PHYSICAL REVIEW A **91**, 052304 (2015)

Generalized quantum state discrimination problems

Kenji Nakahira,[1,4] Kentaro Kato,[2] and Tsuyoshi Sasaki Usuda[3,4]

[1]*Hitachi, Ltd., Research & Development Group, Center for Technology Innovation-Production Engineering, Yokohama, Kanagawa 244-0817, Japan*

[2]*Quantum Communication Research Center, Quantum ICT Research Institute, Tamagawa University, Machida, Tokyo 194-8610, Japan*

[3]*School of Information Science and Technology, Aichi Prefectural University, Nagakute, Aichi 480-1198, Japan*

[4]*Quantum Information Science Research Center, Quantum ICT Research Institute, Tamagawa University, Machida, Tokyo 194-8610, Japan*

(Received 25 January 2015; published 6 May 2015)

We address a broad class of optimization problems of finding quantum measurements, which includes the problems of finding an optimal measurement in the Bayes criterion and a measurement maximizing the average correct probability with a fixed rate of inconclusive results. Our approach can deal with any problem in which each of the objective and constraint functions is formulated by the sum of the traces of the multiplication of a Hermitian operator and a detection operator. We first derive dual problems and necessary and sufficient conditions for an optimal measurement. We also consider the minimax version of these problems and provide necessary and sufficient conditions for a minimax solution. Finally, for optimization problem having a certain symmetry, there exists an optimal solution with the same symmetry. Examples are shown to illustrate how our results can be used.

DOI: 10.1103/PhysRevA.91.052304 PACS number(s): 03.67.Hk

I. INTRODUCTION

Discrimination between quantum states is a fundamental topic in quantum information theory. Research in quantum state discrimination was pioneered by Helstrom, Holevo, and Yuen *et al.* [1–3] in the 1970s and has attracted intensive attention. It is well known in quantum mechanics that nonorthogonal states cannot be discriminated with certainty. Thus, optimal measurement strategies have been proposed under various criteria. Among them, one of the most widely investigated is the Bayes criterion, or the criterion of minimum average error probability [1–3]. In the Bayes criterion, necessary and sufficient conditions for obtaining an optimal measurement have been formulated [1–4], and closed-form analytical expressions for optimal measurements have also been derived in some cases (see, e.g., [5–9]). This criterion is based on the assumption that prior probabilities of the states are known. In contrast, if these prior probabilities are unknown, then minimax criteria are often used [10,11]. Necessary and sufficient conditions for a measurement minimizing the worst case of the average error probability in the minimax strategy have been found [10]. This result has also been extended to the average Bayes cost [12].

Other types of optimal measurements have been investigated. In the case in which prior probabilities of the states are known, an example concerns a measurement that achieves low average error probability at the expense of allowing for a certain fraction of inconclusive results [13–15]. In particular, an unambiguous (or error-free) measurement that maximizes the correct probability, which is called an optimal unambiguous measurement, has been well studied [13–15]. A measurement that maximizes the average correct probability with a fixed average inconclusive (or failure) probability, which is called an optimal inconclusive measurement, has also been studied [16–18]. Moreover, a measurement that maximizes the average correct probability where a certain fixed average error probability is allowed, which we call an optimal error margin measurement, has also been investigated [19–21]. On the other hand, in the case in which prior probabilities are unknown, several types of measurements based on the minimax strategy have been proposed, such as a measurement that minimizes the maximum probability of detection errors [22] and a measurement with a certain fraction of inconclusive results [23]. Properties of optimal measurements in the above criteria, such as necessary and sufficient conditions for optimal solutions, have been derived for each criterion.

In this paper, we investigate optimization problems of finding optimal quantum measurements and their minimax versions that are applicable to a wide range of quantum state discrimination problems. Our approach can deal with any problem in which each of the objective and constraint functions is formulated by the sum of the traces of the multiplication of a Hermitian operator and a detection operator, which implies that any problems related to finding any of the optimal measurements described above can be formulated as our problems. Thus, we can say that our approach can provide a unified treatment in a large class of problems. The results obtained in this paper would be valuable from the practical point of view; for example, they not only provide a broader perspective than the results for a particular problem, but also can apply to many problems that have not been reported previously, some examples of which are presented in this paper. To obtain knowledge about an optimal measurement in a new criterion has the potential to create a new application of quantum state discrimination.

In Sec. II, we provide a generalized optimization problem in which each of the objective and constraint functions is formulated by the sum of the traces of the multiplication of a Hermitian operator and a detection operator. We derive its dual problem and necessary and sufficient conditions for an optimal measurement. In Sec. III, we discuss the minimax version of our generalized problem and provide necessary and sufficient conditions for a minimax solution. In Sec. IV, we demonstrate that if a given problem has a certain symmetry, then there exists an optimal solution with the same symmetry. Finally, we present some examples to illustrate the applicability of our results in Sec. V.

II. GENERALIZED OPTIMAL MEASUREMENT

A. Formulation

We consider a quantum measurement on a Hilbert space $\mathcal{H}$. Such a quantum measurement can be modeled by a positive operator-valued measure (POVM) $\Pi = \{\hat{\Pi}_m : m \in \mathcal{I}_M\}$ on $\mathcal{H}$, where M is the number of the detection operators and $\mathcal{I}_k = \{0,1,\dots,k-1\}$. An example of a quantum measurement is an optimal measurement for distinguishing R quantum states represented by density operators $\hat{\rho}_r (r \in \mathcal{I}_R)$. The density operator $\hat{\rho}_r$ satisfies $\hat{\rho}_r \geqslant 0$ and has unit trace, i.e., $\mathrm{Tr}\hat{\rho}_r = 1$, where $\hat{A} \geqslant 0$ denotes that $\hat{A}$ are positive semidefinite (similarly, $\hat{A} \geqslant \hat{B}$ denotes $\hat{A} - \hat{B} \geqslant 0$). A minimum error measurement is such an optimal measurement, which can be expressed by a POVM with $M = R$ detection operators. A quantum measurement that may return an inconclusive answer can be expressed by a POVM with $M = R + 1$ detection operators; in this case the detection operator $\hat{\Pi}_r$ $(r \in \mathcal{I}_R)$ corresponds to identification of the state $\hat{\rho}_r$, while $\hat{\Pi}_R$ corresponds to the inconclusive answer.

Let $\mathcal{M}$ be the entire set of POVMs on $\mathcal{H}$ that consist of M detection operators. $\Pi \in \mathcal{M}$ satisfies

$$\hat{\Pi}_m \geqslant 0, \ \forall\, m \in \mathcal{I}_M, \qquad \sum_{m=0}^{M-1} \hat{\Pi}_m = \hat{1}, \tag{1}$$

where $\hat{1}$ is the identity operator on $\mathcal{H}$. In addition, let $\mathcal{S}$ and $\mathcal{S}_+$ be the entire sets of Hermitian operators on $\mathcal{H}$ and semidefinite positive operators on $\mathcal{H}$, respectively. Let $\mathcal{R}$ and $\mathcal{R}_+$ be the entire sets of real numbers and non-negative real numbers, respectively, and $\mathcal{R}_+^N$ be the entire set of collections of N nonnegative real numbers.

Here we consider a generalized optimization problem. The conditional probability that the measurement outcome is m when a quantum state $\hat{\rho}$ is given is represented by $\mathrm{Tr}(\hat{\rho}\hat{\Pi}_m)$, and thus there exist many optimization problems of finding optimal quantum measurements such that each of the objective and constraint functions is expressed by a linear combination of forms $\mathrm{Tr}(\hat{\rho}_r\hat{\Pi}_m)$. For this reason, we consider an optimization problem,

$$\text{maximize} \quad f(\Pi) = \sum_{m=0}^{M-1} \mathrm{Tr}(\hat{c}_m \hat{\Pi}_m) \qquad \text{subject to} \quad \Pi \in \mathcal{M}^\circ, \tag{2}$$

where $\hat{c}_m \in \mathcal{S}$ holds for any $m \in \mathcal{I}_M$. (Note that any linear combination of positive semidefinite operators is a Hermitian operator.) $\mathcal{M}^\circ$ is expressed by

$$\mathcal{M}^\circ = \left\{ \Pi \in \mathcal{M} : \sum_{m=0}^{M-1} \mathrm{Tr}(\hat{a}_{j,m}\hat{\Pi}_m) \leqslant b_j, \forall\, j \in \mathcal{I}_J \right\}, \tag{3}$$

where $\hat{a}_{j,m} \in \mathcal{S}$ and $b_j \in \mathcal{R}$ hold for any $m \in \mathcal{I}_M$ and $j \in \mathcal{I}_J$. J is a non-negative integer. We should mention that an equality constraint [e.g., $\mathrm{Tr}(\hat{a}_{j,0}\hat{\Pi}_0) = b_j$] can be replaced by two inequality constraints [e.g., $\mathrm{Tr}(\hat{a}_{j,0}\hat{\Pi}_0) \leqslant b_j$ and $\mathrm{Tr}(-\hat{a}_{j,0}\hat{\Pi}_0) \leqslant -b_j$]. We call an optimal solution to problem (2) a generalized optimal measurement or simply an optimal measurement. Problem (2) is said to be a primal problem. Since $f(\Pi)$ is linear in Π and $\mathcal{M}^\circ$ is convex, problem (2) is a convex optimization problem. Note that since the constraint of $\Pi \in \mathcal{M}$, i.e., Eq. (1), can be rewritten as $\mathrm{Tr}(\hat{\rho}\hat{\Pi}_m) \geqslant 0$ and $\sum_{m=0}^{M-1} \mathrm{Tr}(\hat{\rho}\hat{\Pi}_m) = 1$ for any density operator $\hat{\rho}$, we can say that each of the objective and constraint functions is formulated by the sum of the traces of the multiplication of a Hermitian operator and a detection operator.

B. Examples

We give some examples of optimization problems of finding quantum measurements that can be formulated as problem (2). Let us consider discrimination between R quantum states $\{\hat{\rho}_r : r \in \mathcal{I}_R\}$ with prior probabilities $\{\xi_r : r \in \mathcal{I}_R\}$.

1. Optimal measurement in the Bayes criterion

The optimization problem of finding an optimal measurement in the Bayes criterion is formulated as [1–3]

$$\text{minimize} \quad \sum_{m=0}^{R-1} \mathrm{Tr}(\hat{W}_m \hat{\Pi}_m) \qquad \text{subject to} \quad \Pi \in \mathcal{M}. \tag{4}$$

$\hat{W}_m \in \mathcal{S}_+$ $(m \in \mathcal{I}_R)$ can be expressed by

$$\hat{W}_m = \sum_{r=0}^{R-1} \xi_r B_{m,r} \hat{\rho}_r, \tag{5}$$

where $B_{m,r} \in \mathcal{R}_+$ holds for any $m,r \in \mathcal{I}_R$. This problem can be written as the form of problem (2) with

$$M = R, \quad J = 0, \quad \hat{c}_m = -\hat{W}_m. \tag{6}$$

2. Optimal error margin measurement

An optimal error margin measurement is a measurement maximizing the average correct probability under the constraint that the average error probability is not greater than a given value ε, with $0 \leqslant \varepsilon \leqslant 1$ [19–21]. In particular, if $\varepsilon = 0$, then an optimal error margin measurement is equivalent to an optimal unambiguous measurement. The optimization problem of finding an optimal error margin measurement is formulated as

$$\text{maximize} \quad \sum_{r=0}^{R-1} \xi_r \mathrm{Tr}(\hat{\rho}_r \hat{\Pi}_r) \qquad \text{subject to} \quad \Pi \in \mathcal{M}, \sum_{r=0}^{R-1} \xi_r \mathrm{Tr}[\hat{\rho}_r(\hat{\Pi}_r + \hat{\Pi}_R)] \geqslant 1 - \varepsilon, \tag{7}$$

where we consider that the statement that the average error probability is not greater than ε is equivalent to the statement that the sum of the average correct and inconclusive probabilities is not less than $1 - \varepsilon$. This problem can be written as the form of problem (2), with

$$\begin{aligned} M &= R+1, \\ J &= 1, \\ \hat{c}_m &= \begin{cases} \xi_m \hat{\rho}_m, & m < R, \\ 0, & m = R, \end{cases} \\ \hat{a}_{0,m} &= \begin{cases} -\xi_m \hat{\rho}_m, & m < R, \\ -\hat{G}, & m = R, \end{cases} \\ b_0 &= \varepsilon - 1, \end{aligned} \tag{8}$$

where

$$\hat{G} = \sum_{r=0}^{R-1} \xi_r \hat{\rho}_r. \tag{9}$$

Note that an optimal error margin measurement has strong relationship with an optimal inconclusive measurement [24,25]. However, if one wants to obtain an optimal error margin measurement for a given ε, then one needs to solve problem (7) instead of the problem of finding an optimal inconclusive measurement. Also note that in the case of $\varepsilon = 0$ (i.e., an optimal unambiguous measurement), one can use the fact that any optimal measurement $\{\hat{\Pi}_m^\star : m \in \mathcal{I}_{R+1}\}$ satisfies $\hat{\rho}_r \hat{\Pi}_m^\star = 0$ for any $r \neq m \in \mathcal{I}_R$, although we do not use it in this paper. Several techniques based on this fact have been developed, and in recent years important progress has been made [26–30].

