Iterative methods for finding optimal quantum measurements under minimum-error and minimax criteria

Kenji Nakahira
Yokohama Research Laboratory, Hitachi, Ltd., Yokohama, Kanagawa 244-0817, Japan
and Quantum Information Science Research Center, Quantum ICT Research Institute, Tamagawa University, Machida, Tokyo 194-8610, Japan

Kentaro Kato
Quantum Communication Research Center, Quantum ICT Research Institute, Tamagawa University, Machida, Tokyo 194-8610, Japan

Tsuyoshi Sasaki Usuda
School of Information Science and Technology, Aichi Prefectural University, Nagakute, Aichi 480-1198, Japan
and Quantum Information Science Research Center, Quantum ICT Research Institute, Tamagawa University, Machida, Tokyo 194-8610, Japan
(Received 4 April 2014; revised manuscript received 1 October 2014; published 14 January 2015)

We investigate the problem of computing optimal quantum measurements in both minimal measuring and minimax strategies. A Belavkin weighted square-root measurement (BWSRM) with appropriate weights can represent the measurement that maximizes the correct probability for any given prior probabilities of quantum states. Using this fact, we propose methods for computing optimal solutions by optimizing the weights of the BWSRM. First, we explain the conditions for the BWSRM to be optimal. In particular, we argue that if a BWSRM with certain weights is a minimax measurement, then the minimax probabilities can be immediately obtained. Next, we propose an extension of the iterative algorithm developed by Ježek *et al.* [Phys. Rev. A **65**, 060301 (2002)] for maximizing the correct probability. We prove that, for a linearly independent pure state set, Ježek *et al.*'s algorithm converges to an optimal measurement. We also propose an iterative algorithm for a minimax solution and prove that, for a pure state set, our algorithm monotonically decreases the difference between estimated and true minimax values. Finally the performance of our algorithms is evaluated through numerical experiments.

DOI: 10.1103/PhysRevA.91.012318 PACS number(s): 03.67.Hk

I. INTRODUCTION

Discriminating between quantum states is one of the fundamental problems in quantum information science. Since the pioneering work of Helstrom, Holevo, and Yuen *et al.* [1–3], the problem of finding a measurement that can discriminate a set of quantum states as accurately as possible has been widely studied. There are different criteria for the detection of quantum states. A great deal of work has been devoted to the strategy of minimum-error discrimination (see, e.g., [4–11]). In this criterion, under the assumption that the receiver knows the probability distribution of the quantum states, we find a measurement that minimizes the average probability of a detection error. In contrast, Hirota and Ikehara discussed the quantum minimax strategy, which minimizes the worst case of the average probability of a detection error under the assumption that the receiver has no knowledge of the prior probabilities of the quantum states [12]. Several studies have also been reported on the quantum minimax strategy [13–16].

There are some cases in which analytical solutions for a minimum-error measurement are known. Necessary and sufficient conditions for a minimum-error measurement have been derived [1,3,17]. These conditions are often used to prove the optimality of measurements [5,7,10,11]. Necessary and sufficient conditions for a minimax solution are also known [12,15,16]. However, it is usually difficult to obtain a closed-form analytical expression for an optimal solution. Analytical optimal solutions are not known with some exceptions, such as the case in which the dimension of the state space is small or that in which a state set has some kind of symmetry.

We can use numerical methods instead of analytical approaches. The design of a minimum-error measurement can be formulated as a semidefinite programming (SDP) problem [18]. The problem of obtaining a minimax solution can also be expressed as SDP [16]. In many cases, an optimal solution can be computed in polynomial time by exploiting well-known algorithms for solving semidefinite programs with interior point methods. However, these classical methods are often inefficient since they require an excessive number of iterations in large scale problems (e.g., [19]). In several areas, such as signal and image processing, alternative numerical algorithms have recently been proposed that can effectively solve several types of optimization problems (see, e.g., [20–23]). In quantum signal detection theory, Ježek *et al.* proposed an iterative algorithm for obtaining a minimum-error measurement [24]. Their algorithm can be applied to general problems for minimum-error measurement and can effectively compute a solution in many cases. Later Tyson generalized the iterative scheme of Ježek *et al.* and proved that Ježek *et al.*'s algorithm monotonically increases the correct probability [25] (as he described in Ref. [25], Reimpell *et al.* also proved this in a different context [26,27]). An iterative algorithm for obtaining a minimax solution was also proposed [28]. However, it has not been guaranteed that the solutions of these algorithms are the global optimal ones. Moreover, their approach may require considerable computational effort at least if naively implemented.

To reduce the computational complexity, suboptimal measurements are sometimes used. In particular, the square-root measurement (SRM), which is also referred to as the pretty good measurement (PGM), has several desirable properties, thus widely studied [5,6,8,29–33]. The SRM can be constructed directly from the given collection of states and is optimal for several problems where quantum states have high symmetries. A generalized SRM referred to as a Belavkin weighted SRM (BWSRM) (also called a generalized PGM or a weighted least-squares measurement), has also been investigated [8,25,34–37]. The BWSRM can be described using a few parameters called weights. It is shown that a minimum-error measurement for arbitrary given prior probabilities can be expressed by a BWSRM with appropriate weights [34].

