightarrow A_k \otimes D}^D$ ($k \in \{2, \dots, n-1\}$), $a_n \in \mathbf{Meas}_{A_{n-1} \otimes D}^D$, $b_1 \in \mathbf{Proc}_{D \otimes B \rightarrow D \otimes B_1}^D$, $b_l \in \mathbf{Proc}_{D \otimes B_{l-1} \rightarrow D \otimes B_l}^D$ ($l \in \{2, \dots, n-1\}$), and $b_n \in \mathbf{Meas}_{D \otimes B_{n-1}}$, where n is some natural number.

A measurement $e \in \mathbf{Meas}_{A \otimes B}$ is referred to as *separable* if it can be expressed in the form

$$\begin{array}{c} \text{Alice} \quad \text{Bob} \\ \vdots \quad \vdots \\ \text{C} \\ \boxed{e} \\ \vdots \quad \vdots \\ \text{A} \quad \text{B} \end{array} = \begin{array}{c} \text{Alice} \quad \text{Bob} \\ \vdots \quad \vdots \\ \text{C} \\ \boxed{c} \\ \vdots \quad \vdots \\ \text{D} \quad \text{D}' \\ \boxed{a} \quad \boxed{b} \\ \vdots \quad \vdots \\ \text{A} \quad \text{B} \end{array}, \quad (38)$$

where $a \in \mathbf{Proc}_{A \rightarrow D}$, $b \in \mathbf{Proc}_{B \rightarrow D'}$, and $c \in \mathbf{Proc}_{D \otimes D' \rightarrow C}$. Any $c \in \mathbf{Proc}_{D \otimes D' \rightarrow C}$ can be expressed by

$$\begin{array}{c} \vdots \\ \boxed{c} \\ \vdots \end{array} \stackrel{(10)}{=} \sum_{i,j,m=1}^M \sum_{i,j,m=1}^M \begin{array}{c} \vdots \\ \boxed{m} \\ \vdots \\ \boxed{c} \\ \vdots \quad \vdots \\ \boxed{i} \quad \boxed{j} \\ \vdots \quad \vdots \\ \boxed{i} \quad \boxed{j} \\ \vdots \quad \vdots \end{array} = \sum_{i,j,m=1}^M c_{i,j,m} \begin{array}{c} \vdots \\ \boxed{m} \\ \vdots \\ \boxed{i} \quad \boxed{j} \\ \vdots \quad \vdots \end{array}, \quad (39)$$

where $c_{i,j,m} \in \mathbf{R}_+$ is the scalar enclosed by the auxiliary box. Using Eq. (39), we can easily see that a necessary and sufficient condition for $e \in \mathbf{Meas}_{A \otimes B}$ to be separable is that $(m| \circ e$ is separable for each $m \in \mathcal{I}_M$.

A measurement $e \in \mathbf{Meas}_{A \otimes B}$ is referred to as *PT* if

$$\begin{array}{c} \vdots \text{C} \\ \boxed{e} \\ \vdots \quad \vdots \\ \text{A} \quad \text{B} \\ \boxed{f} \\ \vdots \\ \text{A}' \end{array} \in \mathbf{Meas}_{A' \otimes B}$$

holds for any system A' and any deterministic extended process $\bar{f} \in \mathbf{V}_{A' \rightarrow A}$ that is positive for effects.

Let $\mathbf{Meas}_{A \rightarrow B}^{\text{SEQ}}$ be the set of all sequential measurements from A to B . Also, let $\mathbf{Meas}_{A \otimes B}^{\text{LOCC}}$, $\mathbf{Meas}_{A \otimes B}^{\text{SEP}}$, and $\mathbf{Meas}_{A \otimes B}^{\text{PT}}$ be, respectively, a set of all LOCC, separable, and PT measurements of $A \otimes B$.

