Generalized bipartite quantum state discrimination problems with sequential measurements

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

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

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

(Received 20 July 2017; published 27 February 2018; corrected 13 January 2021)

We investigate an optimization problem of finding quantum sequential measurements, which forms a wide class of state discrimination problems with the restriction that only local operations and one-way classical communication are allowed. Sequential measurements from Alice to Bob on a bipartite system are considered. Using the fact that the optimization problem can be formulated as a problem with only Alice's measurement and is convex programming, we derive its dual problem and necessary and sufficient conditions for an optimal solution. Our results are applicable to various practical optimization criteria, including the Bayes criterion, the Neyman-Pearson criterion, and the minimax criterion. In the setting of the problem of finding an optimal global measurement, its dual problem and necessary and sufficient conditions for an optimal solution have been widely used to obtain analytical and numerical expressions for optimal solutions. Similarly, our results are useful to obtain analytical and numerical expressions for optimal sequential measurements. Examples in which our results can be used to obtain an analytical expression for an optimal sequential measurement are provided.

DOI: 10.1103/PhysRevA.97.022340

I. INTRODUCTION

The study of the power and limitations of local discrimination of quantum states has attracted considerable interest in quantum information theory in recent years. In particular, quantum measurements realized by local operations and one-way classical communication (one-way LOCC), also called sequential measurements, have been widely investigated. Sequential measurements are relatively easy to implement with current technology; for example, when two or more parties receive quantum states at different times, measurements in which individual measurements are performed sequentially would be desirable in practical implementations of quantum measurements. However, it is well known that orthogonal quantum states shared by separated parties may not be perfectly distinguished when only sequential measurements are allowed, while they can be perfectly distinguished by a global measurement. This implies that sequential measurements are less powerful than global measurements for quantum state discrimination. An important question that arises in studies of this kind is how well one can distinguish between given quantum states by a sequential measurement.

Many studies have been developed to tackle the problem of which sets of orthogonal states are distinguishable when only sequential measurements are allowed (e.g., [1–7]). There have also been several investigations of a sequential measurement realizing a measurement that maximizes the average success probability (called a minimum-error measurement) [8–11]. It has also been reported that a measurement that maximizes the average success probability with no error at the expense of allowing for a certain fraction of inconclusive (failure) results (called an optimal unambiguous measurement) can be realized by a sequential measurement for binary pure states [12–14]. However, these results are only applicable to a special class of quantum states. Investigations applicable to a broad class of quantum states would be required.

In the scenario in which all quantum measurements are allowed, optimal measurement strategies have been investigated under various criteria, such as the Bayes criterion [15–17] and the minimax criterion [18–20]. A measurement strategy that allows for inconclusive results has also been well studied. The most well-known example along this line is an optimal unambiguous measurement [21–23]. Other examples are a measurement that maximizes the average success probability with a fixed average inconclusive probability, referred to as an optimal inconclusive measurement [24–26], and a measurement that maximizes the average success probability under the condition that the average error probability should not exceed a certain error, referred to an optimal error margin measurement [27–29]. Recently, a generalized state discrimination problem, which is applicable to the above-mentioned criteria, has also been presented [30]. From these studies, some properties of optimal measurements in the above criteria, such as necessary and sufficient conditions for optimality, have been derived. By contrast, in the case of a sequential measurement, very few studies of an optimal sequential measurement for a strategy other than the minimum-error strategy and the unambiguous strategy have been reported (e.g., [31–35]).

More recently, Croke *et al.* derived a necessary and sufficient condition for a sequential measurement to maximize the average success probability (we call such a measurement a minimum-error sequential measurement) and used it to prove optimality of a candidate solution [36]. Also, the authors have derived the dual problem of finding a minimum-error sequential measurement and utilized it to compute numerical solutions [37]. These results are applicable to arbitrary bipartite quantum states; however, only a few properties of a minimum-error sequential measurement have ever been reported. In addition, these methods cannot directly be applied to other criteria.

In this paper, we address a sequential-measurement version of the generalized state discrimination problem described in

Ref. [30]. Similar to that work, this problem includes problems with various criteria. We consider sequential measurements from Alice to Bob on a bipartite system. Since the problem of finding an optimal sequential measurement is much more complex than that of finding an optimal global measurement, the results proposed in Ref. [30] cannot readily be applied to this problem. However, we can see that the entire set of sequential measurements is convex; thus, the generalized state discrimination problem with sequential measurements can be formulated as a convex programming problem. Useful results available in convex programming help us to further understand an optimal sequential measurement. We derive the original dual problem and necessary and sufficient conditions for an optimal solution. In the problem we address, sequential measurements with a finite number of outcomes are considered, whereas the output of Alice's measurement can be infinite or continuous. We show that there always exists an optimal sequential measurement in which Alice's measurement with a finite number of outcomes as long as a solution exists. These properties would be useful to obtain analytical and numerical expressions for an optimal sequential measurement.

In Sec. II, we discuss the formulation of sequential measurements and provide a sequential-measurement version of the generalized state discrimination problem. In Sec. III, its dual problem is derived. Then we show that the optimal values of the primal and dual problems are the same. Necessary and sufficient conditions for an optimal solution are also obtained. In Sec. IV, we show that if a problem has a certain symmetry, then there exists an optimal solution with the same type of symmetry. In Sec. V, we discuss a sequential-measurement version of the generalized minimax problem described in Ref. [30]. We also derive necessary and sufficient conditions for a minimax solution. In Sec. VI, we show two examples in which analytical expressions of optimal sequential measurements are derived. These examples illustrate that our results are useful to obtain analytical solutions to some problems.

II. GENERALIZED OPTIMAL SEQUENTIAL MEASUREMENT

A. Sequential measurement

We consider a composite system $\mathcal{H} = \mathcal{H}_\mathrm{A} \otimes \mathcal{H}_\mathrm{B}$ of two subsystems, Alice and Bob. Let $\mathcal{S}$ and $\mathcal{S}^+$ be, respectively, the entire sets of Hermitian operators and positive semidefinite operators on $\mathcal{H}$. Here $\mathcal{S}_k$ and $\mathcal{S}_k^+$ ($k \in \{\mathrm{A,B}\}$) are defined in the same way with $\mathcal{H}$ replaced by $\mathcal{H}_k$. Also, let $\mathbf{R}$ and $\mathbf{R}_+$ be, respectively, the entire sets of real numbers and non-negative real numbers and $\mathcal{I}_N \equiv \{0,1,\ldots,N-1\}$. Let $\hat{1}$, $\hat{1}_\mathrm{A}$, and $\hat{1}_\mathrm{B}$ be, respectively, the identity operators on $\mathcal{H}$, $\mathcal{H}_\mathrm{A}$, and $\mathcal{H}_\mathrm{B}$. We denote $\{tb_n\}$ and $\{b_n + b'_n\}$, with $t \in \mathbf{R}$ and $b,b' \in \mathbf{R}^N$ (or $b,b' \in \mathbf{R}_+^N$), by tb and $b + b'$, respectively. In addition, $\hat{x} \geqslant \hat{y}$, with Hermitian operators $\hat{x}$ and $\hat{y}$, means that $\hat{x} - \hat{y}$ is positive semidefinite.

Let us consider a sequential measurement on $\mathcal{H}$. Alice first performs a measurement, which is represented by a positive-operator-valued measure (POVM) $\{\hat{A}_j \in \mathcal{S}_\mathrm{A}^+\}_j$, the output of which can be infinite (or continuous). The measurement result j is sent to Bob. Then Bob chooses a measurement $\{\hat{B}_m^{(j)} \in \mathcal{S}_\mathrm{B}^+\}_{m=0}^{M-1}$ depending on j and obtains the outcome $m \in \mathcal{I}_M$, which represents the final measurement result. The measurement on the joint system is given by the POVM $\{\hat{\Pi}_m = \sum_j \hat{A}_j \otimes \hat{B}_m^{(j)}\}_{m=0}^{M-1}$.

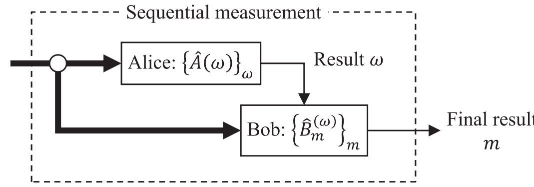

FIG. 1. Schematic diagram of a sequential measurement seen from a different viewpoint. Each of Alice's outcomes ω corresponds one to one to Bob's POVM $\{\hat{B}_m^{(\omega)}\}_m$.

We can consider this sequential measurement from a different viewpoint [37]. Let $\mathcal{M}_\mathrm{B}$ be the entire set of allowed Bob's measurements and Ω be an isomorphic set of $\mathcal{M}_\mathrm{B}$. Each element of $\mathcal{M}_\mathrm{B}$ is uniquely labeled by an index $\omega \in \Omega$; we define Bob's measurement corresponding to $\omega \in \Omega$ as $\hat{B}^{(\omega)} \equiv \{\hat{B}_m^{(\omega)}\}_{m=0}^{M-1}$. Alice first performs a measurement $\hat{A}$ with continuous outcomes in Ω. She sends the result $\omega \in \Omega$ to Bob. He performs the corresponding measurement $\hat{B}^{(\omega)}$. A schematic diagram is depicted in Fig. 1. Alice's POVM $\hat{A}$ uniquely determines this sequential measurement, which is defined as $\hat{\Pi}^{(\hat{A})} \equiv \{\hat{\Pi}_m^{(\hat{A})}\}_{m=0}^{M-1}$ with

$$\hat{\Pi}_m^{(\hat{A})} \equiv \int_\Omega \hat{A}(d\omega) \otimes \hat{B}_m^{(\omega)}. \tag{1}$$

We can interpret that Alice's POVM $\hat{A}$ includes all the information regarding the measurements Bob should perform. Let $\mathcal{M}_\mathrm{A}$ be the entire set of Alice's POVMs. Any sequential measurement can be denoted by $\hat{\Pi}^{(\hat{A})}$ with $\hat{A} \in \mathcal{M}_\mathrm{A}$. In this formulation, the problem of finding an optimal sequential measurement can be formulated as an optimization problem with only $\hat{A}$. It should be noted that $\mathcal{M}_\mathrm{B}$ is not necessarily the entire set of POVMs on $\mathcal{H}_\mathrm{B}$; for example, $\mathcal{H}_\mathrm{B}$ can be a composite system of n subsystems and $\mathcal{M}_\mathrm{B}$ can be the entire set of sequential measurements (or two-way LOCC measurements) on $\mathcal{H}_\mathrm{B}$.

B. State discrimination problem

Here we consider a sequential-measurement version of the optimization problem described in Ref. [30], which is expressed as

$$\text{Problem P: maximize } f(\hat{A}) \equiv \sum_{m=0}^{M-1} \mathrm{Tr}\big[\hat{c}_m \hat{\Pi}_m^{(\hat{A})}\big]$$
$$\text{subject to} \quad \hat{A} \in \mathcal{M}_\mathrm{A}^\circ, \tag{2}$$

where

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

$\hat{c}_m \in \mathcal{S}$, $\hat{a}_{j,m} \in \mathcal{S}$, and $b_j \in \mathbf{R}$. Here J is a non-negative integer that represents the number of constraints.

We can easily verify that $\mathcal{M}_{\rm A}^\circ$ is convex and thus Problem P is a convex programming problem. Let $f^\star$ be the optimal value of Problem P; $f^\star$ is regarded as $-\infty$ if the feasible set $\mathcal{M}_{\rm A}^\circ$ is empty. Note that an equality constraint $\eta_j(\hat{A}) = 0$ can be replaced by two inequality constraints $\eta_j(\hat{A}) \leqslant 0$ and $-\eta_j(\hat{A}) \leqslant 0$.

Problem P can express a large class of problems; one can find some examples in Sec. II B of Ref. [30]. We will also provide two examples in the next section.

C. Examples

We give some examples of problems of finding optimal sequential measurements that can be formulated as Problem P. We consider discrimination between R quantum states $\{\tilde{\rho}_r\}_{r=0}^{R-1}$ with prior probabilities $\{\xi_r\}_{r=0}^{R-1}$. For each $r \in \mathcal{I}_R$, $\tilde{\rho}_r$ is a density operator satisfying $\tilde{\rho}_r \in \mathcal{S}^+$ and $\mathrm{Tr}\,\tilde{\rho}_r = 1$. We let $\hat{\rho}_r \equiv \xi_r \tilde{\rho}_r$.

1. Optimal inconclusive measurement

Let us consider a sequential measurement with inconclusive results $\hat{\Pi}^{(\hat{A})} = \{\hat{\Pi}_r^{(\hat{A})}\}_{r=0}^{R}$ ($\hat{A} \in \mathcal{M}_{\rm A}$ and $R \geqslant 2$). The detection operator $\hat{\Pi}_r^{(\hat{A})}$ with $r < R$ corresponds to identification of the state $\tilde{\rho}_r$, while $\hat{\Pi}_R^{(\hat{A})}$ corresponds to the inconclusive answer. An optimal inconclusive sequential measurement is a sequential measurement that maximizes the average success probability under the constraint that the average inconclusive probability equals a given value $p_{\rm I}$ with $0 \leqslant p_{\rm I} \leqslant 1$. The problem of obtaining such a measurement can be formulated as follows:

$$\begin{aligned} &\text{maximize } P_{\rm S}(\hat{A}) \equiv \sum_{r=0}^{R-1} \mathrm{Tr}\big[\hat{\rho}_r \hat{\Pi}_r^{(\hat{A})}\big] \\ &\text{subject to} \quad \hat{A} \in \mathcal{M}_{\rm A}, \quad \sum_{r=0}^{R-1} \mathrm{Tr}\big[\hat{\rho}_r \hat{\Pi}_R^{(\hat{A})}\big] = p_{\rm I}. \end{aligned} \tag{4}$$

This problem is equivalent to Problem P with

$$\begin{aligned} M &= R+1, \quad J = 1, \\ \hat{c}_m &= \begin{cases} \hat{\rho}_m, & m < R \\ 0, & m = R, \end{cases} \\ \hat{a}_{0,m} &= \begin{cases} 0, & m < R \\ -\displaystyle\sum_{r=0}^{R-1} \hat{\rho}_r, & m = R, \end{cases} \\ b_0 &= -p_{\rm I}, \end{aligned} \tag{5}$$

where we use the fact that the problem remains unchanged when the second constraint of Eq. (4) is replaced with $\sum_{r=0}^{R-1} \mathrm{Tr}[\hat{\rho}_r \hat{\Pi}_R^{(\hat{A})}] \geqslant p_{\rm I}$.

In particular, in the case of $p = 0$, an optimal inconclusive sequential measurement is a minimum-error sequential measurement. Since, in this case, we can assume $\hat{\Pi}_R^{(\hat{A})} = 0$, this problem is rewritten as

$$\begin{aligned} &\text{maximize } \ P_{\rm S}(\hat{A}) \\ &\text{subject to } \ \hat{A} \in \mathcal{M}_{\rm A}, \end{aligned} \tag{6}$$

which is equivalent to Problem P with $M = R$, $J = 0$, and $\hat{c}_m = \hat{\rho}_m$.

2. Optimal measurement in the Bayes criterion with a constraint

Another example is an extension of the problem of finding an optimal sequential measurement in the Bayes criterion. Let us consider the following problem:

$$\begin{aligned} &\text{minimize } C_0(\hat{A}) \\ &\text{subject to } \hat{A} \in \mathcal{M}_{\rm A}, \quad C_1(\hat{A}) \leqslant \chi. \end{aligned} \tag{7}$$

For each $k \in \{0,1\}$, $C_k(\hat{A})$ is a cost function defined by

$$C_k(\hat{A}) \equiv \sum_{m=0}^{R-1} \sum_{r=0}^{R-1} w_{m,r}^{(k)} \mathrm{Tr}\big[\hat{\rho}_r \hat{\Pi}_m^{(\hat{A})}\big], \tag{8}$$

where $w_{m,r}^{(k)} \in \mathbf{R}_+$ holds for any $m,r \in \mathcal{I}_R$. This problem can be interpreted as that of finding a sequential measurement $\hat{\Pi}^{(\hat{A})}$ that minimizes the cost $C_0(\hat{A})$ under the constraint that the other cost $C_1(\hat{A})$ should not be greater than a given value χ. The problem of Eq. (7) is equivalent to Problem P with

$$\begin{aligned} M &= R, \quad J = 1, \\ \hat{c}_m &= \sum_{r=0}^{R-1} w_{m,r}^{(0)} \hat{\rho}_r, \\ \hat{a}_{0,m} &= \sum_{r=0}^{R-1} w_{m,r}^{(1)} \hat{\rho}_r, \quad b_0 = \chi. \end{aligned} \tag{9}$$

For an example, in the case of $\chi = \infty$, Eq. (7) represents the sequential version of the problem of finding an optimal measurement in the traditional Bayes criterion (with no constraint) [15–17]. For another example, in the case of

$$\begin{aligned} R &= 2, \\ w_{m,r}^{(0)} &= \xi_r^{-1} \delta_{m,0} \delta_{r,1}, \\ w_{m,r}^{(1)} &= \xi_r^{-1} \delta_{m,1} \delta_{r,0} \end{aligned} \tag{10}$$

($\delta_{j,j'}$ is Kronecker delta), Eq. (7) represents the problem of finding an optimal sequential measurement $\{\hat{\Pi}_0^{(\hat{A})}, \hat{\Pi}_1^{(\hat{A})}\}$ that minimizes $C_0(\hat{A}) = \mathrm{Tr}[\tilde{\rho}_1 \hat{\Pi}_0^{(\hat{A})}]$ under the constraint $C_1(\hat{A}) = \mathrm{Tr}[\tilde{\rho}_0 \hat{\Pi}_1^{(\hat{A})}] \leqslant \chi$, which is known as the Neyman-Pearson criterion.

III. OPTIMAL SOLUTION TO THE GENERALIZED PROBLEM

A. Dual problem

We will derive the dual problem of Problem P. Let

$$\begin{aligned} \hat{\sigma}_\omega(\lambda) &\equiv \mathrm{Tr}_{\rm B} \sum_{m=0}^{M-1} \hat{z}_m(\lambda) \hat{B}_m^{(\omega)}, \\ \hat{z}_m(\lambda) &\equiv \hat{c}_m - \sum_{j=0}^{J-1} \lambda_j \hat{a}_{j,m}, \end{aligned} \tag{11}$$

where $\mathrm{Tr}_{\rm B}$ is the partial trace with respect to the system $\mathcal{H}_{\rm B}$. Here $\hat{\sigma}_\omega(\lambda) \in \mathcal{S}_{\rm A}$ and $\hat{z}_m(\lambda) \in \mathcal{S}$ obviously hold. From Eqs. (1)

and (11), we have that for any $\hat{A} \in \mathcal{M}_{\mathrm{A}}$,

$$\sum_{m=0}^{M-1} \mathrm{Tr}\big[\hat{z}_m(\lambda)\hat{\Pi}_m^{(\hat{A})}\big] = \mathrm{Tr}\int_\Omega \hat{\sigma}_\omega(\lambda)\hat{A}(d\omega). \tag{12}$$

Thus, from Eqs. (2) and (3), we have that for any $\hat{A} \in \mathcal{M}_{\mathrm{A}}^\circ$ and $\lambda \in \mathbf{R}_+^J$,

$$\begin{aligned} f(\hat{A}) &\leqslant \sum_{m=0}^{M-1} \mathrm{Tr}\big[\hat{c}_m\hat{\Pi}_m^{(\hat{A})}\big] - \sum_{j=0}^{J-1}\lambda_j\eta_j(\hat{A}) \\ &= \mathrm{Tr}\big[\hat{z}_m(\lambda)\hat{\Pi}_m^{(\hat{A})}\big] + \sum_{j=0}^{J-1}\lambda_j b_j \\ &= \mathrm{Tr}\int_\Omega \hat{\sigma}_\omega(\lambda)\hat{A}(d\omega) + \sum_{j=0}^{J-1}\lambda_j b_j. \end{aligned} \tag{13}$$

Here let $\hat{X}$ be a Hermitian operator on $\mathcal{H}_{\mathrm{A}}$ satisfying $\hat{X} \geqslant \hat{\sigma}_\omega(\lambda)$ for any $\omega \in \Omega$. From Eq. (13) we have

$$\begin{aligned} f(\hat{A}) &\leqslant \mathrm{Tr}\int_\Omega \hat{X}\hat{A}(d\omega) + \sum_{j=0}^{J-1}\lambda_j b_j \\ &= \mathrm{Tr}\hat{X} + \sum_{j=0}^{J-1}\lambda_j b_j. \end{aligned} \tag{14}$$

This implies that we can obtain the following dual problem:

$$\begin{aligned} \text{Problem DP: minimize } & s(\hat{X},\lambda) \equiv \mathrm{Tr}\hat{X} + \sum_{j=0}^{J-1}\lambda_j b_j \\ \text{subject to } & (\hat{X},\lambda) \in \mathcal{X}^\circ \end{aligned} \tag{15}$$

with variables $\hat{X}$ and λ, where

$$\begin{aligned} \mathcal{X}^\circ &\equiv \{(\hat{X},\lambda) \in \mathcal{X} : \hat{X} \geqslant \hat{\sigma}_\omega(\lambda)\ \forall\omega \in \Omega\}, \\ \mathcal{X} &\equiv \mathcal{S}_{\mathrm{A}} \otimes \mathbf{R}_+^J. \end{aligned} \tag{16}$$

Let $s^\star$ be the optimal value of Problem DP. Problem DP can also be derived using the Lagrangian method (see Appendix A). We can easily verify that Problem DP is also a convex programming problem.