We propose methods of computing optimal quantum measurements in both the minimal measuring and minimax strategies by optimizing the weights of the BWSRM. In Sec. II, we describe a minimum-error measurement and a minimax solution. In Sec. III, we describe a BWSRM and conditions for the BWSRM to be optimal. In Sec. IV, we explain our algorithm of numerically obtaining a minimum-error measurement. Our algorithm is a generalization of Ježek *et al.*'s iterative algorithm [24] with less computational cost. We also discuss a method for estimating the precision of the estimated maximum correct probability by computing its upper and lower bounds at each iteration. Moreover, we prove that, in the case of a linearly independent pure state set, Ježek *et al.*'s algorithm converges to a minimum-error measurement. In Sec. V, we explain our algorithm of numerically obtaining a minimax solution and of obtaining upper and lower bounds for the correct probability of a minimax solution. We also derive that, in the case of a pure state set, the difference between estimated and true minimax values is monotonically decreasing in our algorithm. We show that computational cost can be reduced for a group covariant state set. In Sec. VI, we evaluate the performance of our algorithms through numerical experiments. In Sec. VII, we investigate the time and space complexity of our algorithms.

II. OPTIMAL DETECTION OF QUANTUM STATES

We consider the discrimination between M quantum states represented by density operators ρ_m ($m \in \mathcal{I}_M$), where $\mathcal{I}_k = \{0,1,\ldots,k-1\}$. The density operator ρ_m satisfies $\rho_m \geqslant 0$ and has unit trace ($\mathrm{Tr}\rho_m = 1$), where $A \geqslant 0$ denotes that A is positive semidefinite (similarly, $A \geqslant B$ denotes that $A - B$ is positive semidefinite). A set of quantum states, $\rho = \{\rho_m : m \in \mathcal{I}_M\}$, is called a quantum state set. Let $\mathcal{H}$ be the Hilbert space spanned by the supports of the operators $\{\rho_m : m \in \mathcal{I}_M\}$, which we refer to as the state space of ρ. A state with a rank-1 density operator is called a pure state. A state that is not a pure state is called a mixed state. A pure state set has only pure states, and a mixed state set has at least one mixed state.

A quantum measurement that distinguishes ρ can be modeled by a positive operator-valued measure (POVM) $\Pi = \{\Pi_m : m \in \mathcal{I}_M\}$. Without loss of generality, we can assume that each detection operator Π_m is on $\mathcal{H}$. In this paper we use the matrix representation with respect to an orthonormal basis of $\mathcal{H}$. Let $N = \dim \mathcal{H}$ and $\mathcal{S}^n$ be the entire set of n-dimensional positive semidefinite matrices. In this case, $\rho_m \in \mathcal{S}^N$ and $\Pi_m \in \mathcal{S}^N$ hold for each $m \in \mathcal{I}_M$.

Let $\mathcal{M}$ be the entire set of POVMs. Each $\Pi \in \mathcal{M}$ satisfies

$$\Pi_m \geqslant 0, \quad \forall m \in \mathcal{I}_M,$$
$$\sum_{m=0}^{M-1} \Pi_m = I, \tag{1}$$

where I is the identity matrix on $\mathcal{H}$.

Let ξ_m be the prior probability of the state ρ_m. To simplify the notation, $\{\xi_m^2 : m \in \mathcal{I}_M\}$ is denoted as ξ^2 for any collection of prior probabilities $\xi = \{\xi_m : m \in \mathcal{I}_M\}$. The average probability of correct detection, $P_{\mathrm{C}}(\xi,\Pi)$, is expressed as

$$P_{\mathrm{C}}(\xi,\Pi) = \sum_{m=0}^{M-1} \xi_m \mathrm{Tr}(\rho_m \Pi_m). \tag{2}$$

The average probability of a detection error is $1 - P_{\mathrm{C}}(\xi,\Pi)$.

In the minimal measuring strategy, the receiver knows the collection of prior probabilities. Let

$$P_{\mathrm{C}}^{\mathrm{opt}}(\xi) = \max_{\Pi \in \mathcal{M}} P_{\mathrm{C}}(\xi,\Pi). \tag{3}$$