3. *Optimal inconclusive measurement with a lower bound on correct probabilities*

Another example is an extension of the problem of finding an optimal inconclusive measurement. An optimal inconclusive measurement is a measurement maximizing the average correct probability under the constraint that the average inconclusive probability equals a given value p with $0 \leqslant p \leqslant 1$ [16–18]. Here we add the constraint that for each $r \in \mathcal{I}_R$ the correct probability of the state $\hat{\rho}_r$, i.e., $\mathrm{Tr}(\hat{\rho}_r \hat{\Pi}_r)$, is not less than a given value q with $0 \leqslant q \leqslant 1$. When $q = 0$, an optimal solution is an optimal inconclusive measurement. This problem is formulated as

$$\begin{aligned} &\text{maximize} \quad \sum_{r=0}^{R-1} \xi_r \mathrm{Tr}(\hat{\rho}_r \hat{\Pi}_r) \\ &\text{subject to} \quad \Pi \in \mathcal{M}, \mathrm{Tr}(\hat{\rho}_r \hat{\Pi}_r) \geqslant q, \quad \forall\, r \in \mathcal{I}_R, \\ &\qquad\qquad \mathrm{Tr}(\hat{G}\hat{\Pi}_R) = p, \end{aligned} \tag{10}$$

where $\hat{G}$ is defined by Eq. (9). Since the optimal value of problem (10) is monotonically decreasing with respect to p, we obtain the same solution if the last constraint of problem (10) is replaced with $\mathrm{Tr}(\hat{G}\hat{\Pi}_R) \geqslant p$. Thus, this problem is equivalent to problem (2) with

$$\begin{aligned} M &= J = R+1, \\ \hat{c}_m &= \begin{cases} \xi_m \hat{\rho}_m, & m < R, \\ 0, & m = R, \end{cases} \\ \hat{a}_{j,m} &= \begin{cases} -\delta_{m,j} \hat{\rho}_m, & j < R, \\ -\delta_{m,R} \hat{G}, & j = R, \end{cases} \\ b_j &= \begin{cases} -q, & j < R, \\ -p, & j = R, \end{cases} \end{aligned} \tag{11}$$

where $\delta_{k,k'}$ is the Kronecker δ. Note that if $q > q'$ holds, where q' is the average correct probability of an optimal inconclusive measurement with the average inconclusive probability of p, then this problem is infeasible; i.e., $\mathcal{M}^\circ$ is empty. We discuss this problem in detail in Sec. V A.

C. Dual problem

In this section, we show the dual problem of problem (2). We also show that the optimal values of primal problem (2) and the dual problem are the same.

Theorem 1. Let us consider problem (2). We also consider an optimization problem,

$$\begin{aligned} &\text{minimize} \quad s(\hat{X},\lambda) = \mathrm{Tr}\hat{X} + \sum_{j=0}^{J-1} \lambda_j b_j \\ &\text{subject to} \quad \hat{X} \geqslant \hat{z}_m(\lambda), \quad \forall\, m \in \mathcal{I}_M, \end{aligned} \tag{12}$$

with variables $\hat{X} \in \mathcal{S}$ and $\lambda = \{\lambda_j \in \mathcal{R}_+ : j \in \mathcal{I}_J\} \in \mathcal{R}_+^J$, where

$$\hat{z}_m(\lambda) = \hat{c}_m - \sum_{j=0}^{J-1} \lambda_j \hat{a}_{j,m}. \tag{13}$$

If $\mathcal{M}^\circ$ is not empty, then the optimal values of problems (2) and (12) are the same.

Problem (12) is called the dual problem of problem (2). Note that, in general, $\hat{X}$ satisfying the constraints of problem (12) is not in $\mathcal{S}_+$; however, it is obvious that if $m \in \mathcal{I}_M$ exists such that $\hat{z}_m \in \mathcal{S}_+$, then $\hat{X} \in \mathcal{S}_+$ holds.

Proof. Let us define the function L as

$$\begin{aligned} &L(\Pi,\sigma,\hat{X},\lambda) \\ &\quad = f(\Pi) + \sum_{m=0}^{M-1} \mathrm{Tr}(\hat{\sigma}_m \hat{\Pi}_m) + \mathrm{Tr}\left[\hat{X}\left(\hat{1} - \sum_{m=0}^{M-1} \hat{\Pi}_m\right)\right] \\ &\qquad + \sum_{j=0}^{J-1} \lambda_j \left[b_j - \sum_{m=0}^{M-1} \mathrm{Tr}(\hat{a}_{j,m}\hat{\Pi}_m)\right], \end{aligned} \tag{14}$$

where $\sigma = \{\hat{\sigma}_m \in \mathcal{S}_+ : m \in \mathcal{I}_M\}$, $\hat{X} \in \mathcal{S}$, and $\lambda \in \mathcal{R}_+^J$. Note that L is called the Lagrangian for problem (2). Substituting Eqs. (2), (12), and (13) into Eq. (14) gives

$$L(\Pi,\sigma,\hat{X},\lambda) = s(\hat{X},\lambda) + \sum_{m=0}^{M-1} \mathrm{Tr}\{[\hat{\sigma}_m + \hat{z}_m(\lambda) - \hat{X}]\hat{\Pi}_m\}. \tag{15}$$

Let us consider an optimization problem,

$$\begin{aligned} \text{minimize} \quad & s_\sigma(\sigma,\hat{X},\lambda) \\ \text{subject to} \quad & \hat{\sigma}_m \in \mathcal{S}_+, \forall\, m \in \mathcal{I}_M, \\ & \hat{X} \in \mathcal{S}, \\ & \lambda \in \mathcal{R}_+^J, \end{aligned} \tag{16}$$

where

$$s_\sigma(\sigma,\hat{X},\lambda) = \max_{\Pi \in \mathcal{S}_+^M} L(\Pi,\sigma,\hat{X},\lambda) \tag{17}$$

and $\mathcal{S}_+^M = \{\hat{\Pi}_m \in \mathcal{S}_+ : m \in \mathcal{I}_M\}$. Let $\mathcal{X} = \{\hat{X} : \hat{X} \geqslant \hat{\sigma}_m + \hat{z}_m(\lambda), \forall\, m \in \mathcal{I}_M\}$. The second term of the right-hand side of Eq. (15) is nonpositive if $\hat{X} \in \mathcal{X}$ and can be infinite if $\hat{X} \notin \mathcal{X}$. Therefore, from Eq. (17), $s_\sigma(\sigma,\hat{X},\lambda)$ can be expressed as

$$s_\sigma(\sigma,\hat{X},\lambda) = \begin{cases} s(\hat{X},\lambda), & \hat{X} \in \mathcal{X}, \\ \infty, & \text{otherwise}. \end{cases} \tag{18}$$

From Eq. (18), it follows that there exists an optimal solution to problem (16) such that $\hat{\sigma}_m = 0$ holds for any $m \in \mathcal{I}_M$. Indeed, if $(\sigma,\hat{X},\lambda)$ is an optimal solution to problem (16) [in this case, $\hat{X} \in \mathcal{X}$ holds from Eq. (18)], then $(\{\hat{\sigma}'_m = 0 : m \in \mathcal{I}_M\},\hat{X},\lambda)$ is also an optimal solution. Hence, problem (16) can be rewritten by problem (12).

Slater's condition is known to a sufficient condition under which, if the primal problem is convex, the optimal values of the primal and dual problems are the same [31]. Since each constraint of primal problem (2), including the constraint of $\Pi \in \mathcal{M}$, is expressed as a form of $u_j(\Pi) \leqslant 0$, where u_j is an affine function of Π, from Ref. [32], (the refined form of) Slater's condition is that the primal problem is feasible; i.e., $\mathcal{M}^\circ$ is not empty. Thus, since Slater's condition holds, the optimal values of problems (2) and (12) are the same. ■

It is worth noting that some attempts have been made to obtain the maximum average correct probability without using the fact that POVMs describe quantum measurements [33–35]. In Ref. [33], the dual problem to the problem of finding a minimum error measurement was derived from general probabilistic theories. In Refs. [34,35], the dual problem was derived from "ensemble steering," which determines what states one party can prepare on the other party's system by sharing a bipartite state. In the same way, we can derive dual problem (12) without using POVMs (see the Appendix). However, it might not be easy to prove that the optimal value of problem (12) is attained by using these approaches.

D. Conditions for an optimal measurement

Necessary and sufficient conditions for an optimal measurement in several problems (such as a minimum error measurement and an optimal inconclusive measurement) have been derived [1–4,17,18,26,36]. The following theorem extends these results to our more general setting.

Theorem 2. Suppose that a POVM Π is in $\mathcal{M}^\circ$. The following statements are all equivalent.

(1) Π is an optimal measurement of problem (2).

(2) $\hat{X} \in \mathcal{S}$ and $\lambda \in \mathcal{R}_+^J$ exist such that

$$\hat{X} - \hat{z}_m(\lambda) \geqslant 0, \quad \forall\, m \in \mathcal{I}_M, \tag{19}$$

$$[\hat{X} - \hat{z}_m(\lambda)]\hat{\Pi}_m = 0, \quad \forall\, m \in \mathcal{I}_M, \tag{20}$$

$$\lambda_j \left[b_j - \sum_{m=0}^{M-1} \mathrm{Tr}(\hat{a}_{j,m}\hat{\Pi}_m) \right] = 0, \quad \forall\, j \in \mathcal{I}_J. \tag{21}$$

(3) $\lambda \in \mathcal{R}_+^J$ exists such that

$$\sum_{n=0}^{M-1} \hat{z}_n(\lambda)\hat{\Pi}_n - \hat{z}_m(\lambda) \geqslant 0, \quad \forall\, m \in \mathcal{I}_M, \tag{22}$$

$$\lambda_j \left[b_j - \sum_{m=0}^{M-1} \mathrm{Tr}(\hat{a}_{j,m}\hat{\Pi}_m) \right] = 0, \quad \forall\, j \in \mathcal{I}_J. \tag{23}$$

Proof. It is sufficient to show (1) ⇒ (2), (2) ⇒ (3), and (3) ⇒ (1).