From Eq. (14), $s^\star \geqslant f^\star$ holds, i.e., the optimal value of Problem DP is not less than that of Problem P. Moreover, as stated in the following theorem, they are always the same (proof in Appendix B).

Theorem 1. $s^\star = f^\star$ always holds.

B. Conditions for an optimal solution

In generalized state discrimination problems with no restriction on measurements, necessary and sufficient conditions for an optimal solution have been derived [30]. In a similar manner, we can derive necessary and sufficient conditions for an optimal solution to Problem P using its dual problem (proof in Appendix C).

Theorem 2. Let $\hat{A}$ be a POVM satisfying $\hat{A} \in \mathcal{M}_{\mathrm{A}}^\circ$. The following statements are all equivalent.

(a) $\hat{A}$ is an optimal solution to Problem P.

(b) There exists $(\hat{X},\lambda) \in \mathcal{X}^\circ$ such that

$$[\hat{X} - \hat{\sigma}_\omega(\lambda)]\hat{A}(\omega) = 0\ \forall\omega \in \Omega, \tag{17}$$

$$\lambda_j\eta_j(\hat{A}) = 0\ \forall j \in \mathcal{I}_J. \tag{18}$$

(c) There exists $\lambda \in \mathbf{R}_+^J$ such that

$$\int_\Omega \hat{\sigma}_{\omega'}(\lambda)\hat{A}(d\omega') \geqslant \hat{\sigma}_\omega(\lambda)\ \forall\omega \in \Omega, \tag{19}$$

$$\lambda_j\eta_j(\hat{A}) = 0\ \forall j \in \mathcal{I}_J. \tag{20}$$

Moreover, if condition (b) holds, then $(\hat{X},\lambda)$ is an optimal solution to Problem DP.

From Eq. (17), for any $\omega \in \Omega$, the kernel of $\hat{X} - \hat{\sigma}_\omega(\lambda)$ includes the support of $\hat{A}(\omega)$. If $\hat{A}$ is discrete valued [i.e., $\hat{A}(\omega) \neq 0$ holds for $\omega \in \Omega$ only if ω is in at most countable set $\{\omega_n\}_n$], then Eq. (19) can be rewritten as

$$\sum_n \hat{\sigma}_{\omega_n}\hat{A}(\omega_n) \geqslant \hat{\sigma}_\omega(\lambda)\ \forall\omega \in \Omega. \tag{21}$$

Note that this equation in the case of the problem of obtaining a minimum-error sequential measurement is given by Eq. (19) of Ref. [36]. Although $\hat{A}$ is continuous valued and optimal, there always exists an optimal solution to Problem P with a finite number of outcomes, as will be shown in Theorem 3.

We should mention that obtaining an optimal solution to Problem P is much more difficult than obtaining an optimal solution to the problem described in Ref. [30], i.e., the state discrimination problem with no restriction on measurements. The reason is that, in the former case, we have to optimize over all of Alice's measurements, which include all the information regarding the measurements Bob should perform. Problem DP is generally difficult to solve as well as Problem P. However, we can obtain an analytical solution by solving Problem DP in some cases (see Sec. VI B).

C. Number of outcomes of Alice's POVM

So far in this paper, we have considered Alice's POVM $\hat{A}$ to be continuous. We find that an optimal solution to Problem P with finite outcomes always exists as long as a feasible solution exists, as shown in the following theorem (proof in Appendix D).

Theorem 3. Let $d_{\mathrm{A}} = \dim \mathcal{H}_{\mathrm{A}}$. If $\mathcal{M}_{\mathrm{A}}^\circ$ is not empty, then an optimal solution to Problem P with at most $(J+1)d_{\mathrm{A}}^2$ outcomes exists.

D. Examples

As specific examples, in Sec. II C, we showed the problem of finding an optimal inconclusive sequential measurement and that of finding an optimal measurement in the Bayes criterion with a constraint. We give their dual problems and necessary and sufficient conditions for optimal solutions.

1. Optimal inconclusive measurement

Substituting Eq. (5) into Problem DP, we obtain the following dual problem of finding an optimal inconclusive sequential

TABLE I. Formulation of the generalized state discrimination problems. The date in column 1 are from [30].

	Arbitrary measurements	Sequential measurements
Primal problems	maximize $\sum_{\rm m=0}^{\rm M-1} {\rm Tr}(\hat{\rm c}_{\rm m}\hat{\Pi}_{\rm m})$ subject to $\hat{\Pi}$: POVM, $\sum_{m=0}^{M-1} {\rm Tr}(\hat{a}_{j,m}\hat{\Pi}_m) \leqslant b_j \quad (\forall j \in \mathcal{I}_J)$	maximize $\sum_{\rm m=0}^{\rm M-1} {\rm Tr}[\hat{\rm c}_{\rm m}\hat{\Pi}_{\rm m}^{(\hat{\rm A})}]$ subject to $\hat{A} \in \mathcal{M}_{\rm A}$, $\sum_{m=0}^{M-1} {\rm Tr}[\hat{a}_{j,m}\hat{\Pi}_m^{(\hat{A})}] \leqslant b_j \quad (\forall j \in \mathcal{I}_J)$ (2) and (3)
Dual problems	minimize ${\rm Tr}\hat{\rm X} + \sum_{\rm j=0}^{\rm J-1} \lambda_{\rm j}{\rm b}_{\rm j}$ subject to $\hat{X} \geqslant \hat{z}_m(\lambda) \quad (\forall m \in \mathcal{I}_M),\ \lambda \in \mathbf{R}_+^J$ where $\hat{z}_m(\lambda) = \hat{c}_m - \sum_{j=0}^{J-1} \lambda_j \hat{a}_{j,m}$	minimize ${\rm Tr}\hat{\rm X} + \sum_{\rm j=0}^{\rm J-1} \lambda_{\rm j}{\rm b}_{\rm j}$ subject to $\hat{X} \geqslant \hat{\sigma}_\omega(\lambda) \quad (\forall \omega \in \Omega),\ \lambda \in \mathbf{R}_+^J$ where $\hat{\sigma}_\omega(\lambda) = {\rm Tr}_{\rm B} \sum_{m=0}^{M-1} \hat{z}_m(\lambda)\hat{B}_m^{(\omega)}$, $\hat{z}_m(\lambda) = \hat{c}_m - \sum_{j=0}^{J-1} \lambda_j \hat{a}_{j,m}$ (11), (15), and (16)
Necessary and sufficient conditions for optimality [condition (c)]	$\lambda \in \mathbf{R}_+^J$ exists such that $\sum_{m=0}^{M-1} \hat{z}_m(\lambda)\hat{\Pi}_m \geqslant \hat{z}_m(\lambda)\ \forall m \in \mathcal{I}_M$, $\lambda_j[b_j - \sum_{m=0}^{M-1} {\rm Tr}(\hat{a}_{j,m}\hat{\Pi}_m)] = 0\ \forall j \in \mathcal{I}_J$	$\lambda \in \mathbf{R}_+^J$ exists such that $\int_\Omega \hat{\sigma}_{\omega'}(\lambda)\hat{A}(d\omega') \geqslant \hat{\sigma}_\omega(\lambda)\ \forall \omega \in \Omega$, $\lambda_j[b_j - \sum_{m=0}^{M-1} {\rm Tr}[\hat{a}_{j,m}\hat{\Pi}_m^{(\hat{A})}]] = 0\ \forall j \in \mathcal{I}_J$ (19) and (20)

measurement:

$$\begin{aligned} &\text{minimize } s(\hat{X},\lambda) = {\rm Tr}\hat{X} - \lambda p_{\rm I} \\ &\text{subject to } \quad (\hat{X},\lambda) \in \mathcal{X}^\circ, \end{aligned} \tag{22}$$

where

$$\begin{aligned} \mathcal{X}^\circ &= \{(\hat{X},\lambda) \in \mathcal{S}_{\rm A} \otimes \mathbf{R}_+ : \hat{X} \geqslant \hat{\sigma}_\omega(\lambda), \forall \omega \in \Omega\}, \\ \hat{\sigma}_\omega(\lambda) &= {\rm Tr}_{\rm B} \sum_{r=0}^{R-1} \hat{\rho}_r \big[\hat{B}_r^{(\omega)} + \lambda \hat{B}_R^{(\omega)}\big]. \end{aligned} \tag{23}$$

Let

$$\mathcal{M}_{\rm A}^\circ = \left\{ \hat{A} \in \mathcal{M}_{\rm A} : \sum_{r=0}^{R-1} {\rm Tr}\big[\hat{\rho}_r \hat{\Pi}_R^{(\hat{A})}\big] = p_{\rm I} \right\}. \tag{24}$$

From Theorem 2, $\hat{\Pi}^{(\hat{A})}$ with $\hat{A} \in \mathcal{M}_{\rm A}^\circ$ is an optimal inconclusive sequential measurement if and only if $\hat{A}$ satisfies

$$[\hat{X}^\star - \hat{\sigma}_\omega(\lambda^\star)]\hat{A}(\omega) = 0\ \forall \omega \in \Omega, \tag{25}$$

where $(\hat{X}^\star,\lambda^\star)$ is an optimal solution to Eq. (22). Equation (25) is a sequential-measurement version of an optimal inconclusive global measurement given by Ref. [25].

2. *Optimal measurement in the Bayes criterion with a constraint*

From Eq. (9), we have the following dual problem:

$$\begin{aligned} &\text{minimize } s(\hat{X},\lambda) = {\rm Tr}\hat{X} + \lambda \chi \\ &\text{subject to } \quad (\hat{X},\lambda) \in \mathcal{X}^\circ, \end{aligned} \tag{26}$$

where

$$\begin{aligned} \mathcal{X}^\circ &= \{(\hat{X},\lambda) \in \mathcal{S}_{\rm A} \otimes \mathbf{R}_+ : \hat{X} \geqslant \hat{\sigma}_\omega(\lambda)\ \forall \omega \in \Omega\}, \\ \hat{\sigma}_\omega(\lambda) &= {\rm Tr}_{\rm B} \sum_{m=0}^{R-1}\sum_{r=0}^{R-1} \big[w_{m,r}^{(0)} - \lambda w_{m,r}^{(1)}\big]\hat{\rho}_r \hat{B}_m^{(\omega)}. \end{aligned} \tag{27}$$

In particular, in the case of Eq. (10), i.e., binary hypothesis testing in the Neyman-Pearson criterion, we have

$$\hat{\sigma}_\omega(\lambda) = {\rm Tr}_{\rm B}\big[\tilde{\rho}_1 \hat{B}_0^{(\omega)} - \lambda \tilde{\rho}_0 \hat{B}_1^{(\omega)}\big]. \tag{28}$$

From Theorem 2, $\hat{A} \in \mathcal{M}_{\rm A}$ with $C_1(\hat{A}) \leqslant \chi$ is an optimal solution to Eq. (7) if and only if the following equations hold:

$$\begin{aligned} [\hat{X}^\star - \hat{\sigma}_\omega(\lambda^\star)]\hat{A}(\omega) &= 0\ \forall \omega \in \Omega, \\ \lambda^\star[C_1(\hat{A}) - \chi] &= 0, \end{aligned} \tag{29}$$

where $(\hat{X}^\star,\lambda^\star)$ is an optimal solution to Eq. (26).

E. Comparison with the problem with no restriction on measurements

Table I summarizes the formulation of the state discrimination problems when arbitrary measurements are allowed and when only sequential measurements are allowed. The dual problem in case (b) (i.e., Problem DP) has an infinite (continuous) number of constraints, while that in case (a) has a finite number M of constraints. This makes it difficult to obtain an optimal sequential measurement.

Let $(\hat{X}_{\rm G},\lambda) \in \mathcal{S} \otimes \mathbf{R}_+^J$ be a feasible solution to the dual problem in case (a), which satisfies $\hat{X}_{\rm G} \geqslant \hat{z}_m(\lambda)$ for any $m \in \mathcal{I}_M$.
Postmultiplying both sides of this inequality by $\hat{B}_m^{(\omega)}$, summing over $m = 0,\dots,M-1$, and taking a partial trace over $\mathcal{H}_{\rm B}$, we have

$${\rm Tr}_{\rm B}\hat{X}_{\rm G} \geqslant {\rm Tr}_{\rm B} \sum_{m=0}^{M-1} \hat{z}_m(\lambda)\hat{B}_m^{(\omega)}. \tag{30}$$

Thus, $({\rm Tr}_{\rm B}\hat{X}_{\rm G},\lambda)$ is a feasible solution to Problem DP. This implies that ${\rm Tr}\hat{X}_{\rm G}$ is not less than the optimal value of Problem DP, which is consistent with the fact that global measurements can be better than sequential measurements.

IV. PROBLEM WITH SYMMETRY

In this section, we discuss the case in which Problem P has a certain symmetry. State discrimination problems with symmetries have been well studied and it is known that, in some cases, there exists an optimal solution with the same type of symmetry [25,38–45]. The existence of a symmetric solution helps us to obtain analytical or numerical optimal solutions (e.g., [46–50]). We use basic terminology from group theory that is the mathematical language of symmetry.

A. Group action

First, we briefly introduce a group action. Let $\mathcal{G}$ be a group and $e \in \mathcal{G}$ be its identity element. Also, let $\overline{g} \in \mathcal{G}$ be the inverse element of $g \in \mathcal{G}$. We assume that $\mathcal{G}$ has at least two elements. Let $|\mathcal{G}|$ be the number of elements in $\mathcal{G}$. A group action of $\mathcal{G}$ on a set T is a set of mappings on T, $\{\pi_g : T \to T\}_{g\in\mathcal{G}}$, such that

$$\pi_{gh}(x) = \pi_g[\pi_h(x)]\ \forall g,h \in \mathcal{G},\ x \in T$$
$$\pi_e(x) = x\ \forall x \in T. \tag{31}$$

In what follows, we denote $\pi_g(x)$ by $g \circ x$. Equation (31) can be rewritten as

$$(gh) \circ x = g \circ (h \circ x)\ \forall g,h \in \mathcal{G},\ x \in T$$
$$e \circ x = x\ \forall x \in T. \tag{32}$$

The action is called faithful if for any distinct $g,h \in \mathcal{G}$ there exists $x \in T$ such that $g \circ x \neq h \circ x$.

Let Θ be the entire set of groups $\mathcal{G}$ whose actions on $\mathcal{I}_M$, $\mathcal{I}_J$, $\mathcal{S}$, $\mathcal{S}_{\rm A}$, $\mathcal{S}_{\rm B}$, and Ω are defined and satisfy the following statements.

(i) The actions of $\mathcal{G}$ on $\mathcal{S}_{\rm A}$ and $\mathcal{S}_{\rm B}$ are expressed by

$$g \circ \hat{Q}^{(\rm A)} \equiv \hat{V}_g \hat{Q}^{(\rm A)} \hat{V}_g^\dagger\ \forall g \in \mathcal{G},\quad \hat{Q}^{(\rm A)} \in \mathcal{S}_{\rm A}$$
$$g \circ \hat{Q}^{(\rm B)} \equiv \hat{W}_g \hat{Q}^{(\rm B)} \hat{W}_g^\dagger\ \forall g \in \mathcal{G},\quad \hat{Q}^{(\rm B)} \in \mathcal{S}_{\rm B}, \tag{33}$$

where $\hat{V}_g$ and $\hat{W}_g$ are unitary or antiunitary operators on $\mathcal{H}_{\rm A}$ and $\mathcal{H}_{\rm B}$, respectively, and the dagger is the conjugate transpose operator. These group actions are not necessarily faithful. Moreover, $\{g \circ \hat{B}_m^{(\omega)}\}_m \in \mathcal{M}_{\rm B}$ holds for any $g \in \mathcal{G}$ and $\omega \in \Omega$.[1]

(ii) The action of $\mathcal{G}$ on $\mathcal{S}$ is faithful and is expressed by

$$g \circ \hat{Q} \equiv \hat{U}_g \hat{Q} \hat{U}_g^\dagger\ \forall g \in \mathcal{G},\quad \hat{Q} \in \mathcal{S}, \tag{34}$$

where

$$\hat{U}_g = \hat{V}_g \otimes \hat{W}_g. \tag{35}$$

(iii) The action of $\mathcal{G}$ on Ω, $\{g \circ \omega (\omega \in \Omega)\}_{g\in\mathcal{G}}$, is defined such that for any $g \in \mathcal{G}$ and $\omega \in \Omega$ Bob's measurement $\hat{B}^{(g\circ\omega)} \equiv \{\hat{B}_m^{(g\circ\omega)}\}_{m=0}^{M-1}$ is given by

$$\hat{B}_m^{(g\circ\omega)} = g \circ \hat{B}_{\overline{g}\circ m}^{(\omega)}\ \forall m \in \mathcal{I}_M. \tag{36}$$

Note that $\hat{B}^{(g\circ\omega)} \in \mathcal{M}_{\rm B}$ holds from statement (i).

[1] This assumption always holds if $\mathcal{M}_{\rm B}$ is the entire set of POVMs on $\mathcal{H}_{\rm B}$; otherwise, it does not hold in general. For example, if $\mathcal{H}_{\rm B}$ is a composite system and $\mathcal{M}_{\rm B}$ is the entire set of sequential measurements on $\mathcal{H}_{\rm B}$, then $\{g \circ \hat{B}_m^{(\omega)}\}_m$ might not be in $\mathcal{M}_{\rm B}$ in spite of $\{\hat{B}_m^{(\omega)}\}_m \in \mathcal{M}_{\rm B}$. In such cases, we need to appropriately set the action of $\mathcal{G}$ on $\mathcal{S}_{\rm B}$.

We stress that actions of $\mathcal{G}$ are different among different sets [see, e.g., Eqs. (33) and (34)].

We can easily verify that $\hat{V}_{gh}$ and $\hat{W}_{gh}$, respectively, equal $\hat{V}_g\hat{V}_h$ and $\hat{W}_g\hat{W}_h$ up to global phases for any $g,h \in \mathcal{G}$ and that, from Eq. (35), $\hat{U}_{gh}$ also equals $\hat{U}_g\hat{U}_h$ up to a global phase for any $g,h \in \mathcal{G}$. We can assume, without loss of generality, that $\hat{V}_e = \hat{1}_{\rm A}$ and $\hat{W}_e = \hat{1}_{\rm B}$ hold, which gives $\hat{U}_e = \hat{1}$. Since the action of $\mathcal{G}$ on $\mathcal{S}$ is faithful, $\hat{U}_g$ and $\hat{U}_h$ are not equivalent up to a global phase for any distinct $g,h \in \mathcal{G}$.