The task of the receiver is to obtain a POVM $\Pi \in \mathcal{M}$ maximizing $P_{\mathrm{C}}(\xi,\Pi)$, i.e., satisfying $P_{\mathrm{C}}(\xi,\Pi) = P_{\mathrm{C}}^{\mathrm{opt}}(\xi)$. We call such Π a minimum-error measurement for ρ with ξ. If we consider a minimum-error measurement, we assume that $\xi_m > 0$ for any $m \in \mathcal{I}_M$. In the case in which ξ is not given, however, we can apply the minimax strategy. Under this strategy, a receiver attempts to find a POVM $\Pi^\star \in \mathcal{M}$ that maximizes the worst-case correct probability over the prior probabilities. We can regard $P_{\mathrm{C}}(\xi,\Pi)$ as both a continuous convex function of ξ for fixed Π and a continuous concave function of Π for fixed ξ. Thus, from the minimax theorem in convex analysis [38], we can prove that there exists a pair $(\xi^\star,\Pi^\star)$, which we call a minimax solution for ρ, such that (see [15] for details)

$$\min_{\xi \in \mathcal{X}} \max_{\Pi \in \mathcal{M}} P_{\mathrm{C}}(\xi,\Pi) = P_{\mathrm{C}}(\xi^\star,\Pi^\star) = \max_{\Pi \in \mathcal{M}} \min_{\xi \in \mathcal{X}} P_{\mathrm{C}}(\xi,\Pi), \tag{4}$$

where $\mathcal{X}$ is the entire set of possible prior probabilities $\xi = \{\xi_m : m \in \mathcal{I}_M\}$. Let us call $\xi^\star$ minimax probabilities, $\Pi^\star$ a minimax measurement, and $P_{\mathrm{C}}^\star = P_{\mathrm{C}}(\xi^\star,\Pi^\star)$ the minimax value. From Eq. (4), $P_{\mathrm{C}}^\star = P_{\mathrm{C}}^{\mathrm{opt}}(\xi^\star)$ holds.

III. BELAVKIN WEIGHTED SQUARE-ROOT MEASUREMENTS (BWSRMS)

A. Definition of BWSRMs

Let $R_m = \mathrm{rank}\rho_m$. Also, let ψ_m be the $N \times R_m$ complex matrix whose kth column is $\sqrt{\lambda_{m,k}}\,|\psi_{m,k}\rangle$, where $\lambda_{m,k}$ and $|\psi_{m,k}\rangle \in \mathcal{H}$ are the nonzero eigenvalue and corresponding unit eigenvector, respectively, of ρ_m. Then, we can write $\rho_m = \psi_m \psi_m^\dagger$ ($A^\dagger$ denotes the conjugate transpose of A).

Let w, which we will call weights, be a set of R_m-dimensional positive semidefinite matrices w_m, i.e., $w = \{w_m \in \mathcal{S}^{R_m} : m \in \mathcal{I}_M\}$. The BWSRM with weights w, which

we denote as $\Pi^{(w)} = \{\Pi_m^{(w)} \in \mathcal{S}^N : m \in \mathcal{I}_M\}$, is the POVM expressed as (see Sec. 2.2 of Ref. [34])

$$\Pi_m^{(w)} = [G^{(w)}]^{-1/2} \psi_m w_m \psi_m^\dagger [G^{(w)}]^{-1/2}, \tag{5}$$

where $G^{(w)} \in \mathcal{S}^N$ is defined by

$$G^{(w)} = \sum_{k=0}^{M-1} \psi_k w_k \psi_k^\dagger. \tag{6}$$

In this paper, we assume $w \in \mathcal{W}$, where $\mathcal{W}$ is the entire set of $w = \{w_m \in \mathcal{S}^{R_m} : m \in \mathcal{I}_M\}$ satisfying $G^{(w)} > 0$ ($A > 0$ denotes that A is strictly positive semidefinite). In this case, since $G^{(w)} > 0$ holds, $[G^{(w)}]^{-1/2}$ exists. Moreover, since $w_m \geqslant 0$ holds, $\Pi_m^{(w)}$ is guaranteed to be positive semidefinite.

In the case of a pure state set, since $R_m = 1, \psi_m$ (also denoted as $|\psi_m\rangle$) is a vector in $\mathcal{H}$ and w_m is a non-negative real number. From Eqs. (5) and (6), $\Pi_m^{(w)}$ is expressed as $\Pi_m^{(w)} = |\pi_m^{(w)}\rangle \langle \pi_m^{(w)}|$, where [4,36]

$$\begin{aligned} \left|\pi_m^{(w)}\right\rangle &= [G^{(w)}]^{-1/2} \sqrt{w_m} \, |\psi_m\rangle \,, \\ G^{(w)} &= \sum_{k=0}^{M-1} w_k \, |\psi_k\rangle \, \langle \psi_k| \,. \end{aligned} \tag{7}$$

Note that $\Pi^{(w)}$ satisfies

$$\mathrm{Tr}\left[\rho_m \Pi_m^{(w)}\right] = \mathrm{Tr}\left[Y_m^{(w)} w_m Y_m^{(w)}\right], \tag{8}$$

where $Y_m^{(w)} \in \mathcal{S}^{R_m}$ is defined by

$$Y_m^{(w)} = \psi_m^\dagger [G^{(w)}]^{-1/2} \psi_m. \tag{9}$$

$Y_m^{(w)} > 0$ holds from $[G^{(w)}]^{-1/2} > 0$. $\Pi^{(w)}$ is scale invariant with respect to w, that is, $\Pi^{(w)} = \Pi^{(\{k w_m\})}$ for any positive real number k.