First, we show (1) ⇒ (2). Suppose that $(\hat{X},\lambda)$ is an optimal solution to dual problem (12). Let $\hat{\sigma}_m = 0$ for any $m \in \mathcal{I}_M$. It is obvious from Eq. (12) that Eq. (19) holds. From Theorem 1, $f(\Pi) = s(\hat{X},\lambda)$ holds. Moreover, from $\Pi \in \mathcal{M}^\circ$, the second and third terms of the right-hand side of Eq. (14) are zero, and the fourth term is non-negative, which yields

$$L(\Pi,\sigma,\hat{X},\lambda) \geqslant f(\Pi) = s(\hat{X},\lambda). \tag{24}$$

In contrast, from Eq. (19) and the fact that the trace of the multiplication of two positive semidefinite operators is non-negative, $\mathrm{Tr}\{[\hat{X} - \hat{z}_m(\lambda)]\hat{\Pi}_m\} \geqslant 0$ holds for any $m \in \mathcal{I}_M$, which yields $L(\Pi,\sigma,\hat{X},\lambda) \leqslant s(\hat{X},\lambda)$ from Eq. (15). Thus, from Eq. (24), we obtain $L(\Pi,\sigma,\hat{X},\lambda) = s(\hat{X},\lambda)$, i.e.,

$$\mathrm{Tr}\{[\hat{X} - \hat{z}_m(\lambda)]\hat{\Pi}_m\} = 0, \quad \forall\, m \in \mathcal{I}_M. \tag{25}$$

Therefore, using the fact that $\hat{A}\hat{B} = 0$ holds for any $\hat{A},\hat{B} \in \mathcal{S}_+$ satisfying $\mathrm{Tr}(\hat{A}\hat{B}) = 0$ yields Eq. (20). From $L(\Pi,\sigma,\hat{X},\lambda) = f(\Pi)$, the fourth term of the right-hand side of Eq. (14) must be zero. Therefore, Eq. (21) holds.

Next, we show (2) ⇒ (3). From Eq. (20), $\hat{X}\hat{\Pi}_m = \hat{z}_m(\lambda)\hat{\Pi}_m$ holds. Summing this equation over $m = 0, \ldots, M-1$ yields $\hat{X} = \sum_{m=0}^{M-1} \hat{z}_m(\lambda)\hat{\Pi}_m$, which gives Eq. (22). Equation (23) obviously holds from Eq. (21).

Finally, we show (3) ⇒ (1). Let $\hat{X} = \sum_{m=0}^{M-1} \hat{z}_m(\lambda)\hat{\Pi}_m$. We have that for any POVM $\Pi' = \{\hat{\Pi}'_m : m \in \mathcal{I}_M\} \in \mathcal{M}^\circ$,

$$\begin{aligned} f(\Pi) - f(\Pi') &\geqslant f(\Pi) + \sum_{j=0}^{J-1} \lambda_j \left[b_j - \sum_{m=0}^{M-1} \mathrm{Tr}(\hat{a}_{j,m}\hat{\Pi}_m) \right] \\ &\quad - f(\Pi') - \sum_{j=0}^{J-1} \lambda_j \left[b_j - \sum_{m=0}^{M-1} \mathrm{Tr}(\hat{a}_{j,m}\hat{\Pi}'_m) \right] \\ &= \mathrm{Tr}\hat{X} - \sum_{m=0}^{M-1} \mathrm{Tr}[\hat{z}_m(\lambda)\hat{\Pi}'_m] \\ &= \sum_{m=0}^{M-1} \mathrm{Tr}\{[\hat{X} - \hat{z}_m(\lambda)]\hat{\Pi}'_m\} \geqslant 0, \end{aligned} \tag{26}$$

where the first inequality follows from Eq. (23) and $\sum_{m=0}^{M-1} \mathrm{Tr}(\hat{a}_{j,m}\hat{\Pi}'_m) \leqslant b_j$. The last inequality follows from Eq. (22); i.e., $\hat{X} \geqslant \hat{z}_m(\lambda)$. Since $f(\Pi) \geqslant f(\Pi')$ holds for any POVM $\Pi' \in \mathcal{M}^\circ$, Π is an optimal measurement of problem (2). ■

Let $\Pi^\star = \{\hat{\Pi}^\star_m : m \in \mathcal{I}_M\}$ and $(\hat{X}^\star, \lambda^\star)$ be, respectively, optimal solutions to primal problem (2) and dual problem (12). From Eq. (20), the support of $\hat{\Pi}^\star_m$ is included in the kernel of $\hat{X}^\star - \hat{z}_m(\lambda^\star)$ for any $m \in \mathcal{I}_M$. In particular, if the supports of the operators $\{\hat{z}_m(\lambda^\star) : m \in \mathcal{I}_M\}$ span $\mathcal{H}$ (which holds in many cases of interest), then $\mathrm{rank}\hat{X}^\star = \dim \mathcal{H}$ holds, which gives $\mathrm{rank}\hat{\Pi}^\star_m \leqslant \mathrm{rank}\hat{z}_m(\lambda^\star)$ from Eq. (20). Statement (3) can be more readily used to verify whether a POVM Π is optimal than statement (2); we only need to check whether $\lambda \in \mathcal{R}^J_+$ exists such that Eqs. (22) and (23) hold, which would be easy if J is sufficiently small.

III. GENERALIZED MINIMAX SOLUTION

A. Formulation

In this section, we consider the quantum minimax strategy, which provides a different type of problem from those discussed in the previous section. The quantum minimax strategy has been investigated [10–12,22,23] under the assumption that the collection of prior probabilities is not given. We investigate the minimax strategy for a generalized quantum state discrimination problem.

Let K be a positive integer. Also, let $\mathcal{P}$ be the entire set of collections of K non-negative real numbers, $\mu = \{\mu_k \geqslant 0 : k \in \mathcal{I}_K\}$, satisfying $\sum_{k=0}^{K-1} \mu_k = 1$, which implies that $\mu \in \mathcal{P}$ can be interpreted as a probability distribution. We consider a function $F(\mu,\Pi)$,

$$\begin{aligned} F(\mu,\Pi) &= \sum_{k=0}^{K-1} \mu_k f_k(\Pi), \\ f_k(\Pi) &= \sum_{m=0}^{M-1} \mathrm{Tr}(\hat{c}_{k,m}\hat{\Pi}_m) + d_k, \end{aligned} \tag{27}$$

where $\hat{c}_{k,m} \in \mathcal{S}$ and $d_k \in \mathcal{R}$ hold for any $m \in \mathcal{I}_M$ and $k \in \mathcal{I}_K$. We want to find a POVM $\Pi \in \mathcal{M}^\circ$ that maximizes the worst-case value of $F(\mu,\Pi)$ over $\mu \in \mathcal{P}$, i.e., $\min_{\mu\in\mathcal{P}} F(\mu,\Pi)$, where $\mathcal{M}^\circ$ is defined by Eq. (3). In the case of $K = 1$, this problem is equivalent to problem (2) with $\hat{c}_m = \hat{c}_{0,m}$ and $d_0 = 0$. Therefore, this problem can be regarded as an extension of problem (2).

We can see that if $\mathcal{M}^\circ$ is not empty, then the so-called minimax theorem holds; that is, there exists $(\mu^\star,\Pi^\star)$ satisfying the following equations:

$$\max_{\Pi\in\mathcal{M}^\circ} \min_{\mu\in\mathcal{P}} F(\mu,\Pi) = F(\mu^\star,\Pi^\star) = \min_{\mu\in\mathcal{P}} \max_{\Pi\in\mathcal{M}^\circ} F(\mu,\Pi). \tag{28}$$

Indeed, $\mathcal{M}^\circ$ and $\mathcal{P}$ are closed convex sets, and $F(\mu,\Pi)$ is a continuous convex function of μ for fixed Π and a continuous concave function of Π for fixed μ, which are sufficient conditions for the minimax theorem to hold [37]. We call $(\mu^\star,\Pi^\star)$, $\mu^\star$, and $\Pi^\star$, respectively, a minimax solution, minimax probabilities, and a minimax measurement. $(\mu^\star,\Pi^\star)$ is a minimax solution if and only if $(\mu^\star,\Pi^\star)$ is a saddle point of $F(\mu,\Pi)$, i.e., the following inequalities hold for any $\mu \in \mathcal{P}$ and $\Pi \in \mathcal{M}^\circ$ [37]:

$$F(\mu^\star,\Pi) \leqslant F(\mu^\star,\Pi^\star) \leqslant F(\mu,\Pi^\star). \tag{29}$$

Let

$$F^\star(\mu) = \max_{\Pi\in\mathcal{M}^\circ} F(\mu,\Pi), \tag{30}$$

with $\mu \in \mathcal{P}$. It follows from Eq. (29) that $F^\star(\mu^\star) = F(\mu^\star,\Pi^\star)$ holds. From Eq. (27), $F(\mu,\Pi)$ can be expressed by

$$\begin{aligned} F(\mu,\Pi) &= \sum_{k=0}^{K-1} \mu_k \left[\sum_{m=0}^{M-1} \mathrm{Tr}(\hat{c}_{k,m}\hat{\Pi}_m) + d_k \right] \\ &= \sum_{m=0}^{M-1} \mathrm{Tr}\left[\left(\sum_{k=0}^{K-1} \mu_k \hat{c}_{k,m} \right) \hat{\Pi}_m \right] + \sum_{k=0}^{K-1} \mu_k d_k. \end{aligned} \tag{31}$$

Thus, $F^\star(\mu)$ for a given $\mu \in \mathcal{P}$ can be obtained by finding $\Pi \in \mathcal{M}^\circ$ that maximizes the first term of the second line of Eq. (31), which is formulated as problem (2) with $c_m = \sum_{k=0}^{K-1} \mu_k \hat{c}_{k,m}$.

B. Examples

We give some examples of minimax problems that can be formulated as Eq. (27). Let us consider discrimination between R quantum states $\{\hat{\rho}_r : r \in \mathcal{I}_R\}$.

1. Minimax solution in the Bayes strategy

The minimax strategy in which the average Bayes cost is used as the objective function has been investigated in Ref. [12]. We regard $\mu \in \mathcal{P}$ with $K = R$ as prior probabilities of the states $\{\hat{\rho}_r : r \in \mathcal{I}_R\}$. The aim of this problem is to find a POVM Π that minimizes the worst-case average Bayes cost $B(\mu,\Pi)$ over $\mu \in \mathcal{P}$. $B(\mu,\Pi)$ is expressed by

$$\begin{aligned} B(\mu,\Pi) &= \sum_{m=0}^{R-1} \mathrm{Tr}[\hat{W}_m(\mu)\hat{\Pi}_m], \\ \hat{W}_m(\mu) &= \sum_{k=0}^{R-1} \mu_k B_{m,k}\hat{\rho}_k, \end{aligned} \tag{32}$$

where $B_{m,k} \in \mathcal{R}_+$ holds for any $m,k \in \mathcal{I}_R$. This problem can be expressed by a form of Eq. (27) with $F(\mu,\Pi) = -B(\mu,\Pi)$. In this case, we have

$$\begin{aligned} f_k(\Pi) &= -\sum_{m=0}^{R-1} \mathrm{Tr}[(B_{m,k}\hat{\rho}_k)\hat{\Pi}_m], \quad \forall\, k \in \mathcal{I}_R, \\ \mathcal{M}^\circ &= \mathcal{M}; \end{aligned} \tag{33}$$

i.e.,

$$M = K = R, \quad J = 0, \quad \hat{c}_{k,m} = -B_{m,k}\hat{\rho}_k, \quad d_k = 0. \tag{34}$$

2. Inconclusive minimax solution

The application to the minimax strategy to state discrimination that allows a nonzero inconclusive probability has been investigated in Ref. [23]. The aim of this problem is to find a POVM Π, which we call an inconclusive minimax measurement, which maximizes the worst-case value of the sum of the average correct and inconclusive probabilities under the constraint that $\mathrm{Tr}(\hat{\rho}_j\hat{\Pi}_R)$ is not greater than a given value

p, with $0 \leqslant p \leqslant 1$ for any $j \in \mathcal{I}_R$. In particular, if $p = 0$, then an inconclusive minimax measurement is a standard minimax measurement without inconclusive results [10]. Let $K = R$ and $\mu \in \mathcal{P}$ be prior probabilities of the states $\{\hat{\rho}_r : r \in \mathcal{I}_R\}$; then, this problem can be expressed by a form of Eq. (27) with

$$f_k(\Pi) = \mathrm{Tr}[\hat{\rho}_k(\hat{\Pi}_k + \hat{\Pi}_R)], \quad \forall\, k \in \mathcal{I}_R,$$
$$\mathcal{M}^\circ = \{\Pi \in \mathcal{M} : \mathrm{Tr}(\hat{\rho}_j \hat{\Pi}_R) \leqslant p, \forall\, j \in \mathcal{I}_R\}. \tag{35}$$