Here Θ includes various types of symmetry. As an example, in the case in which only Bob's system has a certain unitary (or antiunitary) symmetry, we can consider a group $\mathcal{G} \in \Theta$ such that $\hat{V}_g = \hat{1}_{\rm A}$ holds for any $g \in \mathcal{G}$. As another example, if Alice's and Bob's systems independently have different unitary (or antiunitary) symmetries, represented by groups $\mathcal{G}_{\rm A}$ and $\mathcal{G}_{\rm B}$, respectively, then we can consider the direct product of the groups $\mathcal{G} = \mathcal{G}_{\rm A} \times \mathcal{G}_{\rm B} \in \Theta$ such that there exist two sets of unitary (or antiunitary) operators $\{\hat{V}'_{g_{\rm A}} : g_{\rm A} \in \mathcal{G}_{\rm A}\}$ and $\{\hat{W}'_{g_{\rm B}} : g_{\rm B} \in \mathcal{G}_{\rm B}\}$ satisfying $\hat{V}_g = \hat{V}'_{g_{\rm A}}$ and $\hat{W}_g = \hat{W}'_{g_{\rm B}}$ for any $g = (g_{\rm A}, g_{\rm B}) \in \mathcal{G}$. A more complex example is given in Sec. IV C 3.

B. Symmetric properties of optimal solutions

We show that if Problem P has a certain symmetry, then there exists an optimal solution with the same type of symmetry (proof in Appendix E).

Theorem 4. Suppose that, in Problem P, there exist a group $\mathcal{G} \in \Theta$ such that

$$g \circ \hat{a}_{j,m} = \hat{a}_{g\circ j, g\circ m}\ \forall g \in \mathcal{G},\quad j \in \mathcal{I}_J,\quad m \in \mathcal{I}_M$$
$$b_j = b_{g\circ j}\ \forall g \in \mathcal{G},\quad j \in \mathcal{I}_J \tag{37}$$
$$g \circ \hat{c}_m = \hat{c}_{g\circ m}\ \forall g \in \mathcal{G},\quad m \in \mathcal{I}_M.$$

Then, as long as $\mathcal{M}_{\rm A}^\circ$ is not empty, for any $\hat{\Phi} \in \mathcal{M}_{\rm A}^\circ$, there exists $\hat{A} \in \mathcal{M}_{\rm A}^\circ$ such that $f(\hat{A}) = f(\hat{\Phi})$ and

$$g \circ \hat{A}(\omega) = \hat{A}(g \circ \omega)\ \forall g \in \mathcal{G},\quad \omega \in \Omega. \tag{38}$$

Moreover, for any $(\hat{Y},\nu) \in \mathcal{X}^\circ$, there exists $(\hat{X},\lambda) \in \mathcal{X}^\circ$ such that $s(\hat{X},\lambda) = s(\hat{Y},\nu)$ and

$$g \circ \hat{X} = \hat{X}\ \forall g \in \mathcal{G}$$
$$\lambda_j = \lambda_{g\circ j}\ \forall g \in \mathcal{G},\quad j \in \mathcal{I}_J. \tag{39}$$

In particular, there exist an optimal solution $\hat{A}$ to Problem P satisfying Eq. (38) and an optimal solution $(\hat{X},\lambda)$ to Problem DP satisfying Eq. (39).

If Eq. (38) holds, then $\hat{\Pi}^{(\hat{A})}$ has the following symmetry:

$$g \circ \hat{\Pi}_m^{(\hat{A})} = \hat{\Pi}_{g\circ m}^{(\hat{A})}. \tag{40}$$

Indeed, from Eqs. (36), (E1), and (E2) we obtain

$$g \circ \hat{\Pi}_m^{(\hat{A})} = g \circ \left[\int_\Omega \hat{A}(d\omega) \otimes \hat{B}_m^{(\omega)}\right]$$
$$= \int_\Omega [g \circ \hat{A}(d\omega)] \otimes \left[g \circ \hat{B}_m^{(\omega)}\right]$$
$$= \int_\Omega \hat{A}[d(g\circ\omega)] \otimes \hat{B}_{g\circ m}^{g\circ\omega} = \hat{\Pi}_{g\circ m}^{(\hat{A})}. \tag{41}$$

Let $\mathcal{M}^\circ_{\mathrm{A};\mathcal{G}}$ be the entire set of $\hat{A} \in \mathcal{M}^\circ_{\mathrm{A}}$ satisfying Eq. (38) and $\mathcal{X}^\circ_{\mathcal{G}}$ be the entire set of $(\hat{X},\lambda) \in \mathcal{X}^\circ$ satisfying Eq. (39). We can easily verify that $\mathcal{M}^\circ_{\mathrm{A};\mathcal{G}}$ and $\mathcal{X}^\circ_{\mathcal{G}}$ are convex. Thus, Problems P and DP remain in convex programming even if we restrict the feasible sets to $\mathcal{M}^\circ_{\mathrm{A};\mathcal{G}}$ and $\mathcal{X}^\circ_{\mathcal{G}}$, respectively.

C. Examples

1. Optimal inconclusive measurement

In the case of an inconclusive sequential measurement, substituting Eq. (5) into Eq. (37) gives

$$g \circ \hat{\rho}_r = \hat{\rho}_{g\circ r} \ \forall g \in \mathcal{G}, \quad r \in \mathcal{I}_R, \tag{42}$$

where the action of $\mathcal{G}$ on $\mathcal{I}_M = \mathcal{I}_{R+1}$ is set such that $g \circ R = R$ holds. Quantum states $\{\hat{\rho}_r\}_{r=0}^{R-1}$ are called group covariant states with respect to $\mathcal{G}$ if Eq. (37) holds (e.g., [44]). From Theorem 4, if given states are group covariant, then there exist optimal solutions to Problems P and DP with the same type of symmetry.

2. Optimal measurement in the Bayes criterion with a constraint

In the case of the Bayes criterion with a constraint, substituting Eq. (9) into Eq. (37) gives

$$\sum_{r=0}^{R-1} w_{m,r}^{(k)} g \circ \hat{\rho}_r = \sum_{r=0}^{R-1} w_{g\circ m,r}^{(k)} \hat{\rho}_r \quad \forall g \in \mathcal{G}, \quad m \in \mathcal{I}_R, \quad k \in \mathcal{I}_2. \tag{43}$$

In particular, if $w_{g\circ m,g\circ r}^{(k)} = w_{m,r}^{(k)}$ holds, then the left-hand side of Eq. (43) equals

$$\sum_{r=0}^{R-1} w_{m,r}^{(k)} g \circ \hat{\rho}_r = \sum_{r=0}^{R-1} w_{g\circ m,g\circ r}^{(k)} g \circ \hat{\rho}_r = \sum_{r'=0}^{R-1} w_{g\circ m,r'}^{(k)} g \circ \hat{\rho}_{\bar{g}\circ r'}, \tag{44}$$

where $r' \equiv g \circ r$. Thus, in this case, Eq. (42) is sufficient to satisfy Eq. (43)

3. A more specific example

As a more specific example, let us consider the problem of finding a minimum-error sequential measurement for ternary quantum states $\{\hat{\rho}_r = \frac{1}{3}\tilde{\rho}_r^{(\mathrm{A})} \otimes \tilde{\rho}_r^{(\mathrm{B})}\}_{r=0}^2$, where $\{\tilde{\rho}_r^{(\mathrm{A})}\}_{r=0}^2$ and $\{\tilde{\rho}_r^{(\mathrm{B})}\}_{r=0}^2$ are, respectively, ternary phase-shift-keyed (PSK) and amplitude-shift-keyed (ASK) coherent states. These states are given by

$$\tilde{\rho}_r^{(\mathrm{A})} = |\alpha \exp(2\pi\sqrt{-1}r/3)\rangle, \quad \tilde{\rho}_0^{(\mathrm{B})} = |0\rangle, \quad \tilde{\rho}_1^{(\mathrm{B})} = |\beta\rangle, \quad \tilde{\rho}_1^{(\mathrm{B})} = |-\beta\rangle, \tag{45}$$

with $\alpha,\beta \in \mathbf{R}$. The phase-space representation of such states is shown in Fig. 2. Since a minimum-error sequential measurement is a special case of an optimal inconclusive one, Eq. (37) reduces to Eq. (42). We consider a group $\mathcal{G} \in \Theta$ and its actions that satisfy Eq. (42).

Let $\mathcal{G}_\mathrm{A} \equiv \{p_\mathrm{A}^k, p_\mathrm{A}^k q_\mathrm{A}\}_{k\in\mathcal{I}_3}$ and $\mathcal{G}_\mathrm{B} \equiv \{p_\mathrm{B}^k, p_\mathrm{B}^k q_\mathrm{B}\}_{k\in\mathcal{I}_2}$ be dihedral groups with $|\mathcal{G}_\mathrm{A}| = 6$ and $|\mathcal{G}_\mathrm{B}| = 4$. The group $\mathcal{G}_k$ ($k \in \{\mathrm{A},\mathrm{B}\}$) is generated by a rotation p_k and a reflection q_k, which have $p_k q_k p_k = q_k$. We have $p_\mathrm{A}^3 = q_\mathrm{A}^2 = e_\mathrm{A}$ and $p_\mathrm{B}^2 = q_\mathrm{B}^2 = e_\mathrm{B}$, where e_A and e_B are, respectively, the identity elements of $\mathcal{G}_\mathrm{A}$ and $\mathcal{G}_\mathrm{B}$. We define actions of $\mathcal{G}_\mathrm{A}$ on $\mathcal{S}_\mathrm{A}$ and $\mathcal{G}_\mathrm{B}$ on $\mathcal{S}_\mathrm{B}$ as

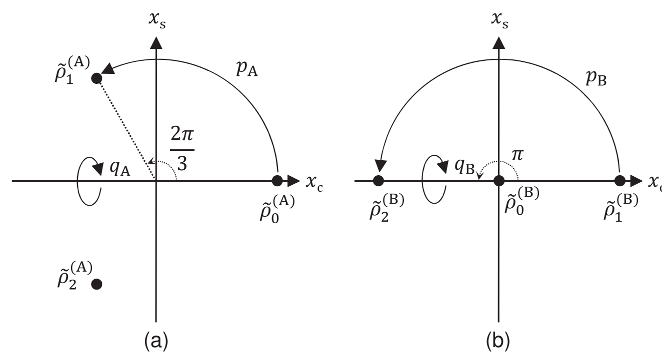

FIG. 2. Phase-space representation of (a) ternary PSK coherent states $\{\tilde{\rho}_r^{(\mathrm{A})}\}$ and (b) ternary ASK coherent states $\{\tilde{\rho}_r^{(\mathrm{B})}\}$.

$$g_\mathrm{A} \circ \hat{Q}^{(\mathrm{A})} \equiv \hat{V}_{g_\mathrm{A}} \hat{Q}^{(\mathrm{A})} \hat{V}_{g_\mathrm{A}}^\dagger \ \forall g_\mathrm{A} \in \mathcal{G}_\mathrm{A}, \quad \hat{Q}^{(\mathrm{A})} \in \mathcal{S}_\mathrm{A},$$
$$g_\mathrm{B} \circ \hat{Q}^{(\mathrm{B})} \equiv \hat{W}_{g_\mathrm{B}} \hat{Q}^{(\mathrm{B})} \hat{W}_{g_\mathrm{B}}^\dagger \ \forall g_\mathrm{B} \in \mathcal{G}_\mathrm{B}, \quad \hat{Q}^{(\mathrm{B})} \in \mathcal{S}_\mathrm{B}, \tag{46}$$

where $\hat{V}_{g_\mathrm{A}}$ and $\hat{W}_{g_\mathrm{B}}$ are, respectively, unitary (or antiunitary) operators on $\mathcal{H}_\mathrm{A}$ and $\mathcal{H}_\mathrm{B}$, satisfying $\hat{V}_{p_\mathrm{A}}^3 = \hat{V}_{q_\mathrm{A}}^2 = \hat{1}_\mathrm{A}$ and $\hat{W}_{p_\mathrm{B}}^2 = \hat{W}_{q_\mathrm{B}}^2 = \hat{1}_\mathrm{B}$. Further, $\hat{V}_{p_\mathrm{A}}$ and $\hat{W}_{p_\mathrm{B}}$, respectively, correspond to the rotation of $2\pi/3$ and π, around the origin in the phase space; $\hat{V}_{q_\mathrm{A}}$ and $\hat{W}_{q_\mathrm{B}}$, respectively, correspond to the reflection about the x_c axis. We can easily see that the following equations hold:

$$p_\mathrm{A} \circ \tilde{\rho}_r^{(\mathrm{A})} = \tilde{\rho}_{r\oplus 1}^{(\mathrm{A})}, \quad q_\mathrm{A} \circ \tilde{\rho}_r^{(\mathrm{A})} = \tilde{\rho}_{\kappa(r)}^{(\mathrm{A})},$$
$$p_\mathrm{B} \circ \tilde{\rho}_r^{(\mathrm{B})} = \tilde{\rho}_{\kappa(r)}^{(\mathrm{B})}, \quad q_\mathrm{B} \circ \tilde{\rho}_r^{(\mathrm{B})} = \tilde{\rho}_r^{(\mathrm{B})}, \tag{47}$$

where $\kappa(0) = 0$, $\kappa(1) = 2$, $\kappa(2) = 1$, and $\oplus$ denotes the addition modulo 3.

Let $e \equiv (e_\mathrm{A},e_\mathrm{B})$, $p \equiv (q_\mathrm{A},p_\mathrm{B})$, and $q \equiv (e_\mathrm{A},q_\mathrm{B})$. Also, we redefine $\hat{V}_g \equiv \hat{V}_{g_\mathrm{A}}$ and $\hat{W}_g \equiv \hat{W}_{g_\mathrm{B}}$ for $g = (g_\mathrm{A},g_\mathrm{B}) \in \mathcal{G}_\mathrm{A} \times \mathcal{G}_\mathrm{B}$. Then $\mathcal{G} \equiv \{e,p,q,pq\}$ is a group such that

$$\begin{aligned} p \circ \hat{\rho}_r &= \tfrac{1}{3}\big[q_\mathrm{A} \circ \tilde{\rho}_r^{(\mathrm{A})}\big] \otimes \big[p_\mathrm{B} \circ \tilde{\rho}_r^{(\mathrm{B})}\big] \\ &= \tfrac{1}{3}\tilde{\rho}_{\kappa(m)}^{(\mathrm{A})} \otimes \tilde{\rho}_{\kappa(m)}^{(\mathrm{B})} = \hat{\rho}_{\kappa(m)}, \\ q \circ \hat{\rho}_r &= \tfrac{1}{3}\big[e_\mathrm{A} \circ \tilde{\rho}_r^{(\mathrm{A})}\big] \otimes \big[q_\mathrm{B} \circ \tilde{\rho}_r^{(\mathrm{B})}\big] \\ &= \tfrac{1}{3}\tilde{\rho}_r^{(\mathrm{A})} \otimes \tilde{\rho}_r^{(\mathrm{B})} = \hat{\rho}_r. \end{aligned} \tag{48}$$

Note that the action of $\mathcal{G}$ on $\mathcal{S}_\mathrm{A}$ is not faithful; indeed, both $\hat{V}_e$ and $\hat{V}_q$ are identical to $\hat{V}_{e_\mathrm{A}}$. Let us define an action of $\mathcal{G}$ on $\mathcal{I}_M$ such that $p \circ m = \kappa(m)$ and $q \circ m = m$; then Eq. (42) holds.

From Theorem 4 there exists $\hat{A} \in \mathcal{M}_\mathrm{A}$ satisfying Eq. (38). From Eq. (40), $\hat{\Pi}^{(\hat{A})} = \{\hat{\Pi}_m^{(\hat{A})}\}_{m=0}^2$ with such $\hat{A}$ has the following symmetry:

$$\hat{U}_p \hat{\Pi}_m^{(\hat{A})} \hat{U}_p^\dagger = \hat{\Pi}_{\kappa(m)}^{(\hat{A})},$$
$$\hat{U}_q \hat{\Pi}_m^{(\hat{A})} \hat{U}_q^\dagger = \hat{\Pi}_m^{(\hat{A})}, \tag{49}$$

where $\hat{U}_p = \hat{V}_{q_\mathrm{A}} \otimes \hat{W}_{p_\mathrm{B}}$ and $\hat{U}_q = \hat{1}_\mathrm{A} \otimes \hat{W}_{q_\mathrm{B}}$ from Eq. (35). Moreover, from Eq. (39) there exists $\hat{X} \in \mathcal{X}^\circ$ commuting with $\hat{V}_{q_\mathrm{A}}$.

Note that, in this example, neither $\hat{A}$ nor $\hat{X}$ has the symmetry expressed by $\{p_{\rm A}^k\}_{k\in\mathcal{I}_3}$, while the states $\{\tilde{\rho}_r^{(\rm A)}\}$ have this symmetry. The reason is that the states $\{\tilde{\rho}_r^{(\rm B)}\}$ do not have this symmetry.

V. GENERALIZED MINIMAX SOLUTION

In the minimax strategy for a quantum state discrimination problem, prior probabilities are unknown and the task is to maximize the worst case of the objective function (such as the average success probability) over all prior probabilities. This strategy has been investigated in several studies [18–20,45,51], whose generalized version appears in Ref. [30]. In this section, we consider a sequential-measurement version of the generalized minimax problem. In a manner similar to the method reported by Ref. [30], we can provide necessary and sufficient conditions for a minimax solution to the sequential-measurement version of the problem. In what follows, we discuss properties that a minimax solution has.

A. Formulation

Let us consider $K \geqslant 1$ objective functions $f_0(\hat{A}),\dots,f_{K-1}(\hat{A})$ expressed as

$$f_k(\hat{A}) \equiv \sum_{m=0}^{M-1} {\rm Tr}\big[\hat{c}_{k,m}\hat{\Pi}_m^{(\hat{A})}\big] + d_k, \tag{50}$$

where $\hat{c}_{k,m}\in\mathcal{S}$ and $d_k\in\mathbf{R}$. Also, let $\mathcal{P}$ be the entire set of collections of K non-negative real numbers $\mu\equiv\{\mu_k\}_{k=0}^{K-1}\in\mathbf{R}_+^K$, satisfying $\sum_{k=0}^{K-1}\mu_k=1$. Here $\mu\in\mathcal{P}$ can be interpreted as a probability distribution. Let $F(\mu,\hat{A})$ be the objective function defined by

$$F(\mu,\hat{A}) \equiv \sum_{k=0}^{K-1}\mu_k f_k(\hat{A}) \tag{51}$$

and $\mathcal{M}_{\rm A}^\circ$ be the set defined by Eq. (3). We investigate the problem of finding a POVM $\hat{A}\in\mathcal{M}_{\rm A}^\circ$ that maximizes the worst-case value of $F(\mu,\hat{A})$ over $\mu\in\mathcal{P}$. This problem can be formulated as follows:

$$\text{Problem P}_{\rm m}\text{:}\ \ \text{maximize}\ \min_{\mu\in\mathcal{P}} F(\mu,\hat{A}) \quad \text{subject to}\ \ \hat{A}\in\mathcal{M}_{\rm A}^\circ. \tag{52}$$

Let $F^\star$ be the optimal value of Problem $\rm P_m$. We call a pair $(\mu^\star,\hat{A}^\star)$ with

$$\mu^\star \in \mathop{\rm argmin}_{\mu\in\mathcal{P}} \max_{\hat{A}\in\mathcal{M}_{\rm A}^\circ} F(\mu,\hat{A}),$$
$$\hat{A}^\star \in \mathop{\rm argmax}_{\hat{A}\in\mathcal{M}_{\rm A}^\circ} \min_{\mu\in\mathcal{P}} F(\mu,\hat{A})$$

a minimax solution to Problem $\rm P_m$. Such $\hat{A}^\star$, which we will call a minimax POVM, is obviously an optimal solution to Problem $\rm P_m$.