Proposition 3:

$$\begin{aligned} \mathbf{Meas}_{A \rightarrow B}^{\text{SEQ}} &\subseteq \mathbf{Meas}_{A \otimes B}^{\text{LOCC}} \subseteq \mathbf{Meas}_{A \otimes B}^{\text{SEP}} \\ &\subseteq \mathbf{Meas}_{A \otimes B}^{\text{PT}} \subseteq \mathbf{Meas}_{A \otimes B}. \end{aligned}$$

Proof. $\mathbf{Meas}_{A \rightarrow B}^{\text{SEQ}} \subseteq \mathbf{Meas}_{A \otimes B}^{\text{LOCC}}$ and $\mathbf{Meas}_{A \otimes B}^{\text{PT}} \subseteq \mathbf{Meas}_{A \otimes B}$

obviously hold. $\mathbf{Meas}_{A \otimes B}^{\text{SEQ}} \subseteq \mathbf{Meas}_{A \otimes B}^{\text{SEP}}$ follows from

$$\begin{array}{c} \vdots \\ \boxed{b} \\ \vdots \\ \boxed{a} \end{array} \stackrel{(10)}{=} \sum_i \begin{array}{c} \vdots \\ \boxed{i} \\ \vdots \\ \boxed{a} \end{array} \stackrel{(17)}{=} \sum_i \begin{array}{c} \vdots \\ \boxed{i} \quad \boxed{i} \\ \vdots \quad \vdots \\ \text{D} \quad \text{D}' \\ \boxed{a} \quad \boxed{b} \\ \vdots \quad \vdots \\ \text{D} \quad \text{C} \end{array} = \begin{array}{c} \vdots \\ \boxed{c} \\ \vdots \\ \text{D} \quad \text{D} \otimes \text{C} \\ \boxed{a} \quad \boxed{b'} \\ \vdots \quad \vdots \end{array},$$

where c and b' are the processes enclosed by the upper and lower auxiliary boxes, respectively. $\mathbf{Meas}_{A \otimes B}^{\text{LOCC}} \subseteq \mathbf{Meas}_{A \otimes B}^{\text{SEP}}$ can be immediately proved in the same way. The proof is completed by showing $\mathbf{Meas}_{A \otimes B}^{\text{SEP}} \subseteq \mathbf{Meas}_{A \otimes B}^{\text{PT}}$. Let $e \in \mathbf{Meas}_{A \otimes B}^{\text{SEP}}$ be expressed in the form of Eq. (38). Arbitrarily choose a deterministic extended process $\bar{f} \in \mathbf{V}_{A' \rightarrow A}$ that is positive for effects. Since $e' := e \circ (\bar{f} \otimes \text{id}_B)$ is obviously deterministic, it remains to prove that e' is a process. We have

$$\begin{array}{c} \vdots \\ \boxed{e'} \\ \vdots \end{array} \stackrel{(38)}{=} \begin{array}{c} \vdots \\ \boxed{c} \\ \vdots \\ \boxed{a} \quad \boxed{b} \\ \vdots \quad \vdots \\ \boxed{f} \end{array} \stackrel{(39)}{=} \sum_{i,j,m=1}^M c_{i,j,m} \begin{array}{c} \vdots \\ \boxed{m} \\ \vdots \\ \boxed{i} \quad \boxed{j} \\ \vdots \quad \vdots \\ \boxed{a} \quad \boxed{b} \\ \vdots \quad \vdots \\ \boxed{f} \end{array}.$$

Since \bar{f} is positive for effects, the extended effect enclosed by the auxiliary box is an effect, and thus e' is a process. Therefore, e is PT. \square

Proposition 4: In quantum theory, $\mathbf{Meas}_{A \otimes B}^{\text{SEP}} = \mathbf{Meas}_{A \otimes B}^{\text{PT}}$ holds.