That is, we have

$$\begin{aligned} M &= R + 1,\\ K &= J = R,\\ \hat{c}_{k,m} &= \begin{cases} \hat{\rho}_k, & m = k \text{ or } m = R,\\ 0, & \text{otherwise},\end{cases}\\ d_k &= 0,\\ \hat{a}_{j,m} &= \delta_{m,R}\hat{\rho}_j,\\ b_j &= p. \end{aligned} \tag{36}$$

3. Minimax solution for plural state sets

We consider a quantum measurement that maximizes the worst-case average correct probabilities for plural quantum state sets $\{\Psi_k : k \in \mathcal{I}_K\}$ with $K \geqslant 2$ as another example, where, for each $k \in \mathcal{I}_K$, Ψ_k is a set of R quantum states, $\Psi_k = \{\hat{\rho}_{k,r} : r \in \mathcal{I}_R\}$, with prior probabilities $\{\xi_{k,r} : r \in \mathcal{I}_R\}$. This problem can be expressed by a form of Eq. (27) with

$$f_k(\Pi) = \sum_{m=0}^{R-1} \mathrm{Tr}(\hat{\rho}'_{k,m}\hat{\Pi}_m),\ \forall\, k \in \mathcal{I}_K,$$
$$\mathcal{M}^\circ = \mathcal{M}, \tag{37}$$

where $\hat{\rho}'_{k,r} = \xi_{k,r}\hat{\rho}_{k,r}$. That is, we have

$$M = R, \quad J = 0, \quad \hat{c}_{k,m} = \hat{\rho}'_{k,m}, \quad d_k = 0. \tag{38}$$

We discuss this problem in detail in Sec. V B.

C. Properties of a minimax solution

We show necessary and sufficient conditions for a minimax solution in Theorem 3 and an optimization problem of obtaining a minimax measurement in Theorem 4.

Theorem 3. Suppose that $\mu^\star \in \mathcal{P}$ and $\Pi^\star \in \mathcal{M}^\circ$ hold. The following statements are all equivalent.

(1) $(\mu^\star,\Pi^\star)$ is a minimax solution to Eq. (27).

(2) We have that for any $k \in \mathcal{I}_K$,

$$f_k(\Pi^\star) \geqslant F^\star(\mu^\star). \tag{39}$$

(3) We have that for any $k,k' \in \mathcal{I}_K$ such that $\mu^\star_{k'} > 0$,

$$f_k(\Pi^\star) \geqslant f_{k'}(\Pi^\star). \tag{40}$$

Proof. It suffices to show (1) ⇔ (2) and (2) ⇔ (3).

First, we show (1) ⇒ (2). Let $\mu^{(k)} = \{\mu_{k'} = \delta_{k,k'} : k' \in \mathcal{I}_K\}$. From Eq. (29) and $F^\star(\mu^\star) = F(\mu^\star,\Pi^\star)$, we have that for any $k \in \mathcal{I}_K$,

$$f_k(\Pi^\star) = F(\mu^{(k)},\Pi^\star) \geqslant F(\mu^\star,\Pi^\star) = F^\star(\mu^\star). \tag{41}$$

Thus, Eq. (39) holds.

Next we show (2) ⇒ (1). From Eqs. (30) and (39), We obtain, for any $\mu \in \mathcal{P}$ and $\Pi \in \mathcal{M}^\circ$,

$$F(\mu^\star,\Pi) \leqslant F^\star(\mu^\star) \leqslant \sum_{k=0}^{K-1} \mu_k f_k(\Pi^\star) = F(\mu,\Pi^\star). \tag{42}$$

Substituting $\mu = \mu^\star$ and $\Pi = \Pi^\star$ into this equation gives $F^\star(\mu^\star) = F(\mu^\star,\Pi^\star)$. Thus, from Eq. (42), Eq. (29) holds, which means that $(\mu^\star,\Pi^\star)$ is a minimax solution to Eq. (27).

Then, we show (2) ⇒ (3). From Eq. (39) and the definition of $F^\star(\mu)$, $F(\mu^\star,\Pi^\star) = F^\star(\mu^\star)$ must hold. Thus, we have

$$\begin{aligned} f_k(\Pi^\star) &= F^\star(\mu^\star), \quad \forall\, k \in \mathcal{I}_K \text{ such that } \mu^\star_k > 0,\\ f_k(\Pi^\star) &\geqslant F^\star(\mu^\star), \quad \forall\, k \in \mathcal{I}_K \text{ such that } \mu^\star_k = 0, \end{aligned} \tag{43}$$

from which we can easily see that Eq. (40) holds.

Finally, we show (3) ⇒ (2). From Eq. (40), $f_k(\Pi^\star) = f_{k'}(\Pi^\star)$ holds for any $k,k' \in \mathcal{I}_K$ satisfying $\mu^\star_k > 0$ and $\mu^\star_{k'} > 0$. Thus, according to the definition of $F^\star(\mu)$, $F^\star(\mu^\star) = f_{k'}(\Pi^\star)$ holds for any $k' \in \mathcal{I}_K$ satisfying $\mu^\star_{k'} > 0$. Substituting this into Eq. (40) gives Eq. (39). ■

Theorem 4. Let us consider the following optimization problem

$$\begin{aligned} &\text{maximize} && f_{\min}(\Pi) = \min_{k\in\mathcal{I}_K} f_k(\Pi)\\ &\text{subject to} && \Pi \in \mathcal{M}^\circ, \end{aligned} \tag{44}$$

with a POVM Π. A POVM Π^+ is an optimal solution to problem (44) if and only if Π^+ is a minimax measurement of Eq. (27).

Proof. Suppose that Π^+ is an optimal solution to problem (44) and that $(\mu^\star,\Pi^\star)$ is a minimax solution to Eq. (27). Equations (30) and (39) give

$$f_{\min}(\Pi^\star) \geqslant F^\star(\mu^\star) = \max_{\Pi\in\mathcal{M}^\circ} F(\mu^\star,\Pi) \geqslant \max_{\Pi\in\mathcal{M}^\circ} f_{\min}(\Pi), \tag{45}$$

which indicates that $\Pi^\star$ is an optimal solution to problem (44). Since Π^+ is also an optimal solution to problem (44), $f_{\min}(\Pi^+) = f_{\min}(\Pi^\star) \geqslant F^\star(\mu^\star)$ holds from Eq. (45), and thus statement (2) of Theorem 3 holds. Therefore, Π^+ is a minimax measurement of Eq. (27). ■

IV. GROUP COVARIANT OPTIMIZATION PROBLEM

In this section, we show that if an optimization problem of obtaining an optimal measurement or a minimax solution has a certain symmetry, the optimal solution also has the same symmetry. A quantum state set that is invariant under the action of a group $\mathcal{G}$ in which each element corresponds to a unitary or antiunitary operator is called a group-covariant (or $\mathcal{G}$-covariant) state set. Similarly, we call an optimal measurement and a minimax solution that are invariant under the same action a group-covariant (or $\mathcal{G}$-covariant) optimal measurement and a minimax solution, respectively. Optimal measurements for group-covariant state sets have been well investigated, and it has been derived that a $\mathcal{G}$-covariant optimal measurement exists for a $\mathcal{G}$-covariant state set under several optimality criteria [5–8,17,23,36,38,39]. These results not only help us to obtain analytical optimal solutions (e.g., [40–42]), but also are useful for developing computationally efficient algorithms for obtaining optimal solutions [43,44]. In this section, we generalize these results to our generalized optimization problems.

A. Group action

First, let us describe a group action. A group action of $\mathcal{G}$ on a set T is a set of mappings from T to T, $\{\pi_g(x)(x \in T) : g \in \mathcal{G}\}$ [we also denote $\pi_g(x)$ as $g \circ x$], such that

(1) for any $g,h \in \mathcal{G}$ and $x \in T$, $(gh) \circ x = g \circ (h \circ x)$ holds.

(2) for any $x \in T$, $e \circ x = x$ holds, where e is the identity element of $\mathcal{G}$.

The action of $\mathcal{G}$ on T is called faithful if, for any distinct $g,h \in \mathcal{G}$, there exists $x \in T$ such that $g \circ x \neq h \circ x$. Here we assume that the number of elements in $\mathcal{G}$, which is denoted as $|\mathcal{G}|$, is greater than 1.

Let us consider an action of $\mathcal{G}$ on the set $\mathcal{I}_N$ with $N \geqslant 1$, that is, $\{g \circ n \in \mathcal{I}_N (n \in \mathcal{I}_N) : g \in \mathcal{G}\}$. This action is not faithful in general. We also consider the action of $\mathcal{G}$ on $\mathcal{S}$, expressed by

$$g \circ \hat{A} = \hat{U}_g \hat{A} \hat{U}_g^\dagger, \tag{46}$$

with $g \in \mathcal{G}$ and $\hat{A} \in \mathcal{S}$, where $\hat{U}_g$ is a unitary or antiunitary operator and $\hat{U}_g^\dagger$ is conjugate transpose of $\hat{U}_g$. (Note that if $\hat{U}_g$ is an antiunitary operator, then $\hat{U}_g^\dagger$ is an antiunitary operator such that $\hat{U}_g^\dagger \hat{U}_g = \hat{U}_g \hat{U}_g^\dagger = \hat{1}$.) $\hat{U}_e = \hat{1}$ and $\hat{U}_{\bar{g}} = \hat{U}_g^\dagger$ obviously hold, where $\bar{g}$ is the inverse element of g. We assume that the action of $\mathcal{G}$ on $\mathcal{S}$ is faithful, which is equivalent to $\hat{U}_g \neq \hat{U}_h$ for any distinct $g,h \in \mathcal{G}$. From Eq. (46), we can easily verify that for any $g \in \mathcal{G}$, $c \in \mathcal{R}$, and $\hat{A},\hat{B} \in \mathcal{S}$, we have

$$\begin{aligned}
g \circ (\hat{A} \pm \hat{B}) &= g \circ \hat{A} \pm g \circ \hat{B},\\
g \circ (c\hat{A}) &= c(g \circ \hat{A}),\\
g \circ \hat{1} &= \hat{1},\\
\mathrm{Tr}(g \circ \hat{A}) &= \mathrm{Tr}\hat{A},\\
\mathrm{Tr}[(g \circ \hat{A})(g \circ \hat{B})] &= \mathrm{Tr}(\hat{A}\hat{B}),\\
g \circ \hat{A} &\in \mathcal{S}_+, \quad \forall\, \hat{A} \in \mathcal{S}_+,\\
g \circ \hat{A} &\geqslant g \circ \hat{B}, \quad \forall\, \hat{A} \geqslant \hat{B}.
\end{aligned} \tag{47}$$

In this section, we use these facts without mentioning them.

B. Group covariant optimal measurement

As a preparation, we first prove the following lemma.

Lemma 1. Suppose that $\mathcal{M}^\circ$ is not empty. Also, suppose that there exist actions of $\mathcal{G}$ on $\mathcal{S}$, $\mathcal{I}_M$, and $\mathcal{I}_J$ such that

$$\begin{aligned}
g \circ \hat{a}_{j,m} &= \hat{a}_{g\circ j, g\circ m}, \quad \forall\, g \in \mathcal{G}, \quad j \in \mathcal{I}_J, \quad m \in \mathcal{I}_M,\\
b_j &= b_{g\circ j}, \quad \forall\, g \in \mathcal{G},\ j \in \mathcal{I}_J.
\end{aligned} \tag{48}$$

Let $\kappa_g(\Phi)$ and $\kappa(\Phi)$ be mappings of $\Phi \in \mathcal{M}^\circ$ expressed by

$$\begin{aligned}
\kappa_g(\Phi) &= \{\bar{g} \circ \hat{\Phi}_{g\circ m} : m \in \mathcal{I}_M\},\\
\kappa(\Phi) &= \left\{ \frac{1}{|\mathcal{G}|} \sum_{g\in\mathcal{G}} \bar{g} \circ \hat{\Phi}_{g\circ m} : m \in \mathcal{I}_M \right\}.
\end{aligned} \tag{49}$$

Then κ_g is a bijective mapping onto $\mathcal{M}^\circ$ for any $g \in \mathcal{G}$, and κ is a mapping onto $\mathcal{M}^\circ$. Moreover, for any $\Phi \in \mathcal{M}^\circ$, we have

$$g \circ \hat{\Pi}_m = \hat{\Pi}_{g\circ m}, \quad \forall\, g \in \mathcal{G}, \quad m \in \mathcal{I}_M, \tag{50}$$

where $\Pi = \kappa(\Phi)$.