B. Example

The minimax strategy in the minimum-error criterion, which has been investigated in Ref. [18], is a simple example. In this strategy, $\mu\in\mathcal{P}$ with $K=R$ can be regarded as prior probabilities of the states $\{\tilde{\rho}_r : r\in\mathcal{I}_R\}$. The aim is to find a sequential measurement $\hat{\Pi}^{(\hat{A})}$ that maximizes the worst-case average success probability $P_{\rm S}(\mu,\hat{A})$ over $\mu\in\mathcal{P}$. Here $P_{\rm S}(\mu,\hat{A})$ is written by

$$P_{\rm S}(\mu,\hat{A}) = \sum_{k=0}^{R-1}\mu_k{\rm Tr}\big[\tilde{\rho}_k\hat{\Pi}_k^{(\hat{A})}\big]. \tag{53}$$

This problem is equivalent to Problem $\rm P_m$ with

$$M=K=R,\quad J=0,$$
$$\hat{c}_{k,m} = \tilde{\rho}_m\delta_{k,m},\quad d_k=0. \tag{54}$$

C. Properties of a minimax solution

We first show the following remark.

Remark 1 (minimax theorem). If $\mathcal{M}_{\rm A}^\circ$ is not empty, then there exists a minimax solution $(\mu^\star,\hat{A}^\star)$ to Problem $\rm P_m$ and it satisfies

$$\min_{\mu\in\mathcal{P}}\max_{\hat{A}\in\mathcal{M}_{\rm A}^\circ} F(\mu,\hat{A}) = F(\mu^\star,\hat{A}^\star) = \max_{\hat{A}\in\mathcal{M}_{\rm A}^\circ}\min_{\mu\in\mathcal{P}} F(\mu,\hat{A}). \tag{55}$$

Proof. $\mathcal{M}_{\rm A}^\circ$ and $\mathcal{P}$ are closed convex sets. $F(\mu,\hat{A})$ is a continuous convex function of μ for fixed $\hat{A}$ and a continuous concave function of $\hat{A}$ for fixed μ. Thus, the minimax theorem holds (e.g., [52]); that is to say, there exists a minimax solution $(\mu^\star,\hat{A}^\star)$ to Problem $\rm P_m$, which satisfies Eq. (55). ■

A minimax solution to Problem $\rm P_m$ can be characterized by a saddle point; i.e., $(\mu^\star,\hat{A}^\star)$ is a minimax solution if and only if, for any $\mu\in\mathcal{P}$ and $\hat{A}\in\mathcal{M}_{\rm A}^\circ$, $(\mu^\star,\hat{A}^\star)$ satisfies [52]

$$F(\mu^\star,\hat{A}) \leqslant F(\mu^\star,\hat{A}^\star) \leqslant F(\mu,\hat{A}^\star). \tag{56}$$

Let

$$F^\star(\mu) \equiv \max_{\hat{A}\in\mathcal{M}_{\rm A}^\circ} F(\mu,\hat{A}). \tag{57}$$

From Eq. (56), $F^\star(\mu^\star)=F(\mu^\star,\hat{A}^\star)$ holds.

Let $\overline{c_m}(\mu)\equiv\sum_{k=0}^{K-1}\mu_k\hat{c}_{k,m}$ and $\overline{d}(\mu)\equiv\sum_{k=0}^{K-1}\mu_k d_k$; then we find that the problem of finding $F^\star(\mu)$ for a fixed $\mu\in\mathcal{P}$ is reduced to Problem P, as shown in the following remark.

Remark 2. Let $\overline{f^\star}(\mu)$ be the optimal value of Problem P with $\hat{c}_m=\overline{c_m}(\mu)$; then $F^\star(\mu)=\overline{f^\star}(\mu)+\overline{d}(\mu)$ holds.

Proof.

$$\begin{aligned} F^\star(\mu) &= \max_{\hat{A}\in\mathcal{M}_{\rm A}^\circ} F(\mu,\hat{A}) \\ &= \max_{\hat{A}\in\mathcal{M}_{\rm A}^\circ}\sum_{k=0}^{K-1}\mu_k\left[\sum_{m=0}^{M-1}{\rm Tr}\big[\hat{c}_{k,m}\hat{\Pi}_m^{(\hat{A})}\big]+d_k\right] \\ &= \max_{\hat{A}\in\mathcal{M}_{\rm A}^\circ}\sum_{m=0}^{M-1}{\rm Tr}\big[\overline{c_m}(\mu)\hat{\Pi}_m^{(\hat{A})}\big]+\overline{d}(\mu) \\ &= \overline{f^\star}(\mu)+\overline{d}(\mu). \end{aligned} \tag{58}$$

■

Theorem 5. Assume $\mu^\star\in\mathcal{P}$ and $\hat{A}^\star\in\mathcal{M}_{\rm A}^\circ$. The following statements are all equivalent.

(a) $(\mu^\star,\hat{A}^\star)$ is a minimax solution to Problem $\rm P_m$.

(b) The following equation holds:

$$f_k(\hat{A}^\star) \geqslant F^\star(\mu^\star)\ \forall k\in\mathcal{I}_K. \tag{59}$$

(c) The following equations hold:

$$F^\star(\mu^\star) = F(\mu^\star,\hat{A}^\star),$$
$$f_k(\hat{A}^\star) \geqslant f_{k'}(\hat{A}^\star)\ \forall k,\quad k'\in\mathcal{I}_K \text{ such that } \mu_{k'}^\star>0. \tag{60}$$

The proof is the same as Theorem 3 of Ref. [30]. ■

Theorem 6. Let us consider the optimization problem

$$\text{maximize} \quad f_{\min}(\hat{A}) \equiv \min_{k\in\mathcal{I}_K} f_k(\hat{A})$$
$$\text{subject to} \quad \hat{A} \in \mathcal{M}_{\rm A}^{\circ} \tag{61}$$

with $\hat{A}$. An optimal solution to the problem given by Eq. (61) is equivalent to a minimax POVM of Problem $\rm P_m$.

The proof is the same as Theorem 4 of Ref. [30]. ■

D. Minimax problem with symmetry

Similar to Theorem 4, if Problem $\rm P_m$ has a certain symmetry, then there exists a minimax solution with the same type of symmetry, as stated in the following theorem (proof in Appendix F).

Theorem 7. Suppose that, in Problem $\rm P_m$, there exist a group $\mathcal{G} \in \Theta$ and its action on $\mathcal{I}_K$ such that

$$\begin{aligned} g \circ \hat{a}_{j,m} &= \hat{a}_{g\circ j, g\circ m} \ \forall g \in \mathcal{G}, \quad j \in \mathcal{I}_J,\ m \in \mathcal{I}_M \\ b_j &= b_{g\circ j} \ \forall g \in \mathcal{G}, \quad j \in \mathcal{I}_J \\ g \circ \hat{c}_{k,m} &= \hat{c}_{g\circ k, g\circ m} \ \forall g \in \mathcal{G}, \quad k \in \mathcal{I}_K,\ m \in \mathcal{I}_M \\ d_k &= d_{g\circ k} \ \forall g \in \mathcal{G}, \quad k \in \mathcal{I}_K. \end{aligned} \tag{62}$$

Then, as long as $\mathcal{M}_{\rm A}^{\circ}$ is not empty, there exists a minimax solution $(\mu^\star, \hat{A}^\star)$ such that

$$\begin{aligned} \mu_k^\star &= \mu_{g\circ k}^\star \ \forall g \in \mathcal{G}, \quad k \in \mathcal{I}_K \\ g \circ \hat{A}^\star(\omega) &= \hat{A}^\star(g \circ \omega) \ \forall g \in \mathcal{G}, \quad \omega \in \Omega. \end{aligned} \tag{63}$$

An an example, in the case of the minimum-error criterion described in Sec. V B, Eq. (62) is equivalent to

$$g \circ \tilde{\rho}_m \delta_{k,m} = \tilde{\rho}_{g\circ m} \delta_{g\circ k, g\circ m} \ \forall g \in \mathcal{G}, \quad k,m \in \mathcal{I}_M. \tag{64}$$

VI. EXAMPLES

In this section, we provide two examples of deriving closed-form analytical expressions for optimal solutions.

A. Optimal inconclusive sequential measurement for double trine states

We derive an optimal solution to the problem of Eq. (4) for double trine states with equal prior probabilities. Note that, in the cases of $p_{\rm I} = 0$ (corresponding to a minimum-error sequential measurement) and $p_{\rm I} = 1/2$ (corresponding to an optimal unambiguous sequential measurement), optimal solutions are given in Refs. [53,54], respectively. However, an optimal solution in the case of $0 < p_{\rm I} < 1/2$ has not been obtained so far and would be difficult to obtain with the method described in these references.

Double trine states with equal prior probabilities can be expressed by $\{\hat{\rho}_m \equiv \frac{1}{3} |\psi_m\rangle\langle\psi_m|\}_{m=0}^2$ with

$$\begin{aligned} |\psi_m\rangle &\equiv |\phi_m\rangle \otimes |\phi_m\rangle, \\ |\phi_m\rangle &\equiv \cos\frac{2\pi m}{3}|0\rangle + \sin\frac{2\pi m}{3}|1\rangle. \end{aligned} \tag{65}$$

Here $\{|\phi_m\rangle\}$ has the symmetry of $|\phi_m\rangle = \hat{V}_{\rm rot}^k |\phi_{m\ominus k}\rangle$, where

$$\hat{V}_{\rm rot} \equiv -\frac{1}{2}\hat{1} + \frac{\sqrt{3}}{2}(|1\rangle\langle 0| - |0\rangle\langle 1|), \tag{66}$$

which is a unitary operator corresponding to a rotation of $\frac{2\pi}{3}$ and $\ominus$ denotes the subtraction modulo 3. Also, since $\langle k|\phi_m\rangle$ ($k \in \{0,1\}$) is real, $\hat{V}_{\rm conj}|\phi_m\rangle = |\phi_m\rangle$ holds, where $\hat{V}_{\rm conj}$ is the antiunitary operator of complex conjugation in the basis $\{|0\rangle, |1\rangle\}$.[2]

First, we derive an optimal solution $(\hat{X}^\star,\lambda^\star)$ to the problem of Eq. (22). Assume, without loss of generality, that $\hat{X}^\star$ commutes with $\hat{V}_{\rm rot}$ and $\hat{V}_{\rm conj}$ (see Theorem 4); then it follows that such $\hat{X}^\star$ must be proportional to $\hat{1}_{\rm A}$. After some computations, we obtain an optimal solution $(\hat{X}^\star,\lambda^\star)$ as follows (see Appendix G):

$$\begin{aligned} \hat{X}^\star &= \left(\frac{1}{2} + \frac{3 - 2p_{\rm I}}{4\sqrt{3 - 4p_{\rm I}}}\right)\hat{1}_{\rm A}, \\ \lambda^\star &= \frac{1}{2} + \frac{1}{2\sqrt{3 - 4p_{\rm I}}}. \end{aligned} \tag{67}$$

Thus, the average success probability of an optimal inconclusive sequential measurement $P_{\rm S}^\star$, which is equivalent to the optimal value $s(\hat{X}^\star,\lambda^\star)$, is given by

$$\begin{aligned} P_{\rm S}^\star &= s(\hat{X}^\star,\lambda^\star) = {\rm Tr}\,\hat{X}^\star - \lambda^\star p_{\rm I} \\ &= \tfrac{1}{2}(1 - p_{\rm I}) + \tfrac{1}{4}\sqrt{3 - 4p_{\rm I}}. \end{aligned} \tag{68}$$

When $p_{\rm I} = 1/2$, $P_{\rm S}^\star + p_{\rm I} = 1$ holds; i.e., the average error probability $1 - P_{\rm S}^\star - p_{\rm I}$ is zero. This indicates that there exists an unambiguous sequential measurement with the average inconclusive probability of $1/2$. Since the case of $p_{\rm I} > 1/2$ is trivial, assume $0 \leqslant p_{\rm I} \leqslant 1/2$ (in this case, $\frac{1}{2} + \frac{1}{2\sqrt{3}} \leqslant \lambda^\star \leqslant 1$ holds).

Next we derive an optimal sequential measurement. Let $|\phi_m^\perp\rangle$ be the vector expressed by

$$|\phi_m^\perp\rangle \equiv -\sin\frac{2\pi m}{3}|0\rangle + \cos\frac{2\pi m}{3}|1\rangle, \tag{69}$$

which satisfies $\langle\phi_m^\perp|\phi_m\rangle = 0$ and $|\phi_m^\perp\rangle = \hat{V}_{\rm rot}^k|\phi_{m\ominus k}^\perp\rangle$. From the discussion in Appendix G and the symmetry of $\{|\phi_m\rangle\}$, $\hat{X}^\star - \hat{\sigma}_\omega(\lambda^\star)$ is rank one [i.e., the largest eigenvalue of $\hat{\sigma}_\omega(\lambda^\star)$ is $\upsilon(\lambda^\star)$, which is defined in Appendix G] if and only if $\{\hat{B}_m^{(\omega)}\}_{m=0}^3$ is expressed as

$$\hat{B}_m^{(\omega)} = \hat{B}_m^{(\omega_k)} \equiv \begin{cases} \hat{V}_{\rm rot}^k \hat{B}_{m\ominus k}^\bullet \hat{V}_{\rm rot}^{-k}, & m < 3 \\ \frac{4}{3}p_{\rm I}|\phi_k^\perp\rangle\langle\phi_k^\perp|, & m = 3, \end{cases} \tag{70}$$

where $\{\hat{B}_m^\bullet\}$ is given by Eq. (G12) with $\alpha = 4p_{\rm I}/3$ and $\omega_k \in \Omega$ ($k \in \mathcal{I}_3$) is an index corresponding to the POVM $\hat{B}^{(\omega_k)} \equiv \{\hat{B}_m^{(\omega_k)}\}$ defined by Eq. (70). In Eq. (70), we use

$$\hat{V}_{\rm rot}^k \hat{B}_3^\bullet \hat{V}_{\rm rot}^{-k} = \alpha \hat{V}_{\rm rot}^k |\phi_0^\perp\rangle\langle\phi_0^\perp| \hat{V}_{\rm rot}^{-k} = \tfrac{4}{3}p_{\rm I}|\phi_k^\perp\rangle\langle\phi_k^\perp|. \tag{71}$$

---

[2] Our discussion in Sec. IV can be used when considering a dihedral group with order 6, $\mathcal{G} = \{p^k, p^k q\}_{k\in\mathcal{I}_3}$, which is generated by a rotation p and a reflection q with $pqp = q$. To be concrete, let $\hat{V}_{p^k q^l} = \hat{V}_{\rm rot}^k \hat{V}_{\rm conj}^l$ for any $k \in \mathcal{I}_3$ and $l \in \mathcal{I}_2$ and let $\hat{U}_g = \hat{V}_g \otimes \hat{V}_g$; then we can consider group actions of $\mathcal{G}$. Note that double trine states also have the symmetry of $(|0\rangle\langle 0| - |1\rangle\langle 1|)|\phi_m\rangle = |\phi_{\kappa(m)}\rangle$ [$\kappa(0) = 0$, $\kappa(1) = 2$, and $\kappa(2) = 1$]; however, we do not need this symmetry to obtain their optimal sequential measurement.

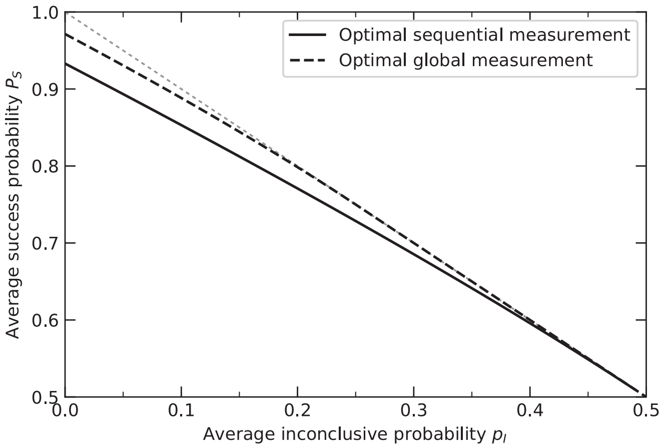

FIG. 3. Average success probabilities of optimal sequential and global measurements for double trine states with equal prior probabilities.

Using the fact that, from Eq. (25), the support of $\hat{A}(\omega)$ is included in the kernel of $\hat{X}^\star - \hat{\sigma}_\omega(\lambda^\star)$, $\hat{A}$ can be obtained in the following way. When $\omega = \omega_k$, since Eq. (G4) with $\theta = \frac{2\pi k}{3}$ holds, $\hat{V}_{\rm rot}^k |1\rangle = |\phi_k^\perp\rangle$ is in the kernel of $\hat{X}^\star - \hat{\sigma}_\omega(\lambda^\star)$. Then $\hat{A}(\omega_k)$ must be proportional to $|\phi_m^\perp\rangle\langle\phi_m^\perp|$. When $\omega \neq \omega_k$, the kernel of $\hat{X}^\star - \hat{\sigma}_\omega(\lambda^\star)$ is $\{0\}$, which implies that $\hat{A}(\omega) = 0$. Thus, an optimal inconclusive sequential measurement $\hat{\Pi}^\star$ is expressed by $\hat{\Pi}^\star = \hat{\Pi}^{(\hat{A})}$, where

$$\hat{A}(\omega) = \begin{cases} \frac{2}{3}|\phi_k^\perp\rangle\langle\phi_k^\perp| & \text{for } \omega = \omega_k (k \in \mathcal{I}_3), \\ 0 & \text{otherwise} \end{cases} \quad (72)$$

holds for any $\omega \in \Omega$. It follows that $\hat{A}$ is a POVM with three outcomes $\{\omega_k\}_{k=0}^2$. From Eqs. (70) and (72), $\hat{\Pi}^\star$ can be rewritten as

$$\hat{\Pi}_m^\star = \begin{cases} \frac{2}{3}\sum_{k=0, k\neq m}^2 |\phi_k^\perp\rangle\langle\phi_k^\perp| \otimes \hat{V}_{\rm rot}^k \hat{B}_{m\ominus k}^\bullet \hat{V}_{\rm rot}^{-k}, & m < 3 \\ \frac{8}{9} p_{\rm I} \sum_{k=0}^2 |\phi_k^\perp\rangle\langle\phi_k^\perp| \otimes |\phi_k^\perp\rangle\langle\phi_k^\perp|, & m = 3. \end{cases} \quad (73)$$

Figure 3 shows the average success probabilities of optimal measurements with and without the restriction that only sequential measurements are allowed. Note that the average success probability of an optimal inconclusive global measurement can be computed by the method described in Ref. [55]. The average error probability is zero when $p_{\rm I} \geqslant 1/2$ and $p_{\rm I} \geqslant 1/4$ in the cases of optimal inconclusive sequential and global measurements, respectively.