Proof. Since $\mathbf{Meas}_{A \otimes B}^{\text{SEP}} \subseteq \mathbf{Meas}_{A \otimes B}^{\text{PT}}$ holds, it suffices to show $\mathbf{Meas}_{A \otimes B}^{\text{SEP}} \supseteq \mathbf{Meas}_{A \otimes B}^{\text{PT}}$, i.e., for any $e \in \mathbf{Meas}_{A \otimes B}^{\text{PT}}$ and $m \in \mathcal{I}_M$, $(m| \circ e$ is separable.

Assume, by contradiction, that $(e_m| := (m| \circ e$ is entangled for some $e \in \mathbf{Meas}_{A \otimes B}^{\text{PT}}$ and $m \in \mathcal{I}_M$. It has been shown in Ref. [37] that there exists an extended state $|\bar{v}\rangle \in \mathbf{V}_{A \otimes B}$ such that $(e_m|\bar{v}\rangle < 0$ and $[(b| \otimes (b'|)|\bar{v}\rangle] \geq 0$ hold for any $(b| \in \mathbf{Eff}_A$ and $(b'| \in \mathbf{Eff}_B$. One can easily verify that, for each $(b| \in \mathbf{Eff}_A$, $|v'\rangle := [(b| \otimes \text{id}_B) \circ |\bar{v}\rangle] \in \mathbf{St}_B$ holds since $(b'|v'\rangle \geq 0$ holds for any $(b'| \in \mathbf{Eff}_B$. Let $|v_B\rangle := [(\bar{\tau}_A| \otimes \text{id}_B) \circ |\bar{v}\rangle] \in \mathbf{St}_B$. Assume, without loss of generality, that $|v_B\rangle$ is full rank (if not, we can replace $|\bar{v}\rangle$ with $|\bar{v}\rangle + c|\rho_0\rangle \otimes |\perp_B\rangle$ with $0 < c \in \mathbf{R}_+$, $|\rho_0\rangle \in \mathbf{St}_A^N$, and $|\perp_B\rangle := \mathbb{1}_{N_B}$, where c is sufficiently small such that $(e_m|\bar{v}\rangle < 0$). It is easily seen that there exists a reversible process $g \in \mathbf{Proc}_{B \rightarrow B}$ such that $g \circ |v_B\rangle = |\perp_B\rangle$. Let g^{-1} be

the inverse of g . It is well-known that, for each system A , there exist $|\cup_A\rangle \in \mathbf{St}_{A\otimes A}$ and $(\cap_A| \in \mathbf{Eff}_{A\otimes A}$ such that $[(\cap_A| \otimes \text{id}_A] \circ [\text{id}_A \otimes |\cup_A\rangle] = \text{id}_A$ and $(\cap_A| \circ [|\perp_{\pm A}\rangle] \otimes \text{id}_A = (\overline{\cap}_A|$ (see, e.g., [32]). Let

$$\begin{array}{c} \overline{\cap}_A| \\ \hline \boxed{f} \\ \hline B \end{array} := \begin{array}{c} \cap_B \\ \hline \begin{array}{c} A \\ \boxed{g} \\ B \end{array} \\ \hline \cup_B \\ \hline \overline{v} \end{array}.$$

One can easily verify

$$\begin{array}{c} \overline{\cap}_A| \\ \hline \boxed{f} \\ \hline B \end{array} = \begin{array}{c} \cap_B \\ \hline \begin{array}{c} A \\ \boxed{g} \\ B \end{array} \\ \hline \cup_B \\ \hline \overline{v} \end{array} = \begin{array}{c} \cap_B \\ \hline \begin{array}{c} B \\ \boxed{g} \\ B \end{array} \\ \hline \cup_B \\ \hline \overline{v} \end{array} = \begin{array}{c} \overline{\cap}_B| \\ \hline B \end{array},$$

i.e., \overline{f} is deterministic. Also, \overline{f} is positive for effects since, for any $(b| \in \mathbf{Eff}_A$, from $[(b| \otimes \text{id}_B] \circ |\overline{v}\rangle \in \mathbf{St}_B$, $(b| \circ \overline{f} \in \mathbf{Eff}_B$ holds. Moreover, we have