Proof. First, we show that κ_g is bijective onto $\mathcal{M}^\circ$. Let $\Phi \in \mathcal{M}^\circ$ and $\Phi^{(g)} = \kappa_g(\Phi)$. Since $\Phi_m^{(g)} = \bar{g} \circ \hat{\Phi}_{g\circ m} \in \mathcal{S}+$ and $\sum_{m=0}^{M-1} \Phi_m^{(g)} = \bar{g} \circ \hat{1} = \hat{1}$ hold, $\Phi^{(g)} \in \mathcal{M}$ holds. We also obtain for any $j \in \mathcal{I}_J$,

$$\begin{aligned}
&\sum_{m=0}^{M-1} \mathrm{Tr}\big(\hat{a}_{j,m} \hat{\Phi}_m^{(g)}\big)\\
&= \sum_{m=0}^{M-1} \mathrm{Tr}[\hat{a}_{j,m}(\bar{g} \circ \hat{\Phi}_{g\circ m})] = \sum_{m=0}^{M-1} \mathrm{Tr}[(g \circ \hat{a}_{j,m})\hat{\Phi}_{g\circ m}]\\
&= \sum_{m=0}^{M-1} \mathrm{Tr}(\hat{a}_{g\circ j, g\circ m} \hat{\Phi}_{g\circ m}) \leqslant b_{g\circ j} = b_j,
\end{aligned} \tag{51}$$

where the inequality follows from the group action being bijective and $\Phi \in \mathcal{M}^\circ$. Thus, $\Phi^{(g)} \in \mathcal{M}^\circ$ holds. Moreover, since $\kappa_{\bar{g}}[\kappa_g(\Phi)] = \kappa_g[\kappa_{\bar{g}}(\Phi)] = \Phi$, $\kappa_{\bar{g}}$ is the inverse mapping of κ_g. Therefore, κ_g is bijective onto $\mathcal{M}^\circ$.

Next, we show that κ is a mapping onto $\mathcal{M}^\circ$ and that Eq. (50) holds. From Eq. (51), we have that for any $j \in \mathcal{I}_J$,

$$\sum_{m=0}^{M-1} \mathrm{Tr}(\hat{a}_{j,m} \hat{\Pi}_m) = \frac{1}{|G|} \sum_{g\in\mathcal{G}} \sum_{m=0}^{M-1} \mathrm{Tr}\big(\hat{a}_{j,m} \hat{\Phi}_m^{(g)}\big) \leqslant b_j, \tag{52}$$

which means that $\Pi \in \mathcal{M}^\circ$ holds for any $\Phi \in \mathcal{M}^\circ$; that is, κ is a mapping onto $\mathcal{M}^\circ$. We also have that for any $g \in \mathcal{G}$ and $m \in \mathcal{I}_M$,

$$g \circ \hat{\Pi}_m = \frac{1}{|\mathcal{G}|} \sum_{h\in\mathcal{G}} g \circ \hat{\Phi}_m^{(h)} = \frac{1}{|\mathcal{G}|} \sum_{h'\in\mathcal{G}} \bar{h}' \circ \hat{\Phi}_{h'\circ g\circ m} = \hat{\Pi}_{g\circ m}, \tag{53}$$

where $h' = h \circ \bar{g}$. Thus, Eq. (50) holds. ■

We now show that a $\mathcal{G}$-covariant optimal measurement exists if optimization problem (2) has a certain symmetry with respect to $\mathcal{G}$.

Theorem 5. Let us consider optimization problem (2). Suppose that $\mathcal{M}^\circ$ is not empty. Also, suppose that there exist actions of $\mathcal{G}$ on $\mathcal{S}$, $\mathcal{I}_M$, and $\mathcal{I}_J$ satisfying Eq. (48) and

$$g \circ \hat{c}_m = \hat{c}_{g\circ m}, \quad \forall\, g \in \mathcal{G}, \quad m \in \mathcal{I}_M. \tag{54}$$

Then, for any $\Phi \in \mathcal{M}^\circ$ there exists $\Pi \in \mathcal{M}^\circ$ such that $f(\Pi) = f(\Phi)$ and Eq. (50) hold, where f is the objective function of problem (2). In particular, an optimal measurement Π exists satisfying Eq. (50). Moreover, there exists an optimal solution $(\hat{X},\lambda)$ to dual problem (12) such that

$$\begin{aligned}
g \circ \hat{X} &= \hat{X}, \quad \forall\, g \in \mathcal{G},\\
\lambda_j &= \lambda_{g\circ j}, \quad \forall\, g \in \mathcal{G},\ j \in \mathcal{I}_J.
\end{aligned} \tag{55}$$

As examples of Theorem 5, we can derive that there exist a minimum error measurement, an optimal unambiguous measurement, and an optimal inconclusive measurement that are $\mathcal{G}$ covariant if a given state set is $\mathcal{G}$ covariant, which is shown in Ref. [36].

If we let $\mathcal{M}^\circ_\mathcal{G}$ be the entire set of $\Pi \in \mathcal{M}^\circ$ satisfying Eq. (50), then we can easily see that, since $\mathcal{M}^\circ_\mathcal{G}$ is convex, problem (2) remains in convex programming even if we restrict the feasible set from $\mathcal{M}^\circ$ to $\mathcal{M}^\circ_\mathcal{G}$.

Proof. First, we show that $\Pi \in \mathcal{M}^\circ$ exists such that $f(\Pi) = f(\Phi)$ and Eq. (50) hold for any $\Phi \in \mathcal{M}^\circ$. Let $\Pi = \kappa(\Phi)$, where κ is defined by Eq. (49). From Lemma 1, Π satisfies $\Pi \in \mathcal{M}^\circ$ and Eq. (50). Moreover, we obtain

$$\begin{aligned} f(\Pi) &= \sum_{m=0}^{M-1} \mathrm{Tr}(\hat{c}_m \hat{\Pi}_m) = \frac{1}{|\mathcal{G}|}\sum_{m=0}^{M-1}\sum_{g\in\mathcal{G}} \mathrm{Tr}[\hat{c}_m(\bar{g}\circ\hat{\Phi}_{g\circ m})] \\ &= \frac{1}{|\mathcal{G}|}\sum_{m=0}^{M-1}\sum_{g\in\mathcal{G}} \mathrm{Tr}[(g\circ\hat{c}_m)\hat{\Phi}_{g\circ m}] \\ &= \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}}\sum_{m=0}^{M-1} \mathrm{Tr}(\hat{c}_{g\circ m}\hat{\Phi}_{g\circ m}) \\ &= \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}} f(\Phi) = f(\Phi). \end{aligned} \tag{56}$$

In particular, if Φ is an optimal measurement, then so is Π.

Next, we show that there exists an optimal solution $(\hat{X},\lambda)$ to dual problem (12) satisfying Eq. (55). Let $\nu = \{\nu_j : j \in \mathcal{I}_J\} \in \mathcal{R}_+^J$. Suppose that $(\hat{Y},\nu)$ is an optimal solution to problem (12). Also, let $\hat{Y}^{(g)} = g\circ\hat{Y}$ and $\nu^{(g)} = \{\nu_j^{(g)} = \nu_{\bar{g}\circ j} : j \in \mathcal{I}_J\}$. $\hat{Y}^{(g)} \in \mathcal{S}$ and $\nu^{(g)} \in \mathcal{R}_+^J$ obviously hold. We obtain for any $g \in \mathcal{G}$ and $m \in \mathcal{I}_M$

$$\begin{aligned} \hat{Y}^{(g)} &\geqslant g\circ\hat{z}_m(\nu) = \hat{c}_{g\circ m} - \sum_{j=0}^{J-1}\nu_j \hat{a}_{g\circ j, g\circ m} \\ &= \hat{c}_{g\circ m} - \sum_{j=0}^{J-1}\nu^{(g)}_{g\circ j}\hat{a}_{g\circ j, g\circ m} = \hat{z}_{g\circ m}(\nu^{(g)}). \end{aligned} \tag{57}$$

We also obtain

$$\begin{aligned} s(\hat{Y}^{(g)},\nu^{(g)}) &= \mathrm{Tr}\hat{Y}^{(g)} + \sum_{j=0}^{J-1}\nu_j^{(g)} b_j \\ &= \mathrm{Tr}\hat{Y} + \sum_{j=0}^{J-1}\nu_{\bar{g}\circ j} b_{\bar{g}\circ j} = s(\hat{Y},\nu). \end{aligned} \tag{58}$$

From Eqs. (57) and (58), $(\hat{Y}^{(g)},\nu^{(g)})$ is also an optimal solution to problem (12). Let $\hat{X} \in \mathcal{S}$ and $\lambda = \{\lambda_j : j \in \mathcal{I}_J\} \in \mathcal{R}_+^J$ be expressed by

$$\hat{X} = \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}}\hat{Y}^{(g)}, \quad \lambda_j = \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}}\nu_j^{(g)}. \tag{59}$$

We can easily see that Eq. (55) holds. For any $m \in \mathcal{I}_M$, we have

$$\begin{aligned} \hat{z}_m(\lambda) &= \hat{c}_m - \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}}\sum_{j=0}^{J-1}\nu_j^{(g)}\hat{a}_{j,m} \\ &= \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}}\left(\hat{c}_m - \sum_{j=0}^{J-1}\nu_j^{(g)}\hat{a}_{j,m}\right) = \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}}\hat{z}_m(\nu^{(g)}). \end{aligned} \tag{60}$$

From Eqs. (57), (59), and (60), we obtain for any $m \in \mathcal{I}_M$,

$$\hat{X} - \hat{z}_m(\lambda) = \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}}[\hat{Y}^{(g)} - \hat{z}_m(\nu^{(g)})] \geqslant 0. \tag{61}$$

Moreover, from Eqs. (58) and (59), we have

$$\begin{aligned} s(\hat{X},\lambda) &= \mathrm{Tr}\hat{X} + \sum_{j=0}^{J-1}\lambda_j b_j \\ &= \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}}(\mathrm{Tr}\hat{Y}^{(b)} + \nu_j^{(g)} b_j) = s(\hat{Y},\nu). \end{aligned} \tag{62}$$

Therefore, $(\hat{X},\nu)$ is also an optimal solution to problem (12). ■

C. Group covariant minimax solution

Similar to Theorem 5, we can show that if Eq. (27) has a certain symmetry with respect to $\mathcal{G}$, then there exists a $\mathcal{G}$-covariant minimax solution.