B. Minimax solution for symmetric states

Let us consider the following N^2 states $\{\tilde{\rho}_{m,n}\}_{m,n=0}^{N-1}$:

$$\tilde{\rho}_{m,n} \equiv |\psi_{m,n}\rangle\langle\psi_{m,n}|, \quad |\psi_{m,n}\rangle \equiv |a_n^{(m)}\rangle \otimes |m\rangle, \quad (74)$$

where N is prime. Here $\{|m\rangle\}_{m=0}^{N-1}$ is an orthonormal basis (ONB) in $\mathcal{H}_{\rm B}$ with $\dim \mathcal{H}_{\rm B} = N$. For each $m \in \mathcal{I}_N$, $\{|a_n^{(m)}\rangle\}_{n=0}^{N-1}$ is also an ONB in $\mathcal{H}_{\rm A}$ with $\dim \mathcal{H}_{\rm A} = N$. A set of ONBs $\{|a_n^{(m)}\rangle\}_{m,n=0}^{N-1}$ constitutes so-called mutually unbiased bases [56], which satisfy $|\langle a_n^{(m)}|a_{n'}^{(m')}\rangle| = 1/\sqrt{N}$ ($\forall n,n' \in \mathcal{I}_N$) for any distinct $m,m' \in \mathcal{I}_N$. More concretely, the ONB $\{|a_n^{(m)}\rangle\}_n$ is the eigenbasis of the operator $\mathcal{S}_{\rm X}\mathcal{S}_{\rm Z}^m$ for each $m \in \mathcal{I}_N$, where $\mathcal{S}_{\rm X}$ and $\mathcal{S}_{\rm Z}$ are generalized Pauli operators expressed by

$$\mathcal{S}_{\rm X} \equiv \sum_{l=0}^{N-1} |l \oplus 1\rangle\langle l|, \quad \mathcal{S}_{\rm Z} \equiv \sum_{l=0}^{N-1} \tau^n |l\rangle\langle l|, \quad (75)$$

with $\tau \equiv \exp(2\pi\sqrt{-1}/N)$ and an ONB $\{|l\rangle\}_{l=0}^{N-1}$ in $\mathcal{H}_{\rm A}$. Here $\oplus$ denotes the addition modulo N. Note that, in the case of $N = 2$, an analytical expression for a minimum-error sequential measurement has been derived in Ref. [2]. Here we will provide an analytical minimax solution in the minimum-error criterion [i.e., an optimal solution to Problem $\rm P_m$ with Eq. (54)] for the states $\{\tilde{\rho}_{m,n}\}_{m,n}$ with $N \geqslant 3$. We can assume without loss of generality that $|a_n^{(m)}\rangle$ is the eigenbasis of $\mathcal{S}_{\rm X}\mathcal{S}_{\rm Z}^m$ with the eigenvalue τ^n, which can be expressed as (e.g., [57])

$$|a_n^{(m)}\rangle = \frac{1}{\sqrt{N}} \sum_{l=0}^{N-1} \tau^{-nl+ml(l-1)/2} |l\rangle. \quad (76)$$

Let us consider the symmetry of the states. Let $\mathcal{G} \in \Theta$ be the group generated by three elements $g_{\rm X}$, $g_{\rm Z}$, and $g_{\rm M}$ whose action on $\mathcal{S}$ is written by Eq. (34) with

$$\hat{U}_{g_{\rm X}} \equiv \mathcal{S}_{\rm X} \otimes \hat{1}_{\rm B}, \quad \hat{U}_{g_{\rm Z}} \equiv \mathcal{S}_{\rm Z} \otimes \hat{1}_{\rm B},$$
$$\hat{U}_{g_{\rm M}} \equiv \left(\sum_{l=0}^{N-1} \tau^{l(l-1)/2} |l\rangle\langle l|\right) \otimes \left(\sum_{m=0}^{N-1} |m \oplus 1\rangle\langle m|\right). \quad (77)$$

From Eqs. (74), (76), and (77) we can easily verify

$$g_{\rm X} \circ \tilde{\rho}_{m,n} = \tilde{\rho}_{m,n\oplus m}, \quad g_{\rm Z} \circ \tilde{\rho}_{m,n} = \tilde{\rho}_{m,n\ominus 1}, \quad g_{\rm M} \circ \tilde{\rho}_{m,n} = \tilde{\rho}_{m\oplus 1,n}, \quad (78)$$

where $\ominus$ denotes the subtraction modulo N. Thus, if we define the action of $\mathcal{G}$ on $\mathcal{I}_{N^2}$ such that $g_{\rm X} \circ (m,n) = (m,n \oplus m)$, $g_{\rm Z} \circ (m,n) = (m,n \ominus 1)$, and $g_{\rm M} \circ (m,n) = (m \oplus 1,n)$, then Eq. (64) holds. Therefore, from Theorem 7 there exists an optimal solution $(\mu^\star,\hat{A}^\star)$ to Problem $\rm P_m$ such that $\mu_{m,n}^\star = \frac{1}{N^2}$ holds for any $m,n \in \mathcal{I}_N$. Here $\hat{\Pi}^{(\hat{A}^\star)}$ is a minimum-error sequential measurement for $\{\tilde{\rho}_{m,n}\}_{m,n}$ with equal prior probabilities, which can be obtained from Eq. (6) with $r = (m,n)$, $R = N^2$, and $\hat{\rho}_{m,n} = \frac{1}{N^2}\tilde{\rho}_{m,n}$. In what follows, we try to obtain the minimum-error measurement $\hat{\Pi}^{(\hat{A}^\star)}$.

First, we obtain an optimal solution $\hat{X}^\star$ to the dual problem given by Eq. (22) with $p_{\rm I} = 0$. From Theorem 4 we can assume $\mathcal{S}_{\rm X}\hat{X}^\star\mathcal{S}_{\rm X}^\dagger = \hat{X}^\star = \mathcal{S}_{\rm Z}\hat{X}^\star\mathcal{S}_{\rm Z}^\dagger$, i.e., $\hat{X}^\star$ commutes with $\mathcal{S}_{\rm X}$ and $\mathcal{S}_{\rm Z}$. On the other hand, since $\mathcal{S}_{\rm X}$ and $\mathcal{S}_{\rm Z}$ do not share any eigenvector, any operator commuting with $\mathcal{S}_{\rm X}$ and $\mathcal{S}_{\rm Z}$ is proportional to $\hat{1}_{\rm A}$. Thus, we have $\hat{X}^\star = c^\star\hat{1}_{\rm A}$ with a constant $c^\star \in \mathbf{R}$. Substituting this into Eq. (22), we obtain the problem

$$\text{minimize } c \quad \text{subject to } c\hat{1}_{\rm A} \geqslant \hat{\sigma}_\omega \ \forall \omega \in \Omega \quad (79)$$

with variable $c \in \mathbf{R}$, whose optimal value is $c^\star$. Thus, it follows that $c^\star$ equals the maximum of the largest eigenvalues of $\hat{\sigma}_\omega$, i.e.,

$$c^\star = \max\{\langle\phi|\hat{\sigma}_\omega|\phi\rangle : |\phi\rangle \in \mathcal{H}_\mathrm{A}, \langle\phi|\phi\rangle = 1, \omega \in \Omega\}. \quad (80)$$

The constant $c^\star$ can be derived from this equation as follows. Substituting Eq. (74) into Eq. (23), we have

$$\hat{\sigma}_\omega = \frac{1}{N^2}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} p_{m,n}^{(\omega)} \left|a_n^{(m)}\right\rangle\!\left\langle a_n^{(m)}\right|, \quad (81)$$

where

$$p_{m,n}^{(\omega)} \equiv \langle m|\hat{B}_{m,n}^{(\omega)}|m\rangle \quad (82)$$

and $\{\hat{B}_{m,n}^{(\omega)}\}_{m,n=0}^{N-1}$ is the POVM on $\mathcal{H}_\mathrm{B}$ corresponding to $\omega \in \Omega$. From Eq. (81), we have that for any normal vector $|\phi\rangle \in \mathcal{H}_\mathrm{A}$,

$$\begin{aligned}\langle\phi|\hat{\sigma}_\omega|\phi\rangle &= \frac{1}{N^2}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} p_{m,n}^{(\omega)} \left|\left\langle\phi\middle|a_k^{(m)}\right\rangle\right|^2 \\ &\leqslant \frac{1}{N^2}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} p_{m,n}^{(\omega)} \left|\left\langle\phi\middle|a_{\kappa(m)}^{(m)}\right\rangle\right|^2 \\ &\leqslant \frac{1}{N^2}\sum_{m=0}^{N-1} \left|\left\langle\phi\middle|a_{\kappa(m)}^{(m)}\right\rangle\right|^2 \\ &= \frac{1}{N^2}\langle\phi|\hat{\Gamma}|\phi\rangle, \end{aligned} \quad (83)$$

where $\kappa(m)$ is a function of m such that

$$\kappa(m) \in \operatorname*{arg\,max}_{n\in\mathcal{I}_N} \left|\left\langle\phi\middle|a_n^{(m)}\right\rangle\right| \ \forall m \in \mathcal{I}_N \quad (84)$$

and

$$\hat{\Gamma} \equiv \sum_{m=0}^{N-1} \left|a_{\kappa(m)}^{(m)}\right\rangle\!\left\langle a_{\kappa(m)}^{(m)}\right|. \quad (85)$$

The third line of Eq. (83) follows from $\sum_{n=0}^{N-1} p_{m,n}^{(\omega)} \leqslant \langle m|\hat{1}_\mathrm{B}|m\rangle = 1$, which is given by $\sum_{n=0}^{N-1} \hat{B}_{m,n}^{(\omega)} \leqslant \hat{1}_\mathrm{B}$.

Due to the symmetry of the states, we can here assume $\kappa(m) = 0$ for each $m \in \mathcal{I}_N$ without loss of generality. Substituting Eq. (76) into Eq. (85), and with some algebra, we can obtain

$$\hat{\Gamma} = \left|\frac{N+1}{2}\right\rangle\!\left\langle\frac{N+1}{2}\right| + 2\sum_{m=0}^{(N-1)/2-1} |\nu_m\rangle\langle\nu_m|, \quad (86)$$

where $|\nu_m\rangle$ is the normal vector defined by

$$|\nu_m\rangle \equiv \begin{cases}(|0\rangle + |1\rangle)/\sqrt{2}, & m = 0\\ (|m+1\rangle + |N-m\rangle)/\sqrt{2}, & m > 0.\end{cases} \quad (87)$$

Equation (86) indicates that the largest eigenvalue of $\hat{\Gamma}$ is 2 and thus, from Eq. (83), the maximum value of $\langle\phi|\hat{\sigma}_\omega|\phi\rangle$ is $\frac{2}{N^2}$, which gives $c^\star = \frac{2}{N^2}$ (i.e., $\hat{X}^\star = \frac{2}{N^2}\hat{1}_\mathrm{A}$). Therefore, we obtain $P_\mathrm{S}(\hat{A}^\star) = \mathrm{Tr}\,\hat{X}^\star = \frac{2}{N}$.[3]

Next we obtain a minimum-error sequential measurement. From Eq. (25), if $\hat{A}(\omega) \neq 0$, then at least one of the eigenvalues of $\hat{X}^\star - \hat{\sigma}_\omega \geqslant 0$ is zero; i.e., $\hat{\sigma}_\omega$ has the eigenvalue $\frac{2}{N^2}$. This implies that the equality in Eq. (83) holds when $|\phi\rangle = |u\rangle$, where $|u\rangle$ is a normalized eigenvector corresponding to the largest eigenvalue of $\hat{\sigma}_\omega$. We consider the case $p_{m,n}^{(\omega)} = \delta_{n,\kappa(m)}$, where $\kappa(m)$ satisfies Eq. (84) with $|\phi\rangle = |u\rangle$, which is sufficient for the equality in Eq. (83) with $|\phi\rangle = |u\rangle$. In this case, $\hat{B}_{m,n}^{(\omega)}$ can be expressed as

$$\hat{B}_{m,n}^{(\omega)} = \delta_{n,\kappa(m)}|m\rangle\langle m|. \quad (88)$$

We can easily verify that $\hat{A}(\omega)$ written by

$$\hat{A}(\omega) = \gamma\,|u\rangle\langle u|, \quad (89)$$

with $\gamma > 0$, satisfies Eq. (25). These conditions help us to find a minimum-error sequential measurement.

Let $\kappa(m) = t \oplus ms$ ($s,t \in \mathcal{I}_N$) and $\omega_{s,t} \in \Omega$ be the corresponding index; then Eq. (88) gives

$$\hat{B}_{m,n}^{(\omega_{s,t})} = \delta_{n,t\oplus ms}|m\rangle\langle m|. \quad (90)$$

From Eq. (81) we have

$$\hat{\sigma}_{\omega_{s,t}} = \frac{1}{N^2}\sum_{m=0}^{N-1}\left|a_{t\oplus ms}^{(m)}\right\rangle\!\left\langle a_{t\oplus ms}^{(m)}\right|. \quad (91)$$

The following is a normalized eigenvector corresponding to the largest eigenvalue $\frac{2}{N^2}$ of $\hat{\sigma}_{\omega_{s,t}}$:

$$|u_{s,t}\rangle \equiv \frac{1}{\sqrt{2N}}\sum_{j=0}^{N-1}\tau^{js(s+1)/2}\left|a_{t\oplus js}^{(j)}\right\rangle. \quad (92)$$

In this case, we can see that Eq. (84) holds with $|\phi\rangle = |u\rangle$. We choose $\hat{A}^\star$ such that

$$\hat{A}^\star(\omega_{s,t}) \equiv \frac{1}{N}|u_{s,t}\rangle\langle u_{s,t}| \quad (93)$$

and $\hat{A}^\star(\omega) \equiv 0$ when $\omega \in \Omega$ is not in $\{\omega_{s,t}\}_{s,t=0}^{N-1}$. We can easily verify $\sum_{s,t}\hat{A}^\star(\omega_{s,t}) = \hat{1}_\mathrm{A}$ and thus $\hat{A}^\star$ is a POVM on $\mathcal{H}_\mathrm{A}$ with N^2 outcomes $\{\omega_{s,t}\}_{s,t}$. Since Eq. (25) with $\hat{A} = \hat{A}^\star$ holds, from Theorem III B, $\hat{A}^\star$ is an optimal solution to Eq. (6). Substituting Eqs. (90) and (93) into Eq. (1), the corresponding minimum-error sequential measurement $\hat{\Pi}^{(\hat{A}^\star)}$ can be expressed by

$$\begin{aligned}\hat{\Pi}_{m,n}^{(\hat{A}^\star)} &= \sum_{s=0}^{N-1}\sum_{t=0}^{N-1}\left(\frac{1}{N}|u_{s,t}\rangle\langle u_{s,t}|\right)\otimes(\delta_{n,t\oplus ms}|m\rangle\langle m|)\\ &= \frac{1}{N}\sum_{s=0}^{N-1}|u_{s,n\ominus ms}\rangle\langle u_{s,n\ominus ms}|\otimes|m\rangle\langle m|.\end{aligned} \quad (94)$$

VII. CONCLUSION

We have studied a sequential-measurement version of the generalized state discrimination problem discussed in Ref. [30]. Since the entire set of sequential measurements is convex, Problem P is convex programming. The corresponding dual problem and necessary and sufficient conditions for an optimal sequential measurement were derived. We also showed that for an optimization problem having a certain group symmetry, there exists an optimal solution with the

[3] If $N = 2$, then Eq. (76) does not hold. However, we can apply the same technique to this case and obtain $\hat{X}^\star = (\frac{1}{4} + \frac{1}{4\sqrt{2}})\hat{1}_\mathrm{A}$ [i.e., $P_\mathrm{S}(\hat{A}^\star) = \mathrm{Tr}\,\hat{X}^\star = \frac{1}{2} + \frac{1}{2\sqrt{2}}$], which is consistent with the result in Ref. [2].

same type of symmetry. Moreover, the minimax version of this problem was studied and necessary and sufficient conditions for a minimax solution were provided. We expect that our results will be useful for the investigation of a broad class of state discrimination problems with sequential measurements.

ACKNOWLEDGMENTS

We are grateful to O. Hirota of Tamagawa University for support. This work was supported by JSPS KAKENHI Grants No. JP17H07115 and No. JP16H04367.

APPENDIX A: ANOTHER METHOD OF DERIVATION OF PROBLEM DP

Let $\sigma(\Omega)$ be the sigma algebra of all measurable subsets of Ω. Alice's POVM $\hat{A} \in \mathcal{M}_{\rm A}$ is a mapping of $\sigma(\Omega)$ into $\mathcal{S}_{\rm A}^+$, which satisfies (i) positivity, $\hat{A}(E) \geqslant 0\ \forall E \in \sigma(\Omega)$; (ii) countable additivity, $\hat{A}(\cup_k E_k) = \sum_k \hat{A}(E_k)$ with mutually disjoint $\{E_k\} \subset \sigma(\Omega)$; and (iii) normalization, $\hat{A}(\Omega) = \hat{1}_{\rm A}$. Let $\overline{\mathcal{M}_{\rm A}}$ be the entire set of (not necessarily normalized) mappings $\hat{A}: \sigma(\Omega) \to \mathcal{S}_{\rm A}^+$ satisfying conditions (i) and (ii). Obviously, $\overline{\mathcal{M}_{\rm A}} \supset \mathcal{M}_{\rm A}$ holds.

We define the Lagrangian for Problem P as

$$L(\hat{A},\hat{X},\lambda) \equiv f(\hat{A}) + {\rm Tr}\{\hat{X}[\hat{1}_{\rm A} - \hat{A}(\Omega)]\} - \sum_{j=0}^{J-1} \lambda_j \eta_j(\hat{A}), \tag{A1}$$

where $L(\hat{A},\hat{X},\lambda)$ is a function of $\hat{A} \in \overline{\mathcal{M}_{\rm A}}$ and $(\hat{X},\lambda) \in \mathcal{X}$. If $\hat{A}(\Omega) \neq \hat{1}_{\rm A}$ holds, then there exists a vector $|x\rangle$ satisfying $|x\rangle \notin {\rm Ker}[\hat{1}_{\rm A} - \hat{A}(\Omega)]$; taking the limit $t \to \infty$ or $t \to -\infty$ yields $L(\hat{A}, t\,|x\rangle\langle x|\,,\lambda) \to -\infty$. Similarly, if there exists $j \in \mathcal{I}_J$ such that $\eta_j(\hat{A}) > 0$, then $L(\hat{A},\hat{X},\lambda) \to -\infty$ when $\lambda_j \to \infty$. Thus, if $\hat{A} \notin \mathcal{M}_{\rm A}^\circ$ holds, then there exists $(\hat{X},\lambda) \in \mathcal{X}$ such that $L(\hat{A},\hat{X},\lambda) \to -\infty$. On the other hand, if $\hat{A} \in \mathcal{M}_{\rm A}^\circ$ holds, then $L(\hat{A},\hat{X},\lambda) \geqslant f(\hat{A})$ holds and the equality holds when $\lambda = 0$. Therefore, we obtain

$$\max_{\hat{A}\in\overline{\mathcal{M}_{\rm A}}} \min_{(\hat{X},\lambda)\in\mathcal{X}} L(\hat{A},\hat{X},\lambda) = \max_{\hat{A}\in\mathcal{M}_{\rm A}^\circ} \min_{(\hat{X},\lambda)\in\mathcal{X}} L(\hat{A},\hat{X},\lambda) = \max_{\hat{A}\in\mathcal{M}_{\rm A}^\circ} f(\hat{A}) = f^\star. \tag{A2}$$

Let

$$\tilde{s}(\hat{X},\lambda) \equiv \max_{\hat{A}\in\overline{\mathcal{M}_{\rm A}}} L(\hat{A},\hat{X},\lambda). \tag{A3}$$

Substituting $F = f$, $x = \hat{A}$, and $y = (\hat{X},\lambda)$ into the formula

$$\min_y \max_x F(x,y) \geqslant \max_x \min_y F(x,y) \tag{A4}$$

and using Eqs. (A2) and (A3) yields

$$\min_{(\hat{X},\lambda)\in\mathcal{X}} \tilde{s}(\hat{X},\lambda) \geqslant f^\star. \tag{A5}$$

Let us consider the problem of finding $(\hat{X},\lambda) \in \mathcal{X}$ that minimizes $\tilde{s}(\hat{X},\lambda)$, which can be regarded as a dual problem of Problem P. From Eqs. (15)–(A1), $L(\hat{A},\hat{X},\lambda)$ is rewritten as

$$L(\hat{A},\hat{X},\lambda) = s(\hat{X},\lambda) + \int_\Omega {\rm Tr}[\hat{\sigma}_\omega(\lambda) - \hat{X}]\hat{A}(d\omega). \tag{A6}$$

If $(\hat{X},\lambda) \notin \mathcal{X}^\circ$ holds [i.e., there exists ω such that $\hat{X} \ngeqslant \hat{\sigma}_\omega(\lambda)$], then there exists a vector $|x\rangle \in \mathcal{H}_{\rm A}$ such that $\langle x|[\hat{X} - \hat{\sigma}_\omega(\lambda)]|x\rangle < 0$; substituting $\hat{A}(\omega) = t\,|x\rangle\langle x|$ into Eq. (A6) and taking the limit $t \to \infty$ gives $L(\hat{A},\hat{X},\lambda) = \infty$. Thus, from Eq. (A3), $\tilde{s}(\hat{X},\lambda) = \infty$ holds. On the other hand, if $(\hat{X},\lambda) \in \mathcal{X}^\circ$ holds, then $L(\hat{A},\hat{X},\lambda)$ reaches its maximum value of $s(\hat{X},\lambda)$ when $\hat{A}(E) = 0$ for any $E \subseteq \Omega$ and thus $\tilde{s}(\hat{X},\lambda) = s(\hat{X},\lambda)$ holds. Therefore, we obtain

$$\min_{(\hat{X},\lambda)\in\mathcal{X}} \tilde{s}(\hat{X},\lambda) = \min_{(\hat{X},\lambda)\in\mathcal{X}^\circ} s(\hat{X},\lambda), \tag{A7}$$

which indicates that the dual problem can be rewritten as Problem DP.