Any minimum-error measurement and minimax measurement can be expressed as a BWSRM, with appropriate weights, as shown later. The probability of correctly detecting the state ρ_m, $\mathrm{Tr}[\rho_m \Pi_m^{(w)}]$ tends to increase as the trace of the corresponding weight $\mathrm{Tr} w_m$ relatively increases.

B. Examples of BWSRMs

The first example is the SRM (also referred to as the Belavkin-Hausladen-Wotters PGM [25,37]). This measurement can be denoted as the BWSRM $\Pi^{(w)}$ with $w = \{\xi_m I_m \in \mathcal{S}^{R_m} : m \in \mathcal{I}_M\}$, where I_m is an R_m-dimensional identity matrix. From Eq. (5), $\Pi_m^{(w)}$ is expressed as

$$\Pi_m^{(w)} = \left(\sum_{k=0}^{M-1} \xi_k \rho_k\right)^{-1/2} \xi_m \rho_m \left(\sum_{k=0}^{M-1} \xi_k \rho_k\right)^{-1/2}. \tag{10}$$

In particular, in the case of a pure state set, $w = \xi = \{\xi_m\}$ holds. From Eq. (7), we have $\Pi_m^{(w)} = |\pi_m^{(w)}\rangle \langle \pi_m^{(w)}|$ with

$$\left|\pi_m^{(w)}\right\rangle = \left(\sum_{k=0}^{M-1} \xi_k \, |\psi_k\rangle \, \langle \psi_k|\right)^{-1/2} \sqrt{\xi_m} \, |\psi_m\rangle \,. \tag{11}$$

The second example is the quadratically weighted measurement (QWM) [25,36,37,39], which is equivalent to the BWSRM $\Pi^{(w)}$ with $w = \{\xi_m^2 \psi_m^\dagger \psi_m \in \mathcal{S}^{R_m} : m \in \mathcal{I}_M\}$. From Eq. (5), we find that

$$\Pi_m^{(w)} = \left(\sum_{k=0}^{M-1} \xi_k^2 \rho_k^2\right)^{-1/2} \xi_m^2 \rho_m^2 \left(\sum_{k=0}^{M-1} \xi_k^2 \rho_k^2\right)^{-1/2}. \tag{12}$$

In particular, in the case of a pure state set, we have $w = \xi^2 = \{\xi_m^2\}$. From Eq. (7), we have $\Pi_m^{(w)} = |\pi_m^{(w)}\rangle \langle \pi_m^{(w)}|$ with

$$\left|\pi_m^{(w)}\right\rangle = \left(\sum_{k=0}^{M-1} \xi_k^2 \, |\psi_k\rangle \, \langle \psi_k|\right)^{-1/2} \xi_m \, |\psi_m\rangle \,. \tag{13}$$

C. Conditions of minimum-error measurements

Now we introduce an important remark for investigating a minimum-error measurement expressed as a BWSRM. The following remark shows that any minimum-error measurement can be always expressed as a BWSRM and provides a necessary and sufficient condition for optimality.

Remark 1 (Theorem 3 of Ref. [34]). We consider a collection of M quantum states, $\rho = \{\rho_m = \psi_m \psi_m^\dagger : m \in \mathcal{I}_M\}$, with prior probabilities $\xi = \{\xi_m : m \in \mathcal{I}_M\}$.

A necessary and sufficient condition for a POVM $\Pi = \{\Pi_m : m \in \mathcal{I}_M\}$ to be a minimum-error measurement is that Π is identical to the BWSRM $\Pi^{(w)} = \{\Pi_m^{(w)} : m \in \mathcal{I}_M\}$ with weights $w = \{w_m \in \mathcal{S}^{R_m} : m \in \mathcal{I}_M\} \in \mathcal{W}$ such that there exists a positive real number c satisfying

$$\xi_m Y_m^{(w)} \leqslant c I_m, \quad \forall m \in \mathcal{I}_M, \tag{14}$$

$$\xi_m Y_m^{(w)} w_m = c w_m, \quad \forall m \in \mathcal{I}_M. \tag{15}$$

From Eq. (15), each column of w_m is an eigenvector of $Y_m^{(w)}$ with eigenvalue $c\xi_m^{-1}$. This implies that Eqs. (14) and (15) are equivalent to the statement that $Y_m^{(w)}$ can be expressed as

$$\begin{aligned} Y_m^{(w)} &= c\xi_m^{-1}\left(I_m^{\mathrm{supp}} \oplus Z_m^{(w)}\right), \\ Z_m^{(w)} &\leqslant I_m^{\mathrm{Ker}}, \end{aligned} \tag{16}$$

where I_m^{supp} is the identity matrix on the space spanned by the supports of w_m, and $Z^{(w)}$ and I_m^{Ker} are respectively a $(R_m - \mathrm{rank} w_m)$-dimensional positive semidefinite matrix and the $(R_m - \mathrm{rank} w_m)$-dimensional identity matrix on the kernel of w_m. In particular, if $w_m > 0$, then Eq. (16) is equivalent to