Theorem 6. Let us consider a minimax solution to Eq. (27). Suppose that $\mathcal{M}^\circ$ is not empty. Also, suppose that there exist actions of $\mathcal{G}$ on $\mathcal{S}$, $\mathcal{I}_M$, $\mathcal{I}_J$, and $\mathcal{I}_K$ satisfying Eq. (48) and

$$\begin{aligned} g\circ\hat{c}_{k,m} &= \hat{c}_{g\circ k, g\circ m}, \quad \forall\, g \in \mathcal{G},\, k \in \mathcal{I}_K, \quad m \in \mathcal{I}_M, \\ d_k &= d_{g\circ k}, \quad \forall\, g \in \mathcal{G}, \quad k \in \mathcal{I}_K. \end{aligned} \tag{63}$$

Then, a minimax solution (μ,Π) exists such that

$$\begin{aligned} \mu_k &= \mu_{g\circ k}, \quad \forall\, g \in \mathcal{G},\, k \in \mathcal{I}_K, \\ g\circ\hat{\Pi}_m &= \hat{\Pi}_{g\circ m}, \quad \forall\, g \in \mathcal{G},\, m \in \mathcal{I}_M. \end{aligned} \tag{64}$$

Proof. Let $(\eta^\star,\Pi^\star)$ be a minimax solution to Eq. (27). Also, let $\mu = \{\mu_k = |\mathcal{G}|^{-1}\sum_{g\in\mathcal{G}}\eta^\star_{g\circ k} : k \in \mathcal{I}_K\}$ and $\Pi = \kappa(\Pi^\star)$, where κ is defined by Eq. (49). Then it follows that $\mu \in \mathcal{P}$, $\Pi \in \mathcal{M}^\circ$, and Eq. (64) hold (also see Lemma 1). Here we show that (μ,Π) is a minimax solution to Eq. (27). From statement (2) of Theorem 3, it suffices to show that $f_k(\Pi) \geqslant F^\star(\mu)$ holds for any $k \in \mathcal{I}_K$. We show $f_k(\Pi) \geqslant F^\star(\eta^\star)$ and $F^\star(\eta^\star) \geqslant F^\star(\mu)$.

First, we show $f_k(\Pi) \geqslant F^\star(\eta^\star)$ for any $k \in \mathcal{I}_K$. Let $\Pi^{(g)} = \kappa_g(\Pi^\star)$; then for any $k \in \mathcal{I}_K$, we have

$$\begin{aligned} f_k(\Pi) &= \frac{1}{|\mathcal{G}|}\sum_{m=0}^{M-1}\sum_{g\in\mathcal{G}}\mathrm{Tr}\big(\hat{c}_{k,m}\hat{\Pi}_m^{(g)}\big) + d_k \\ &= \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}}\left[\sum_{m=0}^{M-1}\mathrm{Tr}\big(\hat{c}_{k,m}\hat{\Pi}_m^{(g)}\big) + d_k\right] \\ &= \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}}\left\{\sum_{m=0}^{M-1}\mathrm{Tr}[(g\circ\hat{c}_{k,m})\hat{\Pi}^\star_{g\circ m}] + d_k\right\} \\ &= \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}}\left[\sum_{m'=0}^{M-1}\mathrm{Tr}(\hat{c}_{g\circ k,m'}\hat{\Pi}^\star_{m'}) + d_k\right] \\ &= \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}} f_{g\circ k}(\Pi^\star) \geqslant F^\star(\eta^\star), \end{aligned} \tag{65}$$

where $m' = g\circ m$. The inequality in the last line follows from $f_k(\Pi^\star) \geqslant F^\star(\eta^\star)$ for any $k \in \mathcal{I}_K$, which is obtained from Theorem 3.

Next, we show $F^\star(\eta^\star) \geqslant F^\star(\mu)$. Let $\eta^{(g)} = \{\eta^\star_{g\circ k} : k \in \mathcal{I}_K\}$. We have that for any $g \in \mathcal{G}$,

$$\begin{aligned} F^\star(\eta^{(g)}) &= \max_{\Phi\in\mathcal{M}^\circ} \sum_{k=0}^{K-1} \eta^\star_{g\circ k} \left[\sum_{m=0}^{M-1} \mathrm{Tr}(\hat{c}_{k,m}\hat{\Phi}_m) + d_k\right] \\ &= \max_{\Phi\in\mathcal{M}^\circ} \sum_{k'=0}^{K-1} \eta^\star_{k'} \left\{\sum_{m=0}^{M-1} \mathrm{Tr}[\hat{c}_{k',m'}(g\circ\hat{\Phi}_m)] + d_{k'}\right\} \\ &= \max_{\Phi'\in\mathcal{M}^\circ} \sum_{k'=0}^{K-1} \eta^\star_{k'} \left[\sum_{m'=0}^{M-1} \mathrm{Tr}(\hat{c}_{k',m'}\hat{\Phi}'_{m'}) + d_{k'}\right] \\ &= F^\star(\eta^\star), \end{aligned} \tag{66}$$

where $k' = g\circ k$, $m' = g\circ m$, and $\Phi' = \kappa_{\bar{g}}(\Phi)$. The third line follows from the mapping $\kappa_{\bar{g}}$ being bijective onto $\mathcal{M}^\circ$ (see Lemma 1). From Eq. (66), we obtain

$$\begin{aligned} F^\star(\mu) &= \max_{\Phi\in\mathcal{M}^\circ} \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}}\sum_{k=0}^{K-1} \eta_k^{(g)} f_k(\Phi) \\ &\leqslant \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}} F^\star(\eta^{(g)}) = F^\star(\eta^\star). \end{aligned} \tag{67}$$

Therefore, (μ,Π) is a minimax solution. ■

V. EXAMPLES OF OPTIMAL MEASUREMENT AND MINIMAX SOLUTION

As an example of a generalized optimal measurement, we discuss the problem of finding an optimal inconclusive measurement with a lower bound on correct probabilities, which is introduced in Sec. II B 3. Also, as an example of a generalized minimax solution, we discuss the problem of finding a minimax solution for plural state sets, which is introduced in Sec. III B 3. Moreover, Tables I and II summarize the problem formulations and their examples shown in Secs. II B and III B, respectively.

A. Optimal inconclusive measurement with a lower bound on correct probabilities

In this example, we can apply Theorems 1 and 2. Substituting Eq. (11) into Eq. (13) gives

$$\hat{z}_m(\lambda) = \begin{cases} (\xi_m + \lambda_m)\hat{\rho}_m, & m < R, \\ \lambda_R \hat{G}, & m = R. \end{cases} \tag{68}$$

Thus, from Theorem 1, dual problem (12) can be rewritten as

$$\begin{aligned} &\text{minimize} \;\; s(\hat{X},\lambda) = \mathrm{Tr}\hat{X} - q\sum_{r=0}^{R-1}\lambda_r - p\lambda_R \\ &\text{subject to} \;\; \hat{X} \geqslant (\xi_r+\lambda_r)\hat{\rho}_r, \quad \forall\, r\in\mathcal{I}_R, \\ &\qquad\qquad\;\; \hat{X} \geqslant \lambda_R\hat{G}. \end{aligned} \tag{69}$$

$\lambda_R\hat{G} \in \mathcal{S}_+$ yields $\hat{X} \in \mathcal{S}_+$. In particular, in the case of $q = 0$, Eq. (69) is equivalent to the dual problem of finding an optimal inconclusive measurement, which is shown in Theorem 1 of Ref. [17].

We can also obtain necessary and sufficient conditions for an optimal measurement from Theorem 2. For example, from statement (3) of this theorem, $\Pi \in \mathcal{M}^\circ$ is an optimal measurement of problem (10) if and only if $\lambda \in \mathcal{R}_+^J$ exists such that

$$\begin{aligned} &\hat{X}(\lambda) - (\xi_r + \lambda_r)\hat{\rho}_r \geqslant 0, \; \forall\, r \in \mathcal{I}_R, \\ &\hat{X}(\lambda) - \lambda_R\hat{G} \geqslant 0, \\ &\lambda_r[\mathrm{Tr}(\hat{\rho}_r\hat{\Pi}_r) - q] = 0, \;\; \forall\, r\in\mathcal{I}_R, \\ &\lambda_R[\mathrm{Tr}(\hat{G}\hat{\Pi}_R) - p] = 0, \end{aligned} \tag{70}$$

where

$$\hat{X}(\lambda) = \sum_{r=0}^{R-1}(\xi_r+\lambda_r)\hat{\rho}_r\hat{\Pi}_r + \lambda_R\hat{G}\hat{\Pi}_R. \tag{71}$$

In the case in which problem (10) has a certain symmetry, we can apply Theorem 5 . Suppose that a given state set is $\mathcal{G}$ covariant, that is, there exist actions of $\mathcal{G}$ on $\mathcal{S}$ and $\mathcal{I}_R$ satisfying $g\circ(\xi_r\hat{\rho}_r) = \xi_{g\circ r}\hat{\rho}_{g\circ r}$, which is equivalent to $g\circ\hat{\rho}_r = \hat{\rho}_{g\circ r}$ and $\xi_r = \xi_{g\circ r}$, for any $g\in\mathcal{G}$ and $r\in\mathcal{I}_R$. Let the action of $\mathcal{G}$ on $\mathcal{I}_M = \mathcal{I}_{R+1}$, $g\circ m$ $(m\in\mathcal{I}_M)$, be $g\circ m = \pi_g^{(R)}(m)$ for any $m\in\mathcal{I}_R$ and $g\circ R = R$, where $\{\pi_g^{(R)} : g\in\mathcal{G}\}$ is the action of $\mathcal{G}$ on $\mathcal{I}_R$. Also, let the action of $\mathcal{G}$ on $\mathcal{I}_J$ be the same as the action of $\mathcal{G}$ on $\mathcal{I}_M$. Then, since Eqs. (48) and (54) hold, there exists an optimal measurement satisfying Eq. (50).

B. Minimax solution for plural state sets

In this example, we can apply Theorem 3, that is, $(\mu^\star,\Pi^\star)$ is a minimax solution if and only if Eq. (39), or Eq. (40), holds. Substituting Eq. (37) into Eq. (40) gives

$$\begin{aligned} &\sum_{m=0}^{R-1}\mathrm{Tr}(\hat{\rho}'_{k,m}\hat{\Pi}^\star_m) \geqslant \sum_{m=0}^{R-1}\mathrm{Tr}(\hat{\rho}'_{k',m}\hat{\Pi}^\star_m), \\ &\forall\, k,k' \in\mathcal{I}_K \text{ such that } \mu^\star_{k'} > 0. \end{aligned} \tag{72}$$

From Eq. (38), $F(\mu,\Pi)$ is expressed by

$$\begin{aligned} F(\mu,\Pi) &= \sum_{k=0}^{K-1}\mu_k\sum_{m=0}^{R-1}\mathrm{Tr}(\hat{\rho}'_{k,m}\hat{\Pi}_m) \\ &= \sum_{m=0}^{R-1}\mathrm{Tr}\left[\left(\sum_{k=0}^{K-1}\mu_k\hat{\rho}'_{k,m}\right)\hat{\Pi}_m\right]. \end{aligned} \tag{73}$$

Thus, $F^\star(\mu)$ is equivalent to the optimal value of $f(\Pi)$ of optimization problem (2) with

$$M = R, \quad J = 0, \quad \hat{c}_m = \sum_{k=0}^{K-1}\mu_k\hat{\rho}'_{k,m}. \tag{74}$$

This indicates that $F^\star(\mu)$ is also equivalent to the average correct probability of a minimum error measurement for the state set $\{\hat{c}_m/\mathrm{Tr}\hat{c}_m : m\in\mathcal{I}_R\}$ with prior probabilities $\{\mathrm{Tr}\hat{c}_m : m\in\mathcal{I}_R\}$ [note that $\hat{c}_m\in\mathcal{S}_+$ holds from Eq. (74)].

We can also apply Theorem 6 in the case in which given state sets have a certain symmetry. Assume that there exist actions of $\mathcal{G}$ on $\mathcal{S}$, $\mathcal{I}_R$, and $\mathcal{I}_K$ satisfying $g\circ\hat{\rho}'_{k,m} = \hat{\rho}'_{g\circ k,g\circ m}$ for any $g\in\mathcal{G}$, $m\in\mathcal{I}_R$, and $k\in\mathcal{I}_K$. For example, this assumption holds if each state set $\Psi_k = \{\hat{\rho}_{k,m} : m\in\mathcal{I}_R\}$ is $\mathcal{G}$ covariant

TABLE I. Basic formulation of generalized optimal measurements and its examples.