APPENDIX B: PROOF OF THEOREM 1

We will prove the cases of $f^\star > -\infty$ and $f^\star = -\infty$ separately.

1. Case of $f^\star > -\infty$

From $s^\star \geqslant f^\star$, it is sufficient to show that there exists $\hat{A} \in \mathcal{M}_{\rm A}^\circ$ satisfying $f(\hat{A}) \geqslant s^\star$. Indeed, in this case, $s^\star = f^\star$ holds from $s^\star \leqslant f(\hat{A}) \leqslant f^\star$.

Let us consider the set

$$\mathcal{Z} \equiv \{(\{\hat{\sigma}_\omega(\lambda) + \hat{x}_\omega - \hat{X}\}_{\omega\in\Omega}, s(\hat{X},\lambda) - u): (\hat{X},\lambda,u,\{\hat{x}_\omega\}_{\omega\in\Omega}) \in \mathcal{T}\}, \tag{B1}$$

where

$$\mathcal{T} \equiv \{(\hat{X},\lambda,u,\{\hat{x}_\omega\}): (\hat{X},\lambda) \in \mathcal{X}, s^\star > u \in \mathbf{R}, \hat{x}_\omega \in \mathcal{S}_{\rm A}^+\}. \tag{B2}$$

Since $\hat{x}_\omega$ is in $\mathcal{S}_{\rm A}^+$, $\hat{\sigma}_\omega(\lambda) + \hat{x}_\omega - \hat{X} = 0$ holds only if $\hat{X} \geqslant \hat{\sigma}_\omega(\lambda)$ holds, which implies that $\{\hat{\sigma}_\omega(\lambda) + \hat{x}_\omega - \hat{X}\}_\omega = \{0\}$ holds only if $(\hat{X},\lambda) \in \mathcal{X}^\circ$ holds. Since $s(\hat{X},\lambda) \geqslant s^\star > u$ holds when $(\hat{X},\lambda) \in \mathcal{X}^\circ$, we have $(\{0\},0) \notin \mathcal{Z}$. Also, we can easily see that $\mathcal{Z}$ is a convex set having a nonempty interior. Thus, from the geometric Hahn-Banach theorem (e.g., [58]), for any $(\hat{X},\lambda,u,\{\hat{x}_\omega\}) \in \mathcal{T}$, there exists $(\{\tilde{A}_\omega\}_{\omega\in\Omega},\alpha) \neq (\{0\},0)$ with $\tilde{A}_\omega \in \mathcal{S}_{\rm A}$ and $\alpha \in \mathbf{R}$ satisfying

$${\rm Tr}\int_\Omega \tilde{A}_\omega[\hat{\sigma}_\omega(\lambda) + \hat{x}_\omega - \hat{X}]\mu(d\omega) + \alpha[s(\hat{X},\lambda) - u] \geqslant 0, \tag{B3}$$

where μ is a strictly positive measure on a sigma algebra $\sigma(\Omega)$ satisfying $\mu(\Omega) = 1$. Let $\delta_\omega(E)$ $(E \subseteq \Omega)$ be the Dirac measure, which is defined by $\delta_\omega(E) = 1$ if $\omega \in E$ holds and $\delta_\omega(E) = 0$ otherwise. By substituting $\hat{x}_\omega = t\hat{x}\delta_{\omega'}(\omega)$ $(\hat{x} \geqslant 0$ and $\omega' \in \Omega)$ into Eq. (B3) and taking the limit $t \to \infty$, we obtain ${\rm Tr}(\tilde{A}_{\omega'}\hat{x}) \geqslant 0$. Since this inequality holds for any $\hat{x} \geqslant 0$ and $\omega' \in \Omega$, $\tilde{A}_\omega \geqslant 0$ holds for any $\omega \in \Omega$. Also, taking the limit $u \to -\infty$ in Eq. (B3) gives $\alpha \geqslant 0$.

To show $\alpha > 0$, assume by contradiction that $\alpha = 0$. Substituting $\hat{X} = t\hat{x}$ $(\hat{x} \geqslant 0)$ and $\hat{x}_\omega = t[1 - \delta_{\omega'}(\omega)]\hat{x}$ into Eq. (B3) and taking the limit $t \to \infty$ gives

$${\rm Tr}\int_\Omega \tilde{A}_\omega\delta_{\omega'}(\omega)\hat{x}\mu(d\omega) = {\rm Tr}(\tilde{A}_{\omega'}\hat{x}) \leqslant 0, \tag{B4}$$

which implies $\tilde{A}_\omega \leqslant 0$ for any $\omega \in \Omega$. Thus, $\tilde{A}_\omega = 0$ must hold, which contradicts $(\{\tilde{A}_\omega\},\alpha) \neq (\{0\},0)$. Therefore, $\alpha > 0$ holds.

Here let $\hat{A} \in \mathcal{M}_{\rm A}$ be a measure satisfying $\hat{A}(\omega) = \tilde{A}_\omega \mu(\omega)/\alpha$ for any $\omega \in \Omega$. To complete the proof, we will show $\hat{A} \in \mathcal{M}_{\rm A}^\circ$ and $f(\hat{A}) \geqslant s^\star$. Dividing both sides of Eq. (B3) by α yields

$$\mathrm{Tr} \int_\Omega [\hat{\sigma}_\omega(\lambda) + \hat{x}_\omega - \hat{X}]\hat{A}(d\omega) + s(\hat{X},\lambda) - u \geqslant 0. \tag{B5}$$

Substituting $\hat{X} = t\hat{x}$ ($\hat{x} \in \mathcal{S}_{\rm A}$) into Eq. (B5) and taking the limit $t \to \infty$ gives

$$\mathrm{Tr}\hat{x} \geqslant \mathrm{Tr}\left[\hat{x} \int_\Omega \hat{A}(d\omega)\right] = \mathrm{Tr}[\hat{x}\hat{A}(\Omega)]. \tag{B6}$$

Since this inequality holds for any $\hat{x} \in \mathcal{S}_{\rm A}$, $\hat{A}(\Omega) = \hat{1}_{\rm A}$ holds. Substituting $\lambda_j = t\delta_{j,j'}$ into Eq. (B5) and taking the limit $t \to \infty$ gives $\eta_{j'}(\hat{A}) \leqslant 0$ and thus $\hat{A} \in \mathcal{M}_{\rm A}^\circ$ holds. Also, substituting $\hat{x}_\omega = \hat{X} = 0$ and $\lambda = 0$ into Eq. (B5) and taking the limit $u \to s^\star$ gives $f(\hat{A}) \geqslant s^\star$. Therefore, $s^\star = f^\star$ holds.

2. Case of $f^\star = -\infty$

Let us consider the following set:

$$\mathcal{W} = \left\{\{\eta_j(\hat{A})\}_{j=0}^{J-1} \in \mathbf{R}^J : \hat{A} \in \mathcal{M}_{\rm A}\right\}. \tag{B7}$$

Since $f^\star = -\infty$ implies that $\mathcal{M}_{\rm A}^\circ$ is empty, for any $\hat{A} \in \mathcal{M}_{\rm A}$, there exists $j \in \mathcal{I}_J$ such that $\eta_j(\hat{A}) > 0$. Therefore, the set $\mathcal{W}' \equiv \{\{\beta_j \leqslant 0\}_{j=0}^{J-1} \in \mathbf{R}^J\}$ has no intersection with $\mathcal{W}$. We can easily verify that $\mathcal{W}$ is compact and $\mathcal{W}'$ is closed; thus, by a separating hyperplane theorem (e.g., [59]), there exist $q \equiv \{q_j\}_{j=0}^{J-1} \in \mathbf{R}_+^J$ and $0 < \epsilon \in \mathbf{R}_+$ such that

$$\sum_{j=0}^{J-1} q_j \eta_j(\hat{A}) > \epsilon \ \forall \hat{A} \in \mathcal{M}_{\rm A}. \tag{B8}$$

Now assume that $\sum_{j=0}^{J-1} q_j = 1$, with no loss of generality. Equations (3) and (B8) give

$$\sum_{j=0}^{J-1} q_j \mathrm{Tr} \sum_{m=0}^{M-1} \hat{a}_{j,m} \hat{\Pi}^{(\hat{A})} \geqslant \sum_{j=0}^{J-1} q_j b_j + \epsilon. \tag{B9}$$

Substituting Eq. (1) into this equation and doing some algebra gives

$$\mathrm{Tr} \int_\Omega \Xi(\omega)\hat{A}(d\omega) \geqslant 0, \tag{B10}$$

where

$$\Xi(\omega) \equiv \mathrm{Tr}_{\rm B} \sum_{j=0}^{J-1} q_j \sum_{m=0}^{M-1} \hat{a}_{j,m} \hat{B}_m^{(\omega)} - \left(\sum_{j=0}^{J-1} q_j b_j + \epsilon\right) \frac{\hat{1}_{\rm A}}{d_{\rm A}} \tag{B11}$$

and $d_{\rm A} = \dim \mathcal{H}_{\rm A}$. Since Eq. (B10) holds for any $\hat{A} \in \mathcal{M}_{\rm A}$, $\Xi(\omega) \geqslant 0$ holds.

Let $\hat{X}_0^\star$ be the optimal solution to the following problem:

$$\text{minimize} \ \ \mathrm{Tr}\hat{X}_0$$
$$\text{subject to} \ \ \hat{X}_0 \geqslant \mathrm{Tr}_{\rm B} \sum_{m=0}^{M-1} \hat{c}_m \hat{B}_m^{(\omega)} \ \forall \omega \in \Omega. \tag{B12}$$

Also, let

$$\hat{Y}(t,q) \equiv \hat{X}_0^\star - t\left(\sum_{j=0}^{J-1} q_j b_j + \epsilon\right)\frac{\hat{1}_{\rm A}}{d_{\rm A}}, \tag{B13}$$

where $t \in \mathbf{R}_+$. From Eqs. (B12) and (B13) we have

$$\begin{aligned}\hat{Y}(t,q) &\geqslant \mathrm{Tr}_{\rm B} \sum_{m=0}^{M-1} \hat{c}_m \hat{B}_m^{(\omega)} - t\,\mathrm{Tr}_{\rm B} \sum_{j=0}^{J-1} q_j \sum_{m=0}^{M-1} \hat{a}_{j,m} \hat{B}_m^{(\omega)} \\ &= \mathrm{Tr}_{\rm B} \sum_{m=0}^{M-1} \hat{z}_m(tq)\hat{B}_m^{(\omega)} = \hat{\sigma}_\omega(tq),\end{aligned} \tag{B14}$$

where the first and second lines follow from $\Xi(\omega) \geqslant 0$ and the definition of $\hat{z}_m(\lambda)$ given by Eq. (11), respectively. Thus, $[\hat{Y}(t,q),tq] \in \mathcal{X}^\circ$ holds, which gives $s[\hat{Y}(t,q),tq] \geqslant s^\star$. From Eq. (B13) we obtain

$$\begin{aligned} s^\star &\leqslant s[\hat{Y}(t,q),tq] = \mathrm{Tr}\hat{Y}(t,q) + t\sum_{j=0}^{J-1} q_j b_j \\ &= \mathrm{Tr}\hat{X}_0^\star - t\epsilon. \end{aligned} \tag{B15}$$

Since $\mathrm{Tr}\hat{X}_0^\star$ is constant, $\mathrm{Tr}\hat{X}_0^\star - t\epsilon \to -\infty$ as $t \to \infty$. Therefore, $s^\star = -\infty$ holds. ■

APPENDIX C: PROOF OF THEOREM 2

We will show (a) $\Rightarrow$ (b), (b) $\Rightarrow$ (c), and (c) $\Rightarrow$ (a) in this order. After that, we will show that $(\hat{X},\lambda)$ is an optimal solution to Problem DP if condition (b) holds.

First, we show (a) $\Rightarrow$ (b). Let $(\hat{X},\lambda)$ be an optimal solution to Problem DP. Since $\hat{A}(\Omega) = \hat{1}_{\rm A}$ and $\eta_j(\hat{A}) \leqslant 0$ hold, the second and third terms of the right-hand side of Eq. (A1) are zero and non-negative, respectively, which gives $L(\hat{A},\hat{X},\lambda) \geqslant f(\hat{A}) = f^\star$. Also, since $\hat{X} \geqslant \hat{\sigma}_\omega(\lambda)$ and $\hat{A}(\omega) \geqslant 0$ hold, the second term of the right-hand side of Eq. (A6) is nonpositive, which gives $L(\hat{A},\hat{X},\lambda) \leqslant s(\hat{X},\lambda) = s^\star$ holds. Since $f^\star = s^\star$ holds from Theorem 1, we obtain

$$f^\star = L(\hat{A},\hat{X},\lambda) = s^\star, \tag{C1}$$

i.e., the third term of the right-hand side of Eq. (A1) and the second term of the right-hand side of Eq. (A6) must be zero. Thus, Eqs. (17) and (18) hold. Note that Eq. (17) follows from the fact that $\hat{x}\hat{y} = 0$ holds for any $\hat{x},\hat{y} \in \mathcal{S}_{\rm A}^+$ satisfying $\mathrm{Tr}(\hat{x}\hat{y}) = 0$.

Next we show (b) $\Rightarrow$ (c). Integrating both sides of Eq. (17) and using $\hat{A}(\Omega) = \hat{1}_{\rm A}$ gives

$$\hat{X} = \int_\Omega \hat{\sigma}_\omega(\lambda)\hat{A}(d\omega). \tag{C2}$$

Further, $\hat{X} \geqslant \hat{\sigma}_\omega(\lambda)$ gives Eq. (19). Equation (20) is equivalent to Eq. (18).

We show (c) ⇒ (a). We define $\hat{X}$ as in Eq. (C2). We have that for any POVM $\hat{A}' \in \mathcal{M}_{\rm A}^\circ$,

$$\begin{aligned} f(\hat{A}) - f(\hat{A}') &\geqslant f(\hat{A}) - \sum_{j=0}^{J-1} \lambda_j \eta_j(\hat{A}) - f(\hat{A}') + \sum_{j=0}^{J-1} \lambda_j \eta_j(\hat{A}') \\ &= \sum_{m=0}^{M-1} \mathrm{Tr}\big[\hat{z}_m(\lambda)\hat{\Pi}_m^{(\hat{A})} - \hat{z}_m(\lambda)\hat{\Pi}_m^{(\hat{A}')}\big] \\ &= \mathrm{Tr}\hat{X} - \mathrm{Tr}\int_\Omega \hat{\sigma}_\omega(\lambda)\hat{A}'(d\omega) \\ &= \mathrm{Tr}\int_\Omega [\hat{X} - \hat{\sigma}_\omega(\lambda)]\hat{A}'(d\omega) \geqslant 0. \end{aligned} \quad \text{(C3)}$$

The second line follows from Eq. (20) and $\eta_j(\hat{A}') \leqslant 0$. The third line follows from Eqs. (2), (3), and (11). The fourth line and the last inequality respectively follow from Eqs. (12) and (19) [i.e., $\hat{X} \geqslant \hat{\sigma}_\omega(\lambda)$]. From Eq. (C3), $\hat{A}$ is an optimal solution to Problem P.

Finally, we will show that $(\hat{X},\lambda)$ is an optimal solution to Problem DP if condition (b) holds. From Eqs. (A6) and (17), $L(\hat{A},\hat{X},\lambda) = s(\hat{X},\lambda)$ holds. Also, from Eqs. (A1) and (18), $L(\hat{A},\hat{X},\lambda) = f(\hat{A}) = f^\star$ holds. Thus, $s(\hat{X},\lambda) = f^\star$ holds, which means that $(\hat{X},\lambda)$ is an optimal solution to Problem DP. ■

APPENDIX D: PROOF OF THEOREM 3

1. Outline

Let $b \equiv \{b_j\}_{j=0}^{J-1} \in \mathbf{R}^J$. Also, let $f^\bullet(\beta)$ be the optimal value of the optimization problem obtained by replacing b of Problem P with $\beta \equiv \{\beta_j\}_{j=0}^{J-1} \in \mathbf{R}^J$. We will first show that $f^\bullet(\beta)$ is a concave function. We will also show that there exists an optimal solution to Problem P with at most $d_{\rm A}^2$ outcomes if $f^\bullet(\beta)$ is strictly concave at $\beta = b$ and with at most $(J+1)d_{\rm A}^2$ outcomes otherwise.

2. Preparations

Before proceeding to the proof, we make some preparations. From Theorem 1 of Ref. [60], any $\hat{A} \in \mathcal{M}_{\rm A}$ can be expressed as

$$\hat{A}(\omega) = \int_\Gamma \hat{E}^{(\gamma)}(\omega)p(d\gamma), \quad \text{(D1)}$$

where $\hat{E}^{(\gamma)} \in \mathcal{M}_{\rm A}$ is a POVM with at most $d_{\rm A}^2$ outcomes, Γ is the entire set of indices γ such that $\hat{E}^{(\gamma)}$ is a POVM with at most $d_{\rm A}^2$ outcomes, and p is a probability measure, which satisfies $p(\Gamma) = 1$. From Eqs. (1), (2), and (D1), we have

$$\begin{aligned} f(\hat{A}) &= \sum_{m=0}^{M-1} \mathrm{Tr}\left[\hat{c}_m \int_\Omega \left(\int_\Gamma \hat{E}^{(\gamma)}(d\omega)p(d\gamma)\right) \otimes \hat{B}_m^{(\omega)}\right] \\ &= \int_\Gamma f[\hat{E}^{(\gamma)}]p(d\gamma). \end{aligned} \quad \text{(D2)}$$

Let us define Γ° as

$$\Gamma^\circ \equiv \big\{\gamma \in \Gamma : \hat{E}^{(\gamma)} \in \mathcal{M}_{\rm A}^\circ\big\}, \quad \text{(D3)}$$

which is the entire set of indices γ such that $\hat{E}^{(\gamma)}$ is a feasible solution to Problem P. Let $\hat{A}$ be an optimal solution to Problem P.

We show the following lemma.

Lemma 1. If $p(\Gamma^\circ) = 1$ holds, then there exists an optimal solution to Problem P with at most $d_{\rm A}^2$ outcomes.