$$\xi_m Y_m^{(w)} = c I_m. \tag{17}$$

Note that a POVM Π is a minimum-error measurement if and only if an operator $X \in \mathcal{S}^N$ exists such that [2,3]

$$X - \xi_m \rho_m \geqslant 0, \quad \forall m \in \mathcal{I}_M, \tag{18}$$

$$(X - \xi_m \rho_m)\Pi_m = 0, \quad \forall m \in \mathcal{I}_M. \tag{19}$$

We can find that Eqs. (14) and (15) are respectively identical to Eqs. (18) and (19). Indeed, substituting $w_m = c^{-2}\xi_m^2 \psi_m^\dagger \Pi_m \psi_m$ and $X = c[G^{(w)}]^{1/2}$, and after some algebra, Eqs. (14) and (15) follow from Eqs. (18) and (19), and vice versa (see [4,34] for details). Also note that Eldar *et al.* argued that a sufficient condition for the SRM to be a minimum-error measurement is

that c exists such that (Theorem 1 of Ref. [40])

$$\xi_m \psi_m^\dagger \left(\sum_{k=0}^{M-1} \xi_k \rho_k \right)^{-1/2} \psi_m = c I_m. \tag{20}$$

Since the SRM is the BWSRM with $\{w_m = \xi_m I_m\}$, Eq. (20) is a special case of Eq. (17). Since $w_m > 0$, Eq. (20) is also a necessary condition.

D. Conditions of minimax solutions

In this subsection, we present a proposition that is useful for considering a minimax solution in which the minimax measurement is expressed by a BWSRM. In preparation, we introduce the following remark.

Remark 2 (Theorem 2 of Ref. [12]). We consider a state set $\rho = \{\rho_m : m \in \mathcal{I}_M\}$. Let $\xi = \{\xi_m : m \in \mathcal{I}_M\}$ be prior probabilities and $\Pi = \{\Pi_m : m \in \mathcal{I}_M\}$ be a POVM. Assume $\xi_m > 0$ for any $m \in \mathcal{I}_M$ [41]. Then, (ξ, Π) is a minimax solution for ρ if and only if Π is a minimum-error measurement for ρ with ξ and $\mathrm{Tr}(\rho_m \Pi_m)$ is a constant independent of $m \in \mathcal{I}_M$.

Using Remarks 1 and 2, we can prove the following proposition.

Proposition 3. We consider a state set $\rho = \{\rho_m = \psi_m \psi_m^\dagger : m \in \mathcal{I}_M\}$. We also consider a BWSRM $\Pi^{(w)} = \{\Pi_m^{(w)} : m \in \mathcal{I}_M\}$, where $w = \{w_m \in \mathcal{S}^{R_m} : m \in \mathcal{I}_M\} \in \mathcal{W}$ are weights. Assume $\mathrm{Tr} w_m > 0$ for any $m \in \mathcal{I}_M$. Let $\xi^{(w)}$ be prior probabilities $\xi^{(w)} = \{\xi_m^{(w)} : m \in \mathcal{I}_M\}$ with

$$\xi_m^{(w)} = \frac{\sqrt{\mathrm{Tr} w_m}}{\sum_{k=0}^{M-1} \sqrt{\mathrm{Tr} w_k}}. \tag{21}$$

The following three statements are equivalent:

(1) $\Pi^{(w)}$ is a minimum-error measurement for ρ with $\xi^{(w)}$.

(2) $\Pi^{(w)}$ is a minimax measurement for ρ.

(3) A positive real number c exists such that for any $m \in \mathcal{I}_M$,

$$\begin{aligned} \xi_m^{(w)} Y_m^{(w)} &\leqslant c I_m, \\ \xi_m^{(w)} Y_m^{(w)} w_m &= c w_m. \end{aligned} \tag{22}$$

Moreover, if one of the above statements is true, then the following statement is also true:

(4) $\xi^{(w)}$ are minimax probabilities.

In particular, if ρ is a pure state set, then the above four statements are equivalent.

In the case of a pure state set, we obtain $\xi_m^{(w)} = \sqrt{w_m}$ by normalizing w_m such that $\sum_{k=0}^{M-1} \sqrt{w_k} = 1$. This means that we can identify w with $[\xi^{(w)}]^2 = \{[\xi_m^{(w)}]^2 : m \in \mathcal{I}_M\}$ when obtaining a minimax solution. Proposition 3 indicates that if minimax probabilities $\xi^\star$ are obtained, then a minimax measurement is obtained as the QWM $\Pi^{([\xi^\star]^2)}$ for a pure state set. Thus, in this case, the problem of obtaining a minimax solution can be reduced to finding minimax probabilities $\xi^\star$.