Primal problems	Dual problems	Necessary and sufficient conditions [Statement (3) of Theorem 2]
Basic formulation		
maximize $\sum_{m=0}^{M-1} \mathrm{Tr}(\hat{c}_m \hat{\Pi}_m)$ subject to $\Pi \in \mathcal{M}$, $\sum_{m=0}^{M-1} \mathrm{Tr}(\hat{a}_{j,m} \hat{\Pi}_m) \leqslant b_j, \forall\, j \in \mathcal{I}_J$ (2), (3)	minimize $\mathrm{Tr}\hat{X} + \sum_{j=0}^{J-1} \lambda_j b_j$ subject to $\hat{X} \geqslant \hat{z}_m(\lambda),\ \forall\, m \in \mathcal{I}_M$, where $\hat{z}_m(\lambda) = \hat{c}_m - \sum_{j=0}^{J-1} \lambda_j \hat{a}_{j,m}$ (12), (13)	$\lambda \in \mathcal{R}_+^J$ exists such that $\hat{X}(\lambda) \geqslant \hat{z}_m(\lambda),\ \forall\, m \in \mathcal{I}_M$, $\lambda_j \left[b_j - \sum_{m=0}^{M-1} \mathrm{Tr}(\hat{a}_{j,m} \hat{\Pi}_m) \right] = 0, \forall\, j \in \mathcal{I}_J$, where $\hat{X}(\lambda) = \sum_{n=0}^{M-1} \hat{z}_n(\lambda) \hat{\Pi}_n$ (19), (20), (21)
Example 1: Optimal measurement in the Bayes criterion (Sec. II B 1) [1–3]		
minimize $\sum_{m=0}^{R-1} \mathrm{Tr}(\hat{W}_m \hat{\Pi}_m)$ subject to $\Pi \in \mathcal{M}$ (4)	maximize $\mathrm{Tr}\hat{X}$ subject to $\hat{W}_m \geqslant \hat{X},\ \forall\, m \in \mathcal{I}_R$	$\hat{W}_m \geqslant \sum_{r=0}^{R-1} \hat{W}_r \hat{\Pi}_r,\ \forall\, m \in \mathcal{I}_R$
Example 2: Optimal error margin measurement (Sec. II B 2) [19–21]		
maximize $\sum_{r=0}^{R-1} \xi_r \mathrm{Tr}(\hat{\rho}_r \hat{\Pi}_r)$ subject to $\Pi \in \mathcal{M}$, $\sum_{r=0}^{R-1} \xi_r \mathrm{Tr}[\hat{\rho}_r(\hat{\Pi}_r + \hat{\Pi}_R)] \geqslant 1 - \varepsilon$ (7)	minimize $\mathrm{Tr}\hat{X} - \lambda(1-\varepsilon)$ subject to $\hat{X} \geqslant (1+\lambda)\xi_r \hat{\rho}_r,\ \forall\, r \in \mathcal{I}_R$, $\hat{X} \geqslant \lambda \hat{G}$	$\lambda \in \mathcal{R}_+$ exists such that $\hat{X}(\lambda) \geqslant (1+\lambda)\xi_r \hat{\rho}_r,\ \forall\, r \in \mathcal{I}_R$, $\hat{X}(\lambda) \geqslant \lambda \hat{G}$, $\lambda \left[\sum_{r=0}^{R-1} \xi_r \mathrm{Tr}[\hat{\rho}_r(\hat{\Pi}_r + \hat{\Pi}_R)] - 1 + \varepsilon \right] = 0$, where $\hat{X}(\lambda) = (1+\lambda) \sum_{r=0}^{R-1} \xi_r \hat{\rho}_r \hat{\Pi}_r + \lambda \hat{G} \hat{\Pi}_R$
Example 3: Optimal inconclusive measurement with a lower bound on correct probabilities (Sec. II B 3)		
maximize $\sum_{r=0}^{R-1} \xi_r \mathrm{Tr}(\hat{\rho}_r \hat{\Pi}_r)$ subject to $\Pi \in \mathcal{M}$, $\mathrm{Tr}(\hat{\rho}_r \hat{\Pi}_r) \geqslant q,\ \forall\, r \in \mathcal{I}_R$, $\mathrm{Tr}(\hat{G} \hat{\Pi}_R) \geqslant p$ (10)	minimize $\mathrm{Tr}\hat{X} - q \sum_{r=0}^{R-1} \lambda_r - p\lambda_R$ subject to $\hat{X} \geqslant (\xi_r + \lambda_r)\hat{\rho}_r,\ \forall\, r \in \mathcal{I}_R$, $\hat{X} \geqslant \lambda_R \hat{G}$ (69)	$\lambda \in \mathcal{R}_+^{R+1}$ exists such that $\hat{X} \geqslant (\xi_r + \lambda_r)\hat{\rho}_r,\ \forall\, r \in \mathcal{I}_R$, $\hat{X} \geqslant \lambda_R \hat{G}$, $\lambda_r[\mathrm{Tr}(\hat{\rho}_r \hat{\Pi}_r) - q] = 0,\ \forall\, r \in \mathcal{I}_R$, $\lambda_R[\mathrm{Tr}(\hat{G}\hat{\Pi}_R) - p] = 0$, where $\hat{X}(\lambda) = \sum_{r=0}^{R-1} (\xi_r + \lambda_r)\hat{\rho}_r \hat{\Pi}_r + \lambda_R \hat{G} \hat{\Pi}_R$ (70), (71)

under the same actions of $\mathcal{G}$ on $\mathcal{S}$ and $\mathcal{I}_R$, i.e., $g \circ \hat{\rho}'_{k,m} = \hat{\rho}'_{k,g\circ m}$ for any $g \in \mathcal{G}$ and $m \in \mathcal{I}_R$ (in this case let $g \circ k = k$ for any $k \in \mathcal{I}_K$). Under this assumption, Eq. (63) holds, and thus a $\mathcal{G}$-covariant minimax solution exists.

VI. CONCLUSION

We investigated a generalized optimization problem of finding quantum measurements. Each of the objective and constraint functions in this problem is formulated by the sum of the traces of the multiplication of a Hermitian operator and a detection operator. We first derived corresponding dual problems and necessary and sufficient conditions for an optimal measurement. The minimax version of this problem was also studied, and necessary and sufficient conditions for a minimax solution were provided. We finally showed that for an optimization problem having a certain symmetry with respect to a group in which each element corresponds to a unitary or antiunitary operator, there exists an optimal solution with the same symmetry.

ACKNOWLEDGMENTS

We are grateful to O. Hirota of Tamagawa University for support. T.S.U. was supported (in part) by JSPS KAKENHI (Grant No. 24360151).

TABLE II. Basic formulation of generalized minimax solutions and its examples.

Problems	Necessary and sufficient conditions [Statement (3) of Theorem 3]
Basic formulation maximize $\min_{\mu\in\mathcal{P}} F(\mu,\Pi)$ subject to $\Pi \in \mathcal{M}^\circ$ where $F(\mu,\Pi) = \sum_{k=0}^{K-1} \mu_k f_k(\Pi)$, $f_k(\Pi) = \sum_{m=0}^{M-1} \mathrm{Tr}(\hat{c}_{k,m}\hat{\Pi}_m) + d_k$, $\mathcal{M}^\circ = \left\{\Pi \in \mathcal{M} : \sum_{m=0}^{M-1} \mathrm{Tr}(\hat{a}_{j,m}\hat{\Pi}_m) \leqslant b_j, \forall\, j \in \mathcal{I}_J\right\}$ (3), (27), (28)	$f_k(\Pi^\star) \geqslant f_{k'}(\Pi^\star), \forall\, k,k' \in \mathcal{I}_K$ such that $\mu^\star_{k'} > 0$ (40)
Example 1: Minimax solution in Bayes strategy [12] minimize $\max_{\mu\in\mathcal{P}} \sum_{k=0}^{R-1} \mu_k \sum_{m=0}^{R-1} B_{m,k}\mathrm{Tr}(\hat{\rho}_k\hat{\Pi}_m)$ subject to $\Pi \in \mathcal{M}$ (32)	$\sum_{m=0}^{R-1} B_{m,k}\mathrm{Tr}(\hat{\rho}_k\hat{\Pi}^\star_m) \leqslant \sum_{m=0}^{R-1} B_{m,k'}\mathrm{Tr}(\hat{\rho}_{k'}\hat{\Pi}^\star_m)$, $\forall k,k' \in \mathcal{I}_R$ such that $\mu^\star_{k'} > 0$
Example 2: Inconclusive minimax solution [23] maximize $\min_{\mu\in\mathcal{P}} \sum_{k=0}^{R-1} \mu_k \mathrm{Tr}[\hat{\rho}_k(\hat{\Pi}_k + \hat{\Pi}_R)]$ subject to $\Pi \in \mathcal{M}^\circ$, where $\mathcal{M}^\circ = \{\Pi \in \mathcal{M} : \mathrm{Tr}(\hat{\rho}_j\hat{\Pi}_R) \leqslant p, \forall\, j \in \mathcal{I}_R\}$ (35)	$\mathrm{Tr}[\hat{\rho}_k(\hat{\Pi}^\star_k + \hat{\Pi}^\star_R)] \geqslant \mathrm{Tr}[\hat{\rho}_{k'}(\hat{\Pi}^\star_{k'} + \hat{\Pi}^\star_R)]$, $\forall\, k,k' \in \mathcal{I}_R$ such that $\mu^\star_{k'} > 0$
Example 3: Minimax solution for plural state sets maximize $\min_{\mu\in\mathcal{P}} \sum_{k=0}^{K-1} \mu_k \sum_{m=0}^{R-1} \mathrm{Tr}(\hat{\rho}'_{k,m}\hat{\Pi}_m)$ subject to $\Pi \in \mathcal{M}$ (37)	$\sum_{m=0}^{R-1} \mathrm{Tr}(\hat{\rho}'_{k,m}\hat{\Pi}^\star_m) \geqslant \sum_{m=0}^{R-1} \mathrm{Tr}(\hat{\rho}'_{k',m}\hat{\Pi}^\star_m)$, $\forall\, k,k' \in \mathcal{I}_K$ such that $\mu^\star_{k'} > 0$ (72)

APPENDIX: DERIVATION OF DUAL PROBLEM WITHOUT BORN RULE

Here we show that the optimal value of problem (12) is an upper bound of the objective function $f(\Pi)$ for a generalized optimal measurement Π without using the Born rule [i.e., the fact that the probability $P(m|\hat{\rho})$ of the outcome $m \in \mathcal{I}_M$ for input state $\hat{\rho}$ is $\mathrm{Tr}(\hat{\rho}\hat{\Pi}_m)$]. We pose the following two requirements.

(1) Any quantum state is given by a density operator, which is positive semidefinite with unit trace.

(2) The probability $P(m|\hat{\rho})$ is affine in $\hat{\rho}$, that is, for any two states $\hat{\rho}$ and $\hat{\rho}'$ and any $t \in \mathcal{R}$ with $0 \leqslant t \leqslant 1$, we have

$$P[m|t\hat{\rho} + (1-t)\hat{\rho}'] = tP(m|\hat{\rho}) + (1-t)P(m|\hat{\rho}'). \quad \text{(A1)}$$

Since $t\hat{\rho} + (1-t)\hat{\rho}'$ can be interpreted as a statistical mixture of the states $\hat{\rho}$ and $\hat{\rho}'$ with probabilities t and $1-t$, the second requirement seems to be natural, which is also pointed out in Ref. [33]. Since $P(m|\hat{\rho})$ is a probability, it must satisfy $P(m|\hat{\rho}) \geqslant 0$ for any $m \in \mathcal{I}_M$ and $\sum_{m=0}^{M-1} P(m|\hat{\rho}) = 1$.