$$\begin{array}{c} e_m \\ \hline \begin{array}{c} A \\ \boxed{f} \\ B \end{array} \\ \hline \cup_B \end{array} = \begin{array}{c} e_m \\ \hline \begin{array}{c} A \\ \cap_B \\ \hline \begin{array}{c} B \\ \boxed{g} \\ B \end{array} \\ \hline \cup_B \\ \hline \overline{v} \end{array} \\ \hline \cup_B \end{array} = \begin{array}{c} e_m \\ \hline \begin{array}{c} A \\ \boxed{g^{-1}} \\ B \end{array} \\ \hline \cup_B \end{array} < 0$$

which means that $(e_m| \circ (\overline{f} \otimes \text{id}_B))$ is not an effect. This contradicts $e \in \mathbf{Meas}_{A\otimes B}^{\text{PT}}$. Therefore, $\mathbf{Meas}_{A\otimes B}^{\text{SEP}} \supseteq \mathbf{Meas}_{A\otimes B}^{\text{PT}}$ holds. \square

Obviously, $\mathbf{Meas}_{A\rightarrow B}^{\text{SEQ}}$, $\mathbf{Meas}_{A\otimes B}^{\text{LOCC}}$, $\mathbf{Meas}_{A\otimes B}^{\text{SEP}}$ and $\mathbf{Meas}_{A\otimes B}^{\text{PT}}$ are convex subsets of $\mathbf{Meas}_{A\otimes B}$ that are symmetric under permutations of the measurement outcomes. Thus, the following corollary follows immediately from Theorem 1 with $A \otimes B$ instead of A :

Corollary 5: Let $\mathcal{M}_{A\otimes B}$ be $\mathbf{Meas}_{A\rightarrow B}^{\text{SEQ}}$, $\mathbf{Meas}_{A\otimes B}^{\text{LOCC}}$, $\mathbf{Meas}_{A\otimes B}^{\text{SEP}}$ or $\mathbf{Meas}_{A\otimes B}^{\text{PT}}$. Let $\tau : \mathcal{G} \rightarrow \text{Aut}(\mathcal{I}_M)$ and $\overline{\pi} : \mathcal{G} \rightarrow \text{Aut}(\mathcal{M}_A)$ be group actions with a group \mathcal{G} . If a state preparation $\rho \in \mathbf{Prep}_{A\otimes B}$ is $(\mathcal{G}, \tau, \overline{\pi})$ -covariant, then, for any measurement $e \in \mathcal{M}_{A\otimes B}$, there exists a $(\mathcal{G}, \tau, \overline{\pi})$ -covariant measurement $e^\diamond \in \mathcal{M}_{A\otimes B}$ that satisfies $P(e^\diamond) = P(e)$.

Example of quantum theory It is easily seen that each element of $\text{Aut}(\mathcal{M}_{A\otimes B})$ with $\mathcal{M}_{A\otimes B} \in$

$\{\mathbf{Meas}_{A\rightarrow B}^{\text{SEQ}}, \mathbf{Meas}_{A\otimes B}^{\text{LOCC}}, \mathbf{Meas}_{A\otimes B}^{\text{SEP}}\}$ maps normalized pure effects to normalized pure effects (where a pure effect is called normalized if its trace is one) and thus maps normalized pure states to normalized pure states. Let \mathcal{P} be the set of all normalized separable pure states of $A \otimes B$. Also, let \mathcal{P}^* be the set of all normalized separable pure effects of $A \otimes B$. For each $(e| \in \mathcal{P}^*$, there exists a sequential measurement $\{(e_m|)_{m \in \mathcal{I}_M}\}$ satisfying $(e_1| = (e|$. Thus, we can easily see that, for any $\overline{f} \in \text{Aut}(\mathbf{Meas}_{A\rightarrow B}^{\text{SEQ}})$, we have $\{(e| \circ \overline{f} : (e| \in \mathcal{P}^*) = \mathcal{P}^*$, i.e.,

$$\{\overline{f} \circ |\psi\rangle : |\psi\rangle \in \mathcal{P}\} = \mathcal{P}. \quad (40)$$