In contrast, in the case of a mixed state set, we cannot identify w with $[\xi^{(w)}]^2$ since w is a set of matrices. However, it follows from Remark 1 that $w^\star \in \mathcal{W}$ exists such that $\Pi^{(w^\star)}$ is a minimax measurement. Moreover, from Proposition 3, if such $w^\star$ is obtained, then minimax probabilities are obtained as $\xi^{(w^\star)}$ of Eq. (21). Thus, in this case, the minimax problem can be reduced to finding $w^\star$.

Proof. Since (1) $\Leftrightarrow$ (3) from Remark 1, it is sufficient to show (2) $\Leftrightarrow$ (3) and (3) $\Rightarrow$ (4) for any state set, and (4) $\Rightarrow$ (1) for any pure state set.

First, we prove (3) $\Rightarrow$ (2) and (3) $\Rightarrow$ (4). Assume that c exists satisfying Eq. (22). From Eq. (22), we find that

$$\left[\xi_m^{(w)}\right]^2 Y_m^{(w)} w_m Y_m^{(w)} = c^2 w_m. \tag{23}$$

By taking the trace on both sides of this equation and substituting Eq. (8), we obtain

$$\begin{aligned} \left[\xi_m^{(w)}\right]^2 \mathrm{Tr}\left[\rho_m \Pi_m^{(w)}\right] &= c^2 \mathrm{Tr} w_m \\ &= c^2 \left[\xi_m^{(w)}\right]^2 \left(\sum_{k=0}^{M-1} \sqrt{\mathrm{Tr} w_k} \right)^2, \end{aligned} \tag{24}$$

where the second line follows from Eq. (21). In contrast, $\xi_m^{(w)} > 0$ holds from $\mathrm{Tr} w_m > 0$. Thus, from Eq. (24), $\mathrm{Tr}[\rho_m \Pi_m^{(w)}]$ is a constant independent of $m \in \mathcal{I}_M$. In contrast, $\Pi^{(w)}$ is a minimum-error measurement for ρ with $\xi^{(w)}$ since (1) $\Leftrightarrow$ (3). Therefore, from Remark 2, $(\xi^{(w)}, \Pi^{(w)})$ is a minimax solution.

Next, we prove (2) $\Rightarrow$ (3). Assume that $\Pi^{(w)}$ is a minimax measurement. Let ξ be the corresponding minimax probabilities. Since $\Pi^{(w)}$ is a minimum-error measurement for ρ with ξ, c exists such that Eqs. (14) and (15) hold. From Eq. (15), we find that

$$\xi_m^2 Y_m^{(w)} w_m Y_m^{(w)} = c^2 w_m. \tag{25}$$

Taking the trace on both sides of this equation and substituting Eq. (8) yield

$$\xi_m^2 \mathrm{Tr}\left[\rho_m \Pi_m^{(w)}\right] = c^2 \mathrm{Tr} w_m. \tag{26}$$

Since $c > 0$ and $\mathrm{Tr} w_m > 0$ hold, $\xi_m > 0$ holds. In contrast, from Remark 2, $\mathrm{Tr}[\rho_m \Pi_m^{(w)}]$ is a constant independent of $m \in \mathcal{I}_M$. Therefore, from Eq. (26), $\xi_m \propto \sqrt{\mathrm{Tr} w_m}$ (i.e., $\xi = \xi^{(w)}$) holds, which indicates from Eqs. (14) and (15) that c exists such that Eq. (22) holds.

Finally we prove (4) $\Rightarrow$ (1) for a pure state set. Assume that $\xi^{(w)}$ are minimax probabilities. From Remark 1, a minimax measurement, which is also a minimum-error measurement for ρ with $\xi^{(w)}$, can be represented as a BWSRM $\Pi^{(w')}$ with certain weights w'. Now we will show that $\Pi^{(w')} = \Pi^{(w)}$. Since $\xi_m^{(w)} > 0$ for any $m \in \mathcal{I}_M$, from Eq. (8) and Remark 2, we have

$$w'_m \left[Y_m^{(w')}\right]^2 = \mathrm{Tr}\left[\rho_m \Pi_m^{(w')}\right] = P_{\mathrm{C}}^{\mathrm{opt}}[\xi^{(w)}]. \tag{27}$$

The second equality follows from the fact that $\mathrm{Tr}[\rho_m \Pi_m^{(w')}]$ is independent of $m \in \mathcal{I}_M$. It is obvious that $P_{\mathrm{C}}^{\mathrm{opt}}[\xi^{(w)}] > 0$ holds, which gives $w'_m > 0$ from Eq. (27). Substituting $\xi_m = \xi_m^{(w)}$ and $w_m = w'_m$ into Eq. (15) and dividing both sides by w'_m yield that $\xi_m^{(w)} Y_m^{(w')}$ is a constant independent of $m \in \mathcal{I}_M$. In contrast, as derived from Eq. (27), $\sqrt{w'_m} Y_m^{(w')}$ is also a constant independent of m. Therefore, $\xi_m^{(w)} Y_m^{(w')} \propto \sqrt{w'_m} Y_m^{(w')}$, i.e., $\xi_m^{(w)} \propto \sqrt{w'_m}$, holds. Since $\xi_m^{(w)} \propto \sqrt{w_m}$ also holds from Eq. (21), we have $w'_m \propto w_m$, which yields $\Pi^{(w')} = \Pi^{(w)}$. ■