To simplify the notation, we extend $P(m|\hat{\rho})$ to a linear mapping, which we denote as $p(m|\hat{\rho})$, as follows. $p(m|\hat{A})$ ($m \in \mathcal{I}_M, \hat{A} \in \mathcal{S}$) is defined such that $p(m|\hat{\rho}) = P(m|\hat{\rho})$ holds for any density operator $\hat{\rho}$ and it satisfies, for any $t,t' \in \mathcal{R}$ and $\hat{A},\hat{A}' \in \mathcal{S}$,

$$p(m|t\hat{A} + t'\hat{A}') = tp(m|\hat{A}) + t'p(m|\hat{A}'). \quad \text{(A2)}$$

This equation means that $p(m|\hat{A})$ is linear in $\hat{A}$. This definition uniquely determines $p(m|\cdot)$ for a given $P(m|\cdot)$. Since any $\hat{A} \in \mathcal{S}_+$ can be expressed by a form of $\hat{A} = t\hat{\rho}$ with $t = \mathrm{Tr}\hat{A} \geqslant 0$ and a density operator $\hat{\rho} = \hat{A}/\mathrm{Tr}\hat{A}$, we obtain

$$p(m|\hat{A}) \geqslant 0, \ \forall\ \hat{A} \in \mathcal{S}_+. \quad \text{(A3)}$$

Moreover, for any $\hat{A} \in \mathcal{S}$, let a Schmidt decomposition of $\hat{A}$ be $\hat{A} = \sum_n \lambda_n \hat{P}_n$ ($\hat{P}_n$ can be regarded as a density operator); then, from the linearity of $p(m|\cdot)$ and $\sum_{m=0}^{M-1} P(m|\hat{P}_n) = 1$, we obtain

$$\sum_{m=0}^{M-1} p(m|\hat{A}) = \sum_n \lambda_n \sum_{m=0}^{M-1} P(m|\hat{P}_n) = \mathrm{Tr}\hat{A}. \quad \text{(A4)}$$

It should be noted that we can derive from the above two requirements [and Eq. (A2)] that, for any quantum measurement, there exists a POVM $\Pi = \{\hat{\Pi}_m : m \in \mathcal{I}_M\}$ satisfying $P(m|\hat{\rho}) = \mathrm{Tr}(\hat{\rho}\hat{\Pi}_m)$; i.e., the Born rule holds. However, we do not use this fact in this section.

To avoid using the Born rule, we consider the following problem instead of problem (2):

$$\begin{aligned} &\text{maximize} \quad f_p(\Pi) = \sum_{m=0}^{M-1} p(m|\hat{c}_m) \\ &\text{subject to} \quad \Pi \in \mathcal{M}^\bullet, \end{aligned} \tag{A5}$$

where $\hat{c}_m \in \mathcal{S}$ holds for any $m \in \mathcal{I}_M$. Π is a quantum measurement, which is expressed as a collection of mappings $p(m|\cdot)$, i.e., $\{p(m|\cdot) : m \in \mathcal{I}_M\}$. $\mathcal{M}^\bullet$ is defined by

$$\mathcal{M}^\bullet = \left\{ \Pi : \sum_{m=0}^{M-1} p(m|\hat{a}_{j,m}) \leqslant b_j, \ \forall\, j \in \mathcal{I}_J \right\}, \tag{A6}$$

where $\hat{a}_{j,m} \in \mathcal{S}$ and $b_j \in \mathcal{R}$ hold for any $m \in \mathcal{I}_M$ and $j \in \mathcal{I}_J$. J is a non-negative integer.

Let $\lambda \in \mathcal{R}_+^J$. Also, choose $\hat{X}$ such that $\hat{X} \geqslant \hat{z}_m(\lambda)$ holds for any $m \in \mathcal{I}_M$, where $\hat{z}_m(\lambda)$ is defined by Eq. (13). From Eq. (A3), for any $m \in \mathcal{I}_M$, we have

$$p(m|\hat{X}) - p[m|\hat{z}_m(\lambda)] = p[m|\hat{X} - \hat{z}_m(\lambda)] \geqslant 0. \tag{A7}$$

Thus, from Eq. (A4), we have

$$\sum_{m=0}^{M-1} p[m|\hat{z}_m(\lambda)] \leqslant \sum_{m=0}^{M-1} p(m|\hat{X}) = \mathrm{Tr}\hat{X}. \tag{A8}$$

Therefore, we obtain for any $\Pi \in \mathcal{M}^\bullet$,

$$\begin{aligned} f_p(\Pi) &\leqslant \sum_{m=0}^{M-1} p(m|\hat{c}_m) + \sum_{j=0}^{J-1} \lambda_j \left[b_j - \sum_{m=0}^{M-1} p(m|\hat{a}_{j,m}) \right] \\ &= \sum_{m=0}^{M-1} p[m|\hat{z}_m(\lambda)] + \sum_{j=0}^{J-1} \lambda_j b_j \leqslant \mathrm{Tr}\hat{X} + \sum_{j=0}^{J-1} \lambda_j b_j, \end{aligned} \tag{A9}$$

where the equality in the second line follows from Eq. (13). Equation (A9) means that the optimal value of problem (12) provides an upper bound of the optimal value of problem (A5).

It is worth mentioning that the discussion given above has a strong relationship with the approach described in Refs. [34,35], in which it is pointed out that the average correct probability of a minimum error measurement is upper bounded by ensemble steering and the no-signaling principle, and its upper bound equals the average correct probability. In preparation, we introduce ensemble steering. Assume that two parties, Alice and Bob, share an entangled state and the reduced state on Bob's side is $\hat{\rho}$ that can represent

$$\hat{\rho} = \sum_{n=0}^{N-1} q_n \hat{\rho}_n, \tag{A10}$$

with $N \geqslant 2$, where $\hat{\rho}_n$ is a density operator and $q_n \geqslant 0$ satisfies $\sum_{n=0}^{N-1} q_n = 1$. Then, there exists an Alice measurement with N outcomes that prepares Bob's state $\hat{\rho}_n$ with probability q_n. This is known as ensemble steering, which was first noted by Schrödinger [45,46] and also formalized as the Gisin-Hughston-Jozsa-Wootters theorem [47,48]. The probability that Bob obtains the result m given that his state is $\hat{\rho}$ can be expressed as

$$P(m|\hat{\rho}) = \sum_{n=0}^{N-1} q_n P(m|\hat{\rho}_n), \tag{A11}$$

where the right-hand side denotes the weighted average of the conditional probabilities that Bob obtains the result m knowing that his state is $\hat{\rho}_0,\ldots,\hat{\rho}_{N-1}$ with weights $q_0,\ldots,q_{N-1}$. Equations (A10) and (A11) mean that $P(m|\hat{\rho})$ is affine in $\hat{\rho}$. In other words, the property that $P(m|\hat{\rho})$ is affine in $\hat{\rho}$ can be derived from ensemble steering; thus, one can apply the discussion described in this section. Note that the approach described in Refs. [34,35] also provides an operational interpretation for the average correct probability of a minimum error measurement.

[1] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, San Diego, 1976).
[2] A. S. Holevo, J. Multivar. Anal. **3**, 337 (1973).
[3] H. P. Yuen, K. S. Kennedy, and M. Lax, IEEE Trans. Inf. Theory **21**, 125 (1975).
[4] Y. C. Eldar, A. Megretski, and G. C. Verghese, IEEE Trans. Inf. Theory **49**, 1007 (2003).
[5] V. P. Belavkin, Stochastics **1**, 315 (1975).
[6] M. Ban, K. Kurokawa, R. Momose, and O. Hirota, Int. J. Theor. Phys. **36**, 1269 (1997).
[7] T. S. Usuda, I. Takumi, M. Hata, and O. Hirota, Phys. Lett. A **256**, 104 (1999).
[8] Y. C. Eldar and G. D. Forney, Jr., IEEE Trans. Inf. Theory **47**, 858 (2001).
[9] K. Nakahira and T. S. Usuda, Phys. Rev. A **86**, 062305 (2012).
[10] O. Hirota and S. Ikehara, Trans. IECE Jpn. E **65**, 627 (1982).
[11] M. Osaki, M. Ban, and O. Hirota, Phys. Rev. A **54**, 1691 (1996).
[12] K. Kato, in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)* (IEEE, Cambridge, MA, USA, 2012), pp. 1077–1081.
[13] I. D. Ivanovic, Phys. Lett. A **123**, 257 (1987).
[14] D. Dieks, Phys. Lett. A **126**, 303 (1988).
[15] A. Peres, Phys. Lett. A **128**, 19 (1988).
[16] A. Chefles and S. M. Barnett, J. Mod. Opt. **45**, 1295 (1998).
[17] Y. C. Eldar, Phys. Rev. A **67**, 042309 (2003).
[18] J. Fiurášek and M. Ježek, Phys. Rev. A **67**, 012321 (2003).
[19] M. A. P. Touzel, R. B. A. Adamson, and A. M. Steinberg, Phys. Rev. A **76**, 062314 (2007).
[20] A. Hayashi, T. Hashimoto, and M. Horibe, Phys. Rev. A **78**, 012333 (2008).
[21] H. Sugimoto, T. Hashimoto, M. Horibe, and A. Hayashi, Phys. Rev. A **80**, 052322 (2009).
[22] G. M. D'Ariano, M. F. Sacchi, and J. Kahn, Phys. Rev. A **72**, 032310 (2005).

[23] K. Nakahira, K. Kato, and T. S. Usuda, Phys. Rev. A **88**, 032314 (2013).
[24] K. Nakahira, T. S. Usuda, and K. Kato, Phys. Rev. A **86**, 032316 (2012).
[25] U. Herzog, Phys. Rev. A **86**, 032314 (2012).
[26] Y. C. Eldar, IEEE Trans. Inf. Theory **49**, 446 (2003).
[27] S. Pang and S. Wu, Phys. Rev. A **80**, 052320 (2009).
[28] L. Roa, C. Hermann-Avigliano, R. Salazar, and A. Klimov, Phys. Rev. A **84**, 014302 (2011).
[29] J. A. Bergou, U. Futschik, and E. Feldman, Phys. Rev. Lett. **108**, 250502 (2012).
[30] S. Bandyopadhyay, Phys. Rev. A **90**, 030301(R) (2014).
[31] M. Slater, Cowles Comm. Discuss. Paper, Math. No. 403 (1950).
[32] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, UK, 2009).
[33] G. Kimura, T. Miyadera, and H. Imai, Phys. Rev. A **79**, 062306 (2009).
[34] J. Bae, W.-Y. Hwang, and Y.-D. Han, Phys. Rev. Lett. **107**, 170403 (2011).
[35] J. Bae, New J. Phys. **15**, 073037 (2013).
[36] K. Nakahira and T. S. Usuda, Phys. Rev. A **87**, 012308 (2013).
[37] I. Ekeland and R. Temam, *Convex Analysis and Variational Problems* (SIAM, North Holland, 1999), Vol. 28.
[38] Y. C. Eldar, A. Megretski, and G. C. Verghese, IEEE Trans. Inf. Theory **50**, 1198 (2004).
[39] Y. C. Eldar, M. Stojnic, and B. Hassibi, Phys. Rev. A **69**, 062318 (2004).
[40] E. Andersson, S. M. Barnett, C. R. Gilson, and K. Hunter, Phys. Rev. A **65**, 052308 (2002).
[41] K. Kato and O. Hirota, IEEE Trans. Inf. Theory **49**, 3312 (2003).
[42] D. Qiu, Phys. Rev. A **77**, 012328 (2008).
[43] A. Assalini, G. Cariolaro, and G. Pierobon, Phys. Rev. A **81**, 012315 (2010).
[44] K. Nakahira, K. Kato, and T. S. Usuda, Phys. Rev. A **91**, 012318 (2015).
[45] E. Schrödinger, in *Proceedings of the Cambridge Philosophical Society* (Cambridge University Press, Cambridge, UK, 1935), Vol. 31, pp. 555–563.
[46] E. Schrödinger, in *Proceedings of the Cambridge Philosophical Society* (Cambridge University Press, Cambridge, UK, 1936), Vol. 32, pp. 446–452.
[47] N. Gisin, Helv. Phys. Acta **62**, 363 (1989).
[48] L. P. Hughston, R. Jozsa, and W. K. Wootters, Phys. Lett. A **183**, 14 (1993).