It follows, by the same argument, that Eq. (40) also holds for any $\overline{f} \in \text{Aut}(\mathbf{Meas}_{A\otimes B}^{\text{LOCC}})$ and for any $\overline{f} \in \text{Aut}(\mathbf{Meas}_{A\otimes B}^{\text{SEP}})$. According to Theorem 3 of Ref. [36], $\overline{f} \in \mathbf{V}_{A\otimes B \rightarrow A\otimes B}$ satisfies Eq. (40) if and only if \overline{f} has one of the following forms:

- 1) $\overline{f} \circ |\rho\rangle = (U_1 \otimes U_2) \cdot |\rho\rangle \cdot (U_1 \otimes U_2)^\dagger$,
- 2) $\overline{f} \circ |\rho\rangle = (U_1 \otimes U_2) \cdot [\times_{A,B} \circ |\rho\rangle] \cdot (U_1 \otimes U_2)^\dagger$,

where $|\rho\rangle \in \mathbf{St}_{A\otimes B}$. U_1 and U_2 are, respectively, unitary or anti-unitary matrices of order N_A and N_B . \overline{f} can have the form of 2) only in the case of $A \cong B$. Note that $\times_{A,B}$ satisfies $\times_{A,B} \circ [|\rho_1\rangle \otimes |\rho_2\rangle] = |\rho_2\rangle \otimes |\rho_1\rangle$ for any $|\rho_1\rangle \in \mathbf{St}_A$ and $|\rho_2\rangle \in \mathbf{St}_B$. In each of these cases, we have:

- 1) $(e| \circ \overline{f} = (U_1 \otimes U_2)^\dagger \cdot (e| \cdot (U_1 \otimes U_2)$,
- 2) $(e| \circ \overline{f} = (U_2 \otimes U_1)^\dagger \cdot [(e| \circ \times_{B,A}) \cdot (U_2 \otimes U_1)$,

where $(e| \in \mathbf{Eff}_{A\otimes B}$.

One can easily see that any $\overline{f} \in \text{Aut}(\mathbf{Meas}_{A\rightarrow B}^{\text{SEQ}})$ must have the form of 1) unless $A \cong B \cong I$. (Indeed, if there exists $\overline{f} \in \text{Aut}(\mathbf{Meas}_{A\rightarrow B}^{\text{SEQ}})$ with the form of 2), then, $e \circ \overline{f} \in \mathbf{Meas}_{B\rightarrow A}^{\text{SEQ}}$ holds for any $e \in \mathbf{Meas}_{A\rightarrow B}^{\text{SEQ}}$. In this case, since $\{e \circ \overline{f} : e \in \mathbf{Meas}_{A\rightarrow B}^{\text{SEQ}}\} = \mathbf{Meas}_{B\rightarrow A}^{\text{SEQ}}$ holds from $\overline{f} \in \text{Aut}(\mathbf{Meas}_{A\rightarrow B}^{\text{SEQ}})$, $\mathbf{Meas}_{A\rightarrow B}^{\text{SEQ}} \subseteq \mathbf{Meas}_{B\rightarrow A}^{\text{SEQ}}$ must hold. However, it is immediately seen that there exists $e \in \mathbf{Meas}_{A\rightarrow B}^{\text{SEQ}}$ satisfying $e \notin \mathbf{Meas}_{B\rightarrow A}^{\text{SEQ}}$.) This type of symmetry with respect to $\mathbf{Meas}_{A\rightarrow B}^{\text{SEQ}}$ has been discussed in Ref. [26]. If $\mathcal{M}_{A\otimes B} \in \{\mathbf{Meas}_{A\otimes B}^{\text{LOCC}}, \mathbf{Meas}_{A\otimes B}^{\text{SEP}}\}$ and $A \cong B$ hold, then $\overline{f} \in \text{Aut}(\mathcal{M}_{A\otimes B})$ can have the form of 2).