Note that in the case of a mixed state set, statement (4) of Proposition 3 is usually not equivalent to statements (1)–(3). This can be explained as follows. Assume that $\Pi^{(w)}$ is a minimax measurement. Let $\varpi = \{\varpi_m = u_m w_m u_m^\dagger \in \mathcal{S}^{R_m} : m \in \mathcal{I}_M\} \in \mathcal{W}$, where u_m is an R_m-dimensional unitary matrix.

Since $\mathrm{Tr}\varpi_m = \mathrm{Tr}w_m$ holds and statement (4) is true for w, $\xi^{(\varpi)} = \xi^{(w)}$ are minimax probabilities. However, in general $\Pi^{(\varpi)}$ is not a minimax measurement.

IV. CALCULATION OF MINIMUM-ERROR MEASUREMENT

We consider a quantum state set $\rho = \{\rho_m = \psi_m\psi_m^\dagger : m \in \mathcal{I}_M\}$ with prior probabilities $\xi = \{\xi_m : m \in \mathcal{I}_M\}$. In this section we present our iterative algorithm for finding weights $w = \{w_m \in \mathcal{S}^{R_m} : m \in \mathcal{I}_M\}$ of the BWSRM $\Pi^{(w)} = \{\Pi_m^{(w)} : m \in \mathcal{I}_M\}$ so that the correct probability $P_{\mathrm{C}}[\xi, \Pi^{(w)}]$ is as close as possible to the maximum one $P_{\mathrm{C}}^{\mathrm{opt}}(\xi)$ (denoted by $P_{\mathrm{C}}^{\mathrm{opt}}$ for simplicity). Our algorithm is based on an iterative algorithm developed by Ježek *et al.* [24]. We then give upper and lower bounds for the maximum correct probability. We also argue that Ježek *et al.*'s algorithm is guaranteed to converge to a global optimization solution for a linearly independent pure state set.

A. Iterative algorithm

Ježek *et al.* proposed the following iterative formula [24]:

$$\Pi_m^{(r+1)} = \xi_m^2[\Lambda^{(r)}]^{-1/2}\rho_m\Pi_m^{(r)}\rho_m[\Lambda^{(r)}]^{-1/2}, \qquad \Lambda^{(r)} = \sum_{k=0}^{M-1}\xi_k^2\rho_k\Pi_k^{(r)}\rho_k, \tag{28}$$

where $r \in \{0,1,2,\dots\}$. We assume that Π^t is the BWSRM with certain weights $w^t = \{w_m^t : m \in \mathcal{I}_M\} \in \mathcal{W}$, i.e., $\Pi^t = \Pi^{(w^t)}$, for a certain $t \geqslant 0$, and will show that $\Pi^{(t+1)}$ is also a BWSRM. Let us denote $G^{(r)}$ and $Y_m^{(r)}$ as $G^{(r)} = G^{(w^{(r)})}$ and $Y_m^{(r)} = Y_m^{(w^{(r)})}$ for simplicity. We also assume, for any $r \geqslant 0$,

$$w_m^{(r+1)} = \xi_m^2 Y_m^{(r)} w_m^{(r)} Y_m^{(r)}. \tag{29}$$

We have

$$\begin{aligned}\xi_k^2\rho_k\Pi_k^t\rho_k &= \xi_k^2\rho_k[G^t]^{-1/2}\psi_k w_k^t\psi_k^\dagger[G^t]^{-1/2}\rho_k \\ &= \xi_k^2\psi_k Y_k^t w_k^t Y_k^t\psi_k^\dagger \\ &= \psi_k w_k^{(t+1)}\psi_k^\dagger,\end{aligned} \tag{30}$$

where the first and second lines follow from Eqs. (5) and (9), respectively. Thus, we have that, from Eqs. (6) and (28),

$$\Lambda^t = \sum_{k=0}^{M-1}\psi_k w_k^{(t+1)}\psi_k^\dagger = G^{(t+1)}. \tag{31}$$

Equations (28), (30), and (31) give

$$\begin{aligned}\Pi_m^{(t+1)} &= [G^{(t+1)}]^{-1/2}\psi_m w_m^{(t+1)}\psi_m^\dagger[G^{(t+1)}]^{-1/2} \\ &= \Pi_m^{(w^{(t+1)})},\end{aligned} \tag{32}$$

where the last line follows from Eq. (5). Equation (32) indicates that $\Pi^{(t+1)}$ is the BWSRM with weights $w^{(t+1)}$. Therefore, as far as an initial measurement $\Pi^{(0)}$ is chosen to be the BWSRM $\Pi^{(w^{(0)})}$ with certain weights $w^{(0)} \in \mathcal{W}$, $\Pi^{(r)} = \Pi^{(w^{(r)})}$ holds for any $r \geqslant 0$, where $w^{(r)} \in \mathcal{W}$ are the weights satisfying Eq. (29). This means that Eq. (29) can also be used as an iterative formula. Initial weights $w^{(0)}$ may be set to $w_m^{(0)} = \xi_m I_m$ (i.e., $\Pi^{(0)}$ are the SRM) or $w_m^{(0)} = \xi_m^2\psi_m^\dagger\psi_m$ (i.e., $\Pi^{(0)}$ are the QWM).