Proof. Let $\gamma^\star$ be an index satisfying

$$\gamma^\star \in \arg\max_{\gamma\in\Gamma^\circ} f[\hat{E}^{(\gamma)}]. \quad \text{(D4)}$$

From Eq. (D2), we have

$$\begin{aligned} f^\star = f(\hat{A}) &= \int_\Gamma f[\hat{E}^{(\gamma)}]p(d\gamma) \\ &= \int_{\Gamma^\circ} f[\hat{E}^{(\gamma)}]p(d\gamma) \leqslant f[\hat{E}^{(\gamma^\star)}]. \end{aligned} \quad \text{(D5)}$$

On the other hand, from $\gamma^\star \in \Gamma^\circ$ (i.e., $\hat{E}^{(\gamma^\star)} \in \mathcal{M}_{\rm A}^\circ$), $f[\hat{E}^{(\gamma^\star)}] \leqslant f^\star$ must hold. Thus, $f[\hat{E}^{(\gamma^\star)}] = f^\star$. Therefore, $\hat{E}^{(\gamma^\star)}$, which is a POVM with at most $d_{\rm A}^2$ outcomes, is an optimal solution to Problem P. ■

3. Proof

We first consider the case $J = 0$. From $\Gamma^\circ = \Gamma$, $p(\Gamma^\circ) = 1$ holds. Thus, from Lemma 1 there exists an optimal solution to Problem P with at most $d_{\rm A}^2$ outcomes. For the remainder of the proof, the case $J \geqslant 1$ is considered.

In the following, we will show that $f^\bullet(\beta)$ is a concave function. It suffices to consider the range of β such that $f^\bullet(\beta) > -\infty$. Let $\mathcal{M}_{\rm A}^\bullet(\beta) \subseteq \mathcal{M}_{\rm A}$ be the feasible set of the optimization problem obtained by replacing b of Problem P with β. Now we consider distinct $\beta^{(1)},\beta^{(2)} \in \mathbf{R}^J$. For each $k \in \{1,2\}$, there exists $\hat{A}_k \in \mathcal{M}_{\rm A}^\bullet[\beta^{(k)}]$ satisfying $f(\hat{A}_k) = f^\bullet[\beta^{(k)}]$. Since $t\hat{A}_1 + (1-t)\hat{A}_2 \in \mathcal{M}_{\rm A}^\bullet[t\beta^{(1)} + (1-t)\beta^{(2)}]$ with $0 \leqslant t \leqslant 1$ holds, we obtain

$$\begin{aligned} f^\bullet[t\beta^{(1)} + (1-t)\beta^{(2)}] &\geqslant f[t\hat{A}_1 + (1-t)\hat{A}_2] \\ &= tf(\hat{A}_1) + (1-t)f(\hat{A}_2) \\ &= tf^\bullet[\beta^{(1)}] + (1-t)f^\bullet[\beta^{(2)}]. \end{aligned} \quad \text{(D6)}$$

Therefore, $f^\bullet(\beta)$ is concave.

Let us consider a linear function $f_{\rm L}(\beta)$ such that

$$f_{\rm L}(\beta) - f^\bullet(\beta) \geqslant f_{\rm L}(b) - f^\bullet(b) = 0. \quad \text{(D7)}$$

Since $-f^\bullet(\beta)$ is convex and thus subdifferentiable at each point [61], there always exists such $f_{\rm L}(\beta)$. Let

$$\mathcal{D} \equiv \{\beta \in \mathbf{R}^J : f_{\rm L}(\beta) = f^\bullet(\beta)\}. \quad \text{(D8)}$$

It follows that $\mathcal{D}$ is a convex set including b. Let $\mathcal{E}_\mathcal{D}$ be the entire set of extremal points of $\mathcal{D}$. Also, let $\mathcal{E}$ be the entire set of $\beta' \in \mathbf{R}^J$ such that $f^\bullet(\beta)$ is strictly concave at $\beta = \beta'$. We can easily verify $\mathcal{E}_\mathcal{D} \subseteq \mathcal{E}$.

First, we consider the case $b \in \mathcal{E}_\mathcal{D}$. From $b \in \mathcal{E}$, $f^\bullet(\beta)$ is strictly concave at $\beta = b$. From Lemma 1 it suffices to show $p(\Gamma^\circ) = 1$; assume by contradiction that $p(\Gamma^\circ) < 1$. Let, for each $j \in \mathcal{I}_J$,

$$\Gamma_j \equiv \{\gamma \in \Gamma : \eta_k[\hat{E}^{(\gamma)}] \leqslant 0 (\forall k \in \mathcal{I}_j), \eta_j[\hat{E}^{(\gamma)}] > 0\}. \quad \text{(D9)}$$

For simplicity, let $\Gamma_J \equiv \Gamma^\circ$. The $\{\Gamma_j\}_{j=0}^J$ are obviously disjoint sets satisfying $\bigcup_{j=0}^J \Gamma_j = \Gamma$. Let $p_j \equiv p(\Gamma_j)$ and, for each

$j \in \mathcal{I}_{J+1}$,

$$\hat{A}_j(\omega) \equiv \begin{cases} \int_{\Gamma_j} \hat{E}^{(\gamma)}(\omega) \frac{p(d\gamma)}{p_j} & \text{for } p_j > 0 \\ 0, & \text{otherwise;} \end{cases} \quad \text{(D10)}$$

then $\hat{A}_j$ is in $\mathcal{M}_{\rm A}$ if $p_j > 0$ holds. From Eqs. (D1) and (D10) we have

$$\hat{A}(\omega) = \sum_{j=0}^{J} \int_{\Gamma_j} \hat{E}^{(\gamma)}(\omega) p(d\gamma) = \sum_{j=0}^{J} p_j \hat{A}_j(\omega). \quad \text{(D11)}$$

Thus, we obtain

$$f^\bullet(b) = f(\hat{A}) = f\left(\sum_{j=0}^{J} p_j \hat{A}_j\right) = \sum_{j=0}^{J} p_j f(\hat{A}_j) \leqslant \sum_{j=0}^{J} p_j f^\bullet[\beta^{(j)}], \quad \text{(D12)}$$

where $\beta^{(k)} \equiv \{\eta_j(\hat{A}_k)\}_{j=0}^{J-1}$. The inequality follows from the fact that $\hat{A}_j \in \mathcal{M}_{\rm A}^\bullet[\beta^{(j)}]$ (i.e., $f(\hat{A}_j) \leqslant f^\bullet[\beta^{(j)}]$) holds when $p_j > 0$. On the other hand, it follows that $p_j < 1$ holds for any $j \in \mathcal{I}_{J+1}$. Indeed, $p_J = p(\Gamma^\circ) < 1$ obviously holds. Also, since $\eta_j[\hat{E}^{(\gamma)}] > 0$ holds for any $j \in \mathcal{I}_J$ and $\gamma \in \Gamma_j$, if $p_j = 1$ holds for some $j \in \mathcal{I}_J$, then $\eta_j(\hat{A}) = \eta_j(\hat{A}_j) > 0$ holds from Eq. (D10), which contradicts $\hat{A} \in \mathcal{M}_{\rm A}^\circ$. Thus, there exist at least two distinct integers $k \in \mathcal{I}_{J+1}$ satisfying $p_k > 0$. This implies that, from Eq. (D12), $f^\bullet(\beta)$ is not strictly concave at $\beta = b$ (i.e., $b \notin \mathcal{E}$), which contradicts $b \in \mathcal{E}_\mathcal{D} \subseteq \mathcal{E}$. Therefore, $p(\Gamma^\circ) = 1$ must hold. From Lemma 1 there exists an optimal solution to Problem P with at most $d_{\rm A}^2$ outcomes.

Next we consider the case $b \notin \mathcal{E}_\mathcal{D}$. Since $\mathcal{D}$ is convex, from the finite-dimensional version of the Krein-Milman theorem [62], $\mathcal{D}$ is the convex hull of $\mathcal{E}_\mathcal{D}$. Thus, from Carathéodory's theorem, there exists a set of $J+1$ points $\{b^{(j)}\}_{j=0}^{J} \subseteq \mathcal{E}_\mathcal{D}$ such that $b \in \mathcal{D}$ lies in the convex hull of $\{b^{(j)}\}$ [note that $b^{(j)}$ and $b^{(j')}$ $(j \neq j')$ can be the same]. This indicates that there exists $\{q_j\}_{j=0}^{J} \in \mathbf{R}_+^{J+1}$ with $\sum_{j=0}^{J} q_j = 1$ such that $b = \sum_{j=0}^{J} q_j b^{(j)}$. From $b^{(j)} \in \mathcal{E}_\mathcal{D}$, similar to the above discussion, it follows that, for each $j \in \mathcal{I}_{J+1}$, there exists $\gamma_j \in \Gamma$ satisfying $f[\hat{E}^{(\gamma_j)}] = f^\bullet[b^{(j)}]$ and $\hat{E}^{(\gamma_j)} \in \mathcal{M}_{\rm A}^\bullet[b^{(j)}]$. Using such γ_j, let

$$\hat{A}' \equiv \sum_{j=0}^{J} q_j \hat{E}^{(\gamma_j)}; \quad \text{(D13)}$$

then we have

$$f(\hat{A}') = f\left[\sum_{j=0}^{J} q_j \hat{E}^{(\gamma_j)}\right] = \sum_{j=0}^{J} q_j f[\hat{E}^{(\gamma_j)}] = \sum_{j=0}^{J} q_j f^\bullet[b^{(j)}] = f^\bullet(b) = f^\star. \quad \text{(D14)}$$

The fourth equality follows from the fact that, from $b, b^{(j)} \in \mathcal{D}$ and Eq. (D8), $f^\bullet(b) = f_{\rm L}(b)$ and $f^\bullet[b^{(j)}] = f_{\rm L}[b^{(j)}]$ hold and $f_{\rm L}(\beta)$ is linear. Also, from $\hat{E}^{(\gamma_j)} \in \mathcal{M}_{\rm A}^\bullet[b^{(j)}]$, $\hat{A}' \in \mathcal{M}_{\rm A}^\bullet(b) = \mathcal{M}_{\rm A}^\circ$ holds. Thus, $\hat{A}'$, which is a POVM with at most $(J+1)d_{\rm A}^2$ outcomes, is an optimal solution to Problem P. ■

APPENDIX E: PROOF OF THEOREM 4

Using Eq. (34), we can easily verify that the following equations hold for any $g \in \mathcal{G}$, $c \in \mathbf{R}$, and $\hat{Q}, \hat{R} \in \mathcal{S}$:

$$\begin{aligned} g \circ (\hat{Q} + \hat{R}) &= g \circ \hat{Q} + g \circ \hat{R}, \\ g \circ (\hat{Q}\hat{R}) &= (g \circ \hat{Q})(g \circ \hat{R}), \\ g \circ (c\hat{Q}) &= c(g \circ \hat{Q}), \\ g \circ \hat{1} &= \hat{1}, \\ \mathrm{Tr}(g \circ \hat{Q}) &= \mathrm{Tr}\,\hat{Q}, \\ g \circ \hat{Q} &\geqslant 0 \ \forall \hat{Q} \geqslant 0, \\ g \circ \mathrm{Tr}_{\rm B}\,\hat{Q} &= \mathrm{Tr}_{\rm B}(g \circ \hat{Q}). \end{aligned} \quad \text{(E1)}$$

The similar equations (except for the last one) for $\mathcal{S}_{\rm A}$ and $\mathcal{S}_{\rm B}$ instead of $\mathcal{S}$ also hold. Also, from Eqs. (35) and (33), we have that for any $\hat{Q}^{({\rm A})} \in \mathcal{S}_{\rm A}$ and $\hat{Q}^{({\rm B})} \in \mathcal{S}_{\rm B}$,

$$g \circ [\hat{Q}^{({\rm A})} \otimes \hat{Q}^{({\rm B})}] = [g \circ \hat{Q}^{({\rm A})}] \otimes [g \circ \hat{Q}^{({\rm B})}]. \quad \text{(E2)}$$

In what follows, we will often make use of these equations without mentioning it.

Let, for any $g \in \mathcal{G}$ and $\hat{\Phi} \in \mathcal{M}_{\rm A}^\circ$,

$$\hat{\Phi}^{(g)}(\omega) \equiv \overline{g} \circ \hat{\Phi}(g \circ \omega). \quad \text{(E3)}$$

From Eq. (36), we obtain

$$\begin{aligned} g \circ \hat{\Pi}_m^{(\hat{\Phi}^{(g)})} &= \int_\Omega [g \circ \hat{\Phi}^{(g)}(d\omega)] \otimes \left[g \circ \hat{B}_m^{(\omega)}\right] \\ &= \int_\Omega \hat{\Phi}[d(g \circ \omega)] \otimes \hat{B}_{g \circ m}^{(g \circ \omega)} \\ &= \hat{\Pi}_{g \circ m}^{(\hat{\Phi})}. \end{aligned} \quad \text{(E4)}$$

We first show that the mapping $\kappa_g : \hat{\Phi} \mapsto \hat{\Phi}^{(g)}$ is bijective on $\mathcal{M}_{\rm A}^\circ$. We can easily verify that $\hat{\Phi}^{(g)}$ is a POVM. We have that for any $j \in \mathcal{I}_J$,

$$\begin{aligned} \sum_{m=0}^{M-1} \mathrm{Tr}\left[\hat{a}_{j,m} \hat{\Pi}_m^{(\hat{\Phi}^{(g)})}\right] &= \sum_{m=0}^{M-1} \mathrm{Tr}\left[(g \circ \hat{a}_{j,m})\left(g \circ \hat{\Pi}_m^{(\hat{\Phi}^{(g)})}\right)\right] \\ &= \sum_{m=0}^{M-1} \mathrm{Tr}\left[\hat{a}_{g \circ j, g \circ m} \hat{\Pi}_{g \circ m}^{(\hat{\Phi})}\right] \\ &= \sum_{m'=0}^{M-1} \mathrm{Tr}\left[\hat{a}_{g \circ j, m'} \hat{\Pi}_{m'}^{(\hat{\Phi})}\right] \\ &\leqslant b_{g \circ j} = b_j, \end{aligned} \quad \text{(E5)}$$

where $m' = g \circ m$. The second and fourth lines follow from Eq. (E4) and $\hat{\Phi} \in \mathcal{M}_{\rm A}^\circ$, respectively. Thus, $\hat{\Phi}^{(g)}$ is in $\mathcal{M}_{\rm A}^\circ$. Also, $\kappa_{\overline{g}}$ is the inverse mapping of κ_g. Therefore, κ_g is bijective on $\mathcal{M}_{\rm A}^\circ$.

We next define $\hat{A} \in \mathcal{M}_{\rm A}^\circ$ as

$$\hat{A}(\omega) \equiv \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \hat{\Phi}^{(g)}(\omega) \quad \text{(E6)}$$

and show Eq. (38), $\hat{A} \in \mathcal{M}_{\rm A}^{\circ}$, and $f(\hat{A}) = f(\hat{\Phi})$. We have that for any $g \in \mathcal{G}$ and $m \in \mathcal{I}_M$,

$$\begin{aligned} g \circ \hat{A}(\omega) &= \frac{1}{|\mathcal{G}|} \sum_{h \in \mathcal{G}} g \circ \hat{\Phi}^{(h)}(\omega) \\ &= \frac{1}{|\mathcal{G}|} \sum_{h' \in \mathcal{G}} \overline{h'} \circ \hat{\Phi}(h' \circ g \circ \omega) \\ &= \frac{1}{|\mathcal{G}|} \sum_{h' \in \mathcal{G}} \hat{\Phi}^{(h')}(g \circ \omega) = \hat{A}(g \circ \omega), \end{aligned} \tag{E7}$$

where $h' = h \circ \overline{g}$. This gives Eq. (38). From Eq. (E5) we have that for any $j \in \mathcal{I}_J$,

$$\sum_{m=0}^{M-1} \mathrm{Tr}\big[\hat{a}_{j,m} \hat{\Pi}_m^{(\hat{A})}\big] = \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \sum_{m=0}^{M-1} \mathrm{Tr}\big[\hat{a}_{j,m} \hat{\Pi}_m^{(\hat{\Phi}^{(g)})}\big] \leqslant b_j. \tag{E8}$$

Thus, $\hat{A} \in \mathcal{M}_{\rm A}^{\circ}$ holds. Moreover, we obtain

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

where the fourth line follows from Eq. (E4). In particular, if $\hat{\Phi}$ is an optimal solution to Problem P, then so is $\hat{A}$.

We finally show that there exists $(\hat{X},\lambda) \in \mathcal{X}^{\circ}$ satisfying Eq. (39). Let

$$\begin{aligned} \hat{Y}^{(g)} &\equiv g \circ \hat{Y}, \\ \nu^{(g)} &\equiv \big\{\nu_j^{(g)} \equiv \nu_{\overline{g} \circ j}\big\}_{j=0}^{J-1}; \end{aligned} \tag{E9}$$

then we have that for any $g \in \mathcal{G}$ and $m \in \mathcal{I}_M$,

$$\begin{aligned} g \circ \hat{z}_m(\nu) &= \hat{c}_{g \circ m} - \sum_{j=0}^{J-1} \nu_j \hat{a}_{g \circ j, g \circ m} \\ &= \hat{c}_{g \circ m} - \sum_{j=0}^{J-1} \nu_{g \circ j}^{(g)} \hat{a}_{g \circ j, g \circ m} \\ &= \hat{z}_{g \circ m}[\nu^{(g)}]. \end{aligned} \tag{E10}$$

Thus, we have that for any $\omega \in \Omega$,

$$\begin{aligned} \hat{Y}^{(g)} &\geqslant g \circ \hat{\sigma}_\omega(\nu) \\ &= \mathrm{Tr}_{\rm B} \sum_{m=0}^{M-1} [g \circ \hat{z}_m(\nu)]\big[g \circ \hat{B}_m^{(\omega)}\big] \\ &= \mathrm{Tr}_{\rm B} \sum_{m=0}^{M-1} \hat{z}_{g \circ m}[\nu^{(g)}] \hat{B}_{g \circ m}^{(g \circ \omega)} \\ &= \hat{\sigma}_{g \circ \omega}[\nu^{(g)}], \end{aligned} \tag{E11}$$

i.e., $[\hat{Y}^{(g)},\nu^{(g)}] \in \mathcal{X}^{\circ}$. Also, we obtain

$$\begin{aligned} s[\hat{Y}^{(g)},\nu^{(g)}] &= \mathrm{Tr} \hat{Y}^{(g)} + \sum_{j=0}^{J-1} \nu_j^{(g)} b_j \\ &= \mathrm{Tr} \hat{Y} + \sum_{j=0}^{J-1} \nu_{\overline{g} \circ j} b_{\overline{g} \circ j} \\ &= s(\hat{Y},\nu). \end{aligned} \tag{E12}$$

Let us define $(\hat{X},\lambda)$ as

$$\hat{X} \equiv \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \hat{Y}^{(g)}, \quad \lambda_j \equiv \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \nu_j^{(g)}. \tag{E13}$$

We can easily verify that Eq. (39) holds. We have that for any $\omega \in \Omega$,

$$\begin{aligned} \hat{\sigma}_\omega(\lambda) &= \frac{1}{|\mathcal{G}|} \mathrm{Tr}_{\rm B} \sum_{g \in \mathcal{G}} \sum_{m=0}^{M-1} \hat{z}_m[\nu^{(g)}] \hat{B}_m^{(\omega)} \\ &= \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \hat{\sigma}_\omega[\nu^{(g)}], \end{aligned} \tag{E14}$$

which gives

$$\hat{X} - \hat{\sigma}_\omega(\lambda) = \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} [\hat{Y}^{(g)} - \hat{\sigma}_\omega[\nu^{(g)}]] \geqslant 0, \tag{E15}$$

i.e., $(\hat{X},\lambda) \in \mathcal{X}^{\circ}$. Moreover, from Eqs. (E12) and (E13) we obtain

$$\begin{aligned} s(\hat{X},\lambda) &= \mathrm{Tr} \hat{X} + \sum_{j=0}^{J-1} \lambda_j b_j \\ &= \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \left[\mathrm{Tr} \hat{Y}^{(b)} + \sum_{j=0}^{J-1} \nu_j^{(g)} b_j \right] \\ &= \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} s[\hat{Y}^{(g)},\nu^{(g)}] = s(\hat{Y},\nu). \end{aligned} \tag{E16}$$

In particular, if $(\hat{Y},\nu)$ is an optimal solution to Problem DP, then so is $(\hat{X},\lambda)$. ■

APPENDIX F: PROOF OF THEOREM 9

Let $(\zeta,\hat{\Phi})$ be a minimax solution to Problem $\mathrm{P_m}$. Also, let $\mu^\star \equiv \{\mu_k^\star\}_{k=0}^{K-1}$ with

$$\mu_k^\star \equiv \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \zeta_{g \circ k}. \tag{F1}$$

We can see that $\mu^\star \in \mathcal{P}$ and the first line of Eq. (63) hold. Moreover, similar to Eq. (E6), let $\hat{A}^\star(\omega) \equiv |\mathcal{G}|^{-1} \sum_{g \in \mathcal{G}} \hat{\Phi}^{(g)}(\omega)$, where $\hat{\Phi}^{(g)}$ is defined by Eq. (E3); then, from Eq. (E7), the second line of Eq. (63) holds. The only thing we have to show now is that $(\mu^\star,\hat{A}^\star)$ is also a minimax solution. From Theorem

5 it suffices to show $f_k(\hat{A}^\star) \geqslant F^\star(\mu^\star)$ for any $k \in \mathcal{I}_K$. In what follows, we will show $f_k(\hat{A}^\star) \geqslant F^\star(\zeta)$ and $F^\star(\zeta) \geqslant F^\star(\mu^\star)$.