V. CONCLUSION

The problem of discrimination of symmetric states in an OPT has been investigated in diagrammatic terms. It is well-known that, in quantum theory, if states have a certain symmetry, then there exists a minimum-error measurement that has the same type of symmetry. We showed in Theorem 1 that this property is also valid in a more general OPT. We also showed that this result can be utilized to optimize over a restricted class of measurements. Although we discuss only the minimum-error strategy to simplify the discussion, this result can be easily applied to other various criteria, such as the Bayes criterion or the minimax criterion.

ACKNOWLEDGMENT

I am grateful to O. Hirota, K. Kato, and T. S. Usuda for support. This work was supported by JSPS KAKENHI Grant Number JP19K03658.

REFERENCES

- [1] J. Barrett, "Information processing in generalized probabilistic theories," *Phys. Rev. A*, vol. 75, no. 3, 032304, 2007.
- [2] H. Barnum, J. Barrett, L. O. Clark, M. Leifer, R. Spekkens, N. Stepanik, A. Wilce, and R. Wilke, "Entropy and information causality in general probabilistic theories," *New J. Phys.*, vol. 12, no. 3, 033024, 2010.
- [3] G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Probabilistic theories with purification," *Phys. Rev. A*, vol. 81, no. 6, 062348, 2010.
- [4] G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Informational derivation of quantum theory," *Phys. Rev. A*, vol. 84, no. 1, 012311, 2011.
- [5] S. Gogioso and C. M. Scandolo, "Categorical probabilistic theories," *arXiv preprint arXiv:1701.08075*, 2017.
- [6] S. Popescu and D. Rohrlich, "Quantum nonlocality as an axiom," *Found. Phys.*, vol. 24, no. 3, pp. 379–385, 1994.
- [7] A. S. Holevo, "Statistical decision theory for quantum systems," *J. Multivar. Anal.*, vol. 3, pp. 337–394, 1973.
- [8] H. P. Yuen, K. S. Kennedy, and M. Lax, "Optimum testing of multiple hypotheses in quantum detection theory," *IEEE Trans. Inf. Theory*, vol. 21, no. 2, pp. 125–134, 1975.
- [9] C. W. Helstrom, *Quantum detection and estimation theory*. Academic Press, 1976.
- [10] Y. C. Eldar, A. Megretski, and G. C. Verghese, "Designing optimal quantum detectors via semidefinite programming," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 1007–1012, 2003.
- [11] V. P. Belavkin, "Optimal multiple quantum statistical hypothesis testing," *Stochastics*, vol. 1, no. 1, pp. 315–345, 1975.
- [12] M. Ban, K. Kurokawa, R. Momose, and O. Hirota, "Optimum measurements for discrimination among symmetric quantum states and parameter estimation," *Int. J. Theor. Phys.*, vol. 36, no. 6, pp. 1269–1288, 1997.
- [13] T. S. Usuda, I. Takumi, M. Hata, and O. Hirota, "Minimum error detection of classical linear code sending through a quantum channel," *Phys. Lett. A*, vol. 256, pp. 104–108, 1999.
- [14] Y. C. Eldar and G. D. Forney Jr., "On quantum detection and the square-root measurement," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 858–872, 2001.
- [15] S. M. Barnett and S. Croke, "On the conditions for discrimination between quantum states with minimum error," *J. Phys. A: Math. Theor.*, vol. 42, no. 6, 062001, 2009.
- [16] K. Nakahira, K. Kato, and T. S. Usuda, "Generalized quantum state discrimination problems," *Phys. Rev. A*, vol. 91, no. 5, 052304, 2015.