First, we show $f_k(\hat{A}^\star) \geqslant F^\star(\zeta)$ for any $k \in \mathcal{I}_K$. We have that for any $k \in \mathcal{I}_K$,

$$\begin{aligned} f_k(\hat{A}^\star) &= \frac{1}{|\mathcal{G}|}\sum_{m=0}^{M-1}\sum_{g\in\mathcal{G}}\mathrm{Tr}\big[\hat{c}_{k,m}\hat{\Pi}_m^{(\hat{\Phi}^{(g)})}\big] + d_k \\ &= \frac{1}{|\mathcal{G}|}\sum_{m=0}^{M-1}\sum_{g\in\mathcal{G}}\mathrm{Tr}\big[(g\circ\hat{c}_{k,m})\hat{\Pi}_{g\circ m}^{(\hat{\Phi})}\big] + d_k \\ &= \frac{1}{|\mathcal{G}|}\sum_{m'=0}^{M-1}\sum_{g\in\mathcal{G}}\mathrm{Tr}\big[\hat{c}_{g\circ k,m'}\hat{\Pi}_{m'}^{(\hat{\Phi})}\big] + d_k \\ &= \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}}\left[\sum_{m'=0}^{M-1}\mathrm{Tr}\big[\hat{c}_{g\circ k,m'}\hat{\Pi}_{m'}^{(\hat{\Phi})}\big] + d_{g\circ k}\right] \\ &= \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}} f_{g\circ k}(\hat{\Phi}) \geqslant F^\star(\zeta), \end{aligned} \tag{F2}$$

where $m' = g\circ m$. The second line follows from Eq. (E4). The inequality follows from the fact that, from Theorem V C, $f_k(\hat{\Phi}) \geqslant F^\star(\zeta)$ holds for any $k \in \mathcal{I}_K$.

Next we show $F^\star(\zeta) \geqslant F^\star(\mu^\star)$. Let $\zeta^{(g)} \equiv \{\zeta_k^{(g)} \equiv \zeta_{g\circ k}\}_{k=0}^{K-1}$; then we have that for any $g \in \mathcal{G}$,

$$\begin{aligned} F^\star[\zeta^{(g)}] &= \max_{\Phi\in\mathcal{M}_{\mathrm{A}}^\circ}\sum_{k=0}^{K-1}\zeta_{g\circ k}\left[\sum_{m=0}^{M-1}\mathrm{Tr}\big[\hat{c}_{k,m}\hat{\Pi}_m^{(\hat{\Phi})}\big] + d_k\right] \\ &= \max_{\Phi\in\mathcal{M}_{\mathrm{A}}^\circ}\sum_{k'=0}^{K-1}\zeta_{k'}\left[\sum_{m'=0}^{M-1}\mathrm{Tr}\big[\hat{c}_{k',m'}\hat{\Pi}_{m'}^{(\hat{\Phi}^{(g)})}\big] + d_{k'}\right] \\ &= F^\star(\zeta), \end{aligned} \tag{F3}$$

where $k' = g\circ k$ and $m' = g\circ m$. From Eq. (F3) we obtain

$$\begin{aligned} F^\star(\mu^\star) &= \max_{\Phi\in\mathcal{M}_{\mathrm{A}}^\circ}\frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}}\sum_{k=0}^{K-1}\zeta_k^{(g)}f_k(\Phi) \\ &\leqslant \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}}\max_{\Phi\in\mathcal{M}_{\mathrm{A}}^\circ}\sum_{k=0}^{K-1}\zeta_k^{(g)}f_k(\Phi) \\ &= \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}}F^\star[\zeta^{(g)}] = F^\star(\zeta); \end{aligned} \tag{F4}$$

thus, $(\mu^\star,\hat{A}^\star)$ is a minimax solution. ■

APPENDIX G: DERIVATION OF $(\hat{X}^\star,\lambda^\star)$

We will obtain an optimal solution $(\hat{X}^\star,\lambda^\star)$ to the problem of Eq. (22). This can be derived by extending methods described in Refs. [36,53], in which a minimum-error sequential measurement for double trine states is obtained.

Now we consider the problem of Eq. (22) in which λ is fixed. An optimal solution, denoted by $\hat{X}^\star(\lambda)$, to this problem can be expressed by $\hat{X}^\star(\lambda) = \upsilon(\lambda)\hat{1}_{\mathrm{A}}$, where $\upsilon(\lambda)$ is a real-valued function of λ. It follows that $\upsilon(\lambda)$ is the minimum value satisfying $\upsilon(\lambda)\hat{1}_{\mathrm{A}} \geqslant \hat{\sigma}_\omega(\lambda)$ for any $\omega \in \Omega$, which means that $\upsilon(\lambda)$ is the maximum value of the largest eigenvalues of $\hat{\sigma}_\omega(\lambda)$ over all $\omega \in \Omega$.

Substituting Eq. (65) into Eq. (23) gives

$$\hat{\sigma}_\omega(\lambda) = \sum_{m=0}^{2} l_m^{(\omega)}\,|\phi_m\rangle\langle\phi_m|\,, \tag{G1}$$

where

$$l_m^{(\omega)} \equiv \tfrac{1}{3}\,\langle\phi_m|\big[\hat{B}_m^{(\omega)} + \lambda\hat{B}_3^{(\omega)}\big]|\phi_m\rangle\,. \tag{G2}$$

Let $\upsilon_\omega^+(\lambda)$ and $\upsilon_\omega^-(\lambda)$ be the eigenvalues of $\hat{\sigma}_\omega(\lambda)$ with $\upsilon_\omega^+(\lambda) \geqslant \upsilon_\omega^-(\lambda)$. Here $\hat{U}_\theta$ is defined as

$$\hat{U}_\theta \equiv (\cos\theta)\hat{1} + \sin\theta(|1\rangle\langle 0| - |0\rangle\langle 1|), \tag{G3}$$

which is a unitary operator corresponding to a rotation of θ. There exists θ such that

$$\hat{\sigma}_\omega(\lambda) = \hat{U}_{\theta/2}[\upsilon_\omega^-(\lambda)\,|0\rangle\langle 0| + \upsilon_\omega^+(\lambda)\,|1\rangle\langle 1|]\hat{U}_{\theta/2}^\dagger. \tag{G4}$$

Using Eqs. (G1), (G3), and (G4) and doing some algebra gives

$$\upsilon_\omega^+(\lambda) = \sum_{m=0}^{2}\frac{1}{2}\left[1 - \cos\left(\theta - \frac{2\pi m}{3}\right)\right]l_m^{(\omega)}. \tag{G5}$$

Substituting Eq. (G2) into Eq. (G5) yields

$$\upsilon_\omega^+(\lambda) = \frac{\lambda+1}{2}\sum_{m=0}^{3}\mathrm{Tr}\big[\hat{\rho}_m^{(\theta)}\hat{B}_m^{(\omega)}\big], \tag{G6}$$

where

$$\hat{\rho}_m^{(\theta)} \equiv \begin{cases}\frac{1-\cos(\theta-\frac{2\pi m}{3})}{3(\lambda+1)}\,|\phi_m\rangle\langle\phi_m|\,, & m < 3\\ \lambda\sum_{r=0}^{2}\hat{\rho}_r^{(\theta)}, & m = 3.\end{cases} \tag{G7}$$

We can easily see $\sum_{m=0}^{3}\mathrm{Tr}\,\hat{\rho}_m^{(\theta)} = 1$.

The $\mathrm{Tr}[\hat{\rho}_m^{(\theta)}\hat{B}_m^{(\omega)}]$ in Eq. (G6) equals the average success probability of the POVM $\{\hat{B}_m^{(\omega)}\}_{m=0}^{3}$ for the quaternary states $\{\hat{\rho}_m^{(\theta)}\}_{m=0}^{3}$. Let $P_\theta^\star$ be the average success probability of a minimum-error measurement for $\{\hat{\rho}_m^{(\theta)}\}$; then, from Eq. (G6), we have

$$\upsilon_\omega^+(\lambda) \leqslant \frac{\lambda+1}{2}P_\theta^\star. \tag{G8}$$

This gives

$$\upsilon(\lambda) = \max_\omega \upsilon_\omega^+(\lambda) \leqslant \frac{\lambda+1}{2}\max_\theta P_\theta^\star. \tag{G9}$$

By the symmetry of the problem, we may, without loss of generality, consider only the case $0 \leqslant \theta \leqslant \pi/3$ (i.e., $\mathrm{Tr}\,\hat{\rho}_0^{(\theta)} \leqslant \mathrm{Tr}\,\hat{\rho}_2^{(\theta)} \leqslant \mathrm{Tr}\,\hat{\rho}_1^{(\theta)}$). Using the method described in Ref. [63] (the method of Ref. [64] can also be used), we can obtain an analytical expression of $P_\theta^\star$ for each θ. To avoid cumbersome details, we do not give an analytical expression of $P_\theta^\star$, but note that $P_\theta^\star$ achieves its maximum value if and only if $\theta = 0$ holds and satisfies

$$P_\theta^\star \leqslant P_0^\star = \begin{cases}\frac{2+\sqrt{3}}{4(\lambda+1)} & \text{for } \lambda \leqslant \frac{1}{2} + \frac{1}{2\sqrt{3}}\\ \frac{\lambda(3\lambda-1)}{2(\lambda+1)(2\lambda-1)} & \text{otherwise,}\end{cases} \tag{G10}$$

where we assume $\lambda \leqslant 1$ to simplify the discussion (it is sufficient to consider only this case, as described in the main text). From Eqs. (G9) and (G10) we have

$$\upsilon(\lambda) \leqslant \frac{\lambda+1}{2} P_0^\star = \begin{cases} \frac{2+\sqrt{3}}{8} & \text{for } \lambda \leqslant \frac{1}{2} + \frac{1}{2\sqrt{3}} \\ \frac{\lambda(3\lambda-1)}{4(2\lambda-1)} & \text{otherwise.} \end{cases} \tag{G11}$$

A minimum-error measurement, denoted by $\{\hat{B}_m^\bullet\}_{m=0}^3$, for the states $\{\hat{\rho}_m^{(0)}\}_{m=0}^3$ (i.e., in the case of $\theta = 0$) is given by

$$\hat{B}_m^\bullet = |B_m\rangle\langle B_m|, \quad |B_0\rangle = 0,$$
$$|B_1\rangle = \sqrt{\frac{1}{2}}\,|0\rangle - \sqrt{\frac{1-\alpha}{2}}\,|1\rangle, \tag{G12}$$
$$|B_2\rangle = \sqrt{\frac{1}{2}}\,|0\rangle + \sqrt{\frac{1-\alpha}{2}}\,|1\rangle,$$
$$|B_3\rangle = \sqrt{\alpha}\,|1\rangle,$$

where

$$\alpha = \frac{2(6\lambda^2 - 6\lambda + 1)}{3(2\lambda-1)^2}. \tag{G13}$$

Let ω_0 be in Ω such that $\{\hat{B}_m^{(\omega_0)}\} = \{\hat{B}_m^\bullet\}$. In the case of $\omega = \omega_0$, from Eqs. (G1) and (G2), Eq. (G4) with $\theta = 0$ holds. Since, in this case, $\upsilon_{\omega_0}^+(\lambda) = \frac{\lambda+1}{2} P_0^\star$ holds, the equality in Eq. (G11) holds. By substituting this into Eq. (22) and optimizing λ, we obtain Eq. (67). From Eq. (67), in the case of $\lambda = \lambda^\star$, we have

$$\alpha = \frac{4p_{\rm I}}{3}. \tag{G14}$$

[1] J. Walgate, A. J. Short, L. Hardy, and V. Vedral, Phys. Rev. Lett. **85**, 4972 (2000).
[2] B. Groisman and L. Vaidman, J. Phys. A: Math. Gen. **34**, 6881 (2001).
[3] J. Walgate and L. Hardy, Phys. Rev. Lett. **89**, 147901 (2002).
[4] H. Fan, Phys. Rev. Lett. **92**, 177905 (2004).
[5] M. Nathanson, J. Math. Phys. **46**, 062103 (2005).
[6] S. Bandyopadhyay, S. Ghosh, and G. Kar, New J. Phys. **13**, 123013 (2011).
[7] Z.-C. Zhang, Q.-Y. Wen, F. Gao, G.-J. Tian, and T.-Q. Cao, Quantum Inf. Process. **13**, 795 (2014).
[8] S. Virmani, M. F. Sacchi, M. B. Plenio, and D. Markham, Phys. Lett. A **288**, 62 (2001).
[9] D. Brody and B. Meister, Phys. Rev. Lett. **76**, 1 (1996).
[10] A. Acín, E. Bagan, M. Baig, L. Masanes, and R. Muñoz-Tapia, Phys. Rev. A **71**, 032338 (2005).
[11] A. Assalini, N. Dalla Pozza, and G. Pierobon, Phys. Rev. A **84**, 022342 (2011).
[12] Y.-X. Chen and D. Yang, Phys. Rev. A **64**, 064303 (2001).
[13] Y.-X. Chen and D. Yang, Phys. Rev. A **65**, 022320 (2002).
[14] Z. Ji, H. Cao, and M. Ying, Phys. Rev. A **71**, 032323 (2005).
[15] A. S. Holevo, J. Multivariate Anal. **3**, 337 (1973).
[16] H. P. Yuen, K. S. Kennedy, and M. Lax, IEEE Trans. Inf. Theory **21**, 125 (1975).
[17] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
[18] O. Hirota and S. Ikehara, IEICE Trans. **E65**, 627 (1982).
[19] G. M. D'Ariano, M. F. Sacchi, and J. Kahn, Phys. Rev. A **72**, 032310 (2005).
[20] K. Kato, in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)* (IEEE, Piscataway, NJ, 2012), pp. 1077–1081.
[21] I. D. Ivanovic, Phys. Lett. A **123**, 257 (1987).
[22] D. Dieks, Phys. Lett. A **126**, 303 (1988).
[23] A. Peres, Phys. Lett. A **128**, 19 (1988).
[24] A. Chefles and S. M. Barnett, J. Mod. Opt. **45**, 1295 (1998).
[25] Y. C. Eldar, Phys. Rev. A **67**, 042309 (2003).
[26] J. Fiurášek and M. Ježek, Phys. Rev. A **67**, 012321 (2003).
[27] M. A. P. Touzel, R. B. A. Adamson, and A. M. Steinberg, Phys. Rev. A **76**, 062314 (2007).
[28] A. Hayashi, T. Hashimoto, and M. Horibe, Phys. Rev. A **78**, 012333 (2008).
[29] H. Sugimoto, T. Hashimoto, M. Horibe, and A. Hayashi, Phys. Rev. A **80**, 052322 (2009).
[30] K. Nakahira, K. Kato, and T. S. Usuda, Phys. Rev. A **91**, 052304 (2015).
[31] M. Ban, K. Yamazaki, and O. Hirota, Phys. Rev. A **55**, 22 (1997).
[32] M. Owari and M. Hayashi, New J. Phys. **10**, 013006 (2008).
[33] K. Nakahira and T. S. Usuda, Phys. Rev. A **86**, 052323 (2012).
[34] K. Nakahira and T. S. Usuda, IEEE Trans. Inf. Theory **64**, 613 (2018).
[35] M. Rosati, A. Mari, and V. Giovannetti, Phys. Rev. A **96**, 012317 (2017).
[36] S. Croke, S. M. Barnett, and G. Weir, Phys. Rev. A **95**, 052308 (2017).
[37] K. Nakahira, K. Kato, and T. S. Usuda, Phys. Rev. A **97**, 022320 (2018).
[38] V. P. Belavkin, Stochastics **1**, 315 (1975).
[39] M. Ban, K. Kurokawa, R. Momose, and O. Hirota, Int. J. Theor. Phys. **36**, 1269 (1997).
[40] T. S. Usuda, I. Takumi, M. Hata, and O. Hirota, Phys. Lett. A **256**, 104 (1999).
[41] Y. C. Eldar and G. D. Forney, Jr., IEEE Trans. Inf. Theory **47**, 858 (2001).
[42] Y. C. Eldar, A. Megretski, and G. C. Verghese, IEEE Trans. Inf. Theory **50**, 1198 (2004).
[43] Y. C. Eldar, M. Stojnic, and B. Hassibi, Phys. Rev. A **69**, 062318 (2004).
[44] K. Nakahira and T. S. Usuda, Phys. Rev. A **87**, 012308 (2013).
[45] K. Nakahira, K. Kato, and T. S. Usuda, Phys. Rev. A **88**, 032314 (2013).
[46] E. Andersson, S. M. Barnett, C. R. Gilson, and K. Hunter, Phys. Rev. A **65**, 052308 (2002).
[47] K. Kato and O. Hirota, IEEE Trans. Inf. Theory **49**, 3312 (2003).
[48] D. Qiu, Phys. Rev. A **77**, 012328 (2008).
[49] A. Assalini, G. Cariolaro, and G. Pierobon, Phys. Rev. A **81**, 012315 (2010).
[50] K. Nakahira, K. Kato, and T. S. Usuda, Phys. Rev. A **91**, 012318 (2015).
[51] M. Osaki, M. Ban, and O. Hirota, Phys. Rev. A **54**, 1691 (1996).
[52] I. Ekeland and R. Temam, *Convex Analysis and Variational Problems* (SIAM, Philadelphia, 1999), Vol. 28.

[53] E. Chitambar and M.-H. Hsieh, Phys. Rev. A **88**, 020302 (2013).
[54] E. Chitambar, M. Hsieh, and R. Duan, IEEE Trans. Inf. Theory **60**, 1549 (2014).
[55] K. Nakahira, T. S. Usuda, and K. Kato, Phys. Rev. A **86**, 032316 (2012).
[56] I. Ivonovic, J. Phys. A: Math. Gen. **14**, 3241 (1981).
[57] T. Durt, B.-G. Englert, I. Bengtsson, and K. Życzkowski, Int. J. Quantum Inf. **8**, 535 (2010).
[58] D. G. Luenberger, *Optimization by Vector Space Methods* (Wiley, New York, 1969).
[59] A. Dhara and J. Dutta, *Optimality Conditions in Convex Optimization: A Finite-Dimensional View* (CRC Press, Boca Raton, FL, 2011).
[60] G. Chiribella, G. M. D'Ariano, and D. Schlingemann, Phys. Rev. Lett. **98**, 190403 (2007).
[61] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, 2009).
[62] M. Krein and D. Milman, Stud. Math. **19**, 133 (1940).
[63] M. E. Deconinck and B. M. Terhal, Phys. Rev. A **81**, 062304 (2010).
[64] M. Rosati, G. De Palma, A. Mari, and V. Giovannetti, Phys. Rev. A **95**, 042307 (2017).

Correction: The text and in-line expressions that previously appeared as the second sentence below Eq. (52) have been expanded and corrected.