- [17] T. S. Usuda, S. Usami, I. Takumi, and M. Hata, "Superadditivity in capacity of quantum channel for q -ary linearly dependent real symmetric-state signals," *Phys. Lett. A*, vol. 305, pp. 125–134, 2002.
- [18] E. Andersson, S. M. Barnett, C. R. Gilson, and K. Hunter, "Minimum-error discrimination between three mirror-symmetric states," *Phys. Rev. A*, vol. 65, no. 5, 052308, 2002.
- [19] K. Kato and O. Hirota, "Square-root measurement for quantum symmetric mixed state signals," *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3312–3317, 2003.
- [20] C. L. Chou and L. Y. Hsu, "Minimum-error discrimination between symmetric mixed quantum states," *Phys. Rev. A*, vol. 68, no. 4, 042305, 2003.
- [21] Y. C. Eldar, A. Megretski, and G. C. Verghese, "Optimal detection of symmetric mixed quantum states," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1198–1207, 2004.
- [22] T. Sawada, T. Tsuchimoto, and T. S. Usuda, "Quantum detection of artificial mixed-state signals," in *Proceedings of the 8th International Conference on Quantum Communication, Measurement and Computing*, pp. 405–408, 2007.
- [23] G. Cariolaro and A. Vigato, "Helstrom's theory on quantum binary decision revisited," in *Information Theory Workshop (ITW), 2011 IEEE*, pp. 242–246, IEEE, 2011.
- [24] K. Nakahira and T. S. Usuda, "Minimum-error discrimination between self-symmetric mixed quantum state signals," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1215–1222, 2012.
- [25] K. Nakahira and T. S. Usuda, "Quantum measurement for a group-covariant state set," *Phys. Rev. A*, vol. 87, no. 1, 012308, 2013.
- [26] K. Nakahira, K. Kato, and T. S. Usuda, "Generalized bipartite quantum state discrimination problems with sequential measurements," *Phys. Rev. A*, vol. 97, no. 2, 022340, 2018.
- [27] L. Hardy, "Reformulating and reconstructing quantum theory," *arXiv preprint arXiv:1104.2066*, 2011.
- [28] P. Janotta and R. Lal, "Generalized probabilistic theories without the no-restriction hypothesis," *Phys. Rev. A*, vol. 87, no. 5, 052131, 2013.
- [29] K. Nakahira, "Derivation of quantum theory with superselection rules," *arXiv preprint arXiv:1910.02649*, 2019.
- [30] B. Coecke, "The logic of entanglement: an invitation," *Technical Report RR-03-12, Department of Computer Science, Oxford University*, 2003.
- [31] S. Abramsky and B. Coecke, "A categorical semantics of quantum protocols," in *Proceedings of the 19th Annual IEEE Symposium on Logic in Computer Science*, pp. 415–425, IEEE, 2004.
- [32] B. Coecke and A. Kissinger, *Picturing quantum processes*. Cambridge University Press, 2017.
- [33] G. Kimura, T. Miyadera, and H. Imai, "Optimal state discrimination in general probabilistic theories," *Phys. Rev. A*, vol. 79, no. 6, 062306, 2009.
- [34] K. Nuida, G. Kimura, and T. Miyadera, "Optimal observables for minimum-error state discrimination in general probabilistic theories," *J. Math. Phys.*, vol. 51, no. 9, 093505, 2010.
- [35] J. Bae, "Distinguishability, ensemble steering, and the no-signaling principle," *arXiv preprint arXiv:1210.3125*, 2012.
- [36] S. Friedland, C.-K. Li, Y.-T. Poon, and N.-S. Sze, "The automorphism group of separable states in quantum information theory," *J. Math. Phys.*, vol. 52, no. 4, 042203, 2011.
- [37] M. Horodecki, P. Horodecki, and R. Horodecki, "Separability of mixed states: necessary and sufficient conditions," *Phys. Lett. A*, vol. 223, no. 1, pp. 1–8, 1996.