Now let us consider the following iterative formula instead of Eq. (29):

$$w_m^{(r+1)} = \xi_m^{2d}\left[Y_m^{(r)}\right]^d w_m^{(r)}\left[Y_m^{(r)}\right]^d, \tag{33}$$

where d is a positive real number, which we call the acceleration parameter. It is easy to verify that Eq. (33) with $d = 1$ is equivalent to Eq. (29). The entire algorithm for finding a minimum-error measurement is found in the following pseudocode (the stopping criterion, i.e., steps 3 and 4, will be shown in Sec. IV C).

Finding a minimum-error measurement:
Input: prior probabilities $\xi = \{\xi_m\}$,
quantum states $\rho = \{\rho_m = \psi_m\psi_m^\dagger\}$,
a constant for the stopping criteria δP_{C},
and an acceleration parameter d.
1. Initialize $w^{(0)}$
(e.g., $w_m^{(0)} = \xi_m I_m$ or $w_m^{(0)} = \xi_m^2\psi_m^\dagger\psi_m$).
2.**for** $r = 0,1,\dots$ **do**
/* Decide whether to stop */
3. Compute $\underline{P_{\mathrm{C}}}^{(w^{(r)})}$ and $\overline{P_{\mathrm{C}}}^{(w^{(r)})}$ from Eqs. (39) and (40), respectively.
4. **if** $\overline{P_{\mathrm{C}}}^{(w^{(r)})} - \underline{P_{\mathrm{C}}}^{(w^{(r)})} < \delta P_{\mathrm{C}}$ **then break**
/* Update the weights */
5. Compute
$G^{(r)} = \sum_{k=0}^{M-1}\psi_k w_k^{(r)}\psi_k^\dagger$,
$Y_m^{(r)} = \psi_m^\dagger[G^{(r)}]^{-1/2}\psi_m$,
$w_m^{(r+1)} = \xi_m^{2d}[Y_m^{(r)}]^d w_m^{(r)}[Y_m^{(r)}]^d$.
6. **end for**
7. Compute $\Pi^{(w^{(r)})}$ from Eq. (5).
Output: the POVM $\Pi^{(w^{(r)})}$.

Like the Ježek *et al.*'s algorithm, the proposed algorithm is based on a necessary condition for a minimum-error measurement. Equation (28) implies that if $\Pi^{(r)}$ converges to Π, then Π satisfies

$$\Pi_m = \left(\sum_{k=0}^{M-1}\xi_k^2\rho_k\Pi_k\rho_k\right)^{-1/2}\xi_m\rho_m\Pi_m, \tag{34}$$

which is a necessary condition for a minimum-error measurement [24]. Similarly, Eq. (33) guarantees that if $w_m^{(r)}$ converges to w_m, then w_m satisfies Eq. (15) with $c = 1$. Indeed, from Remark 10 in Appendix A with $A = \xi_m Y_m^{(w)}$ and $B = w_m$, Eq. (15) with $c = 1$ is equivalent to

$$w_m = \xi_m^{2d}\left[Y_m^{(w)}\right]^d w_m\left[Y_m^{(w)}\right]^d. \tag{35}$$

We can interpret that our algorithm tries to find a fixed point of the function $f(w_m) = \xi_m^{2d}[Y_m^{(w)}]^d w_m[Y_m^{(w)}]^d$ using Eq. (33).

We expect to speed up our algorithm using an appropriate acceleration parameter d with $d > 1$. For example, if $Y_m^{(r+1)} \approx Y_m^{(r)}$ holds, then we obtain

$$\begin{aligned}&\xi_m^2 Y_m^{(r+1)}\left[\xi_m^2 Y_m^{(r)} w_m^{(r)} Y_m^{(r)}\right]Y_m^{(r+1)} \\ &\approx \xi_m^4\left[Y_m^{(r)}\right]^2 w_m^{(r)}\left[Y_m^{(r)}\right]^2.\end{aligned} \tag{36}$$

Therefore, one iteration of Eq. (33) with $d = 2$ is approximately equivalent to two iterations of Eq. (29). As discussed in Sec. VI, numerical experiments show that Eq. (33) with $d > 1$ tends to convergence faster than Eq. (29).

Our algorithm updates weights $w^{(r)}$ instead of a POVM $\Pi^{(r)}$, which provides two advantages. One advantage is low computational cost. Since the size of $w^{(r)}$ is usually smaller than that of $\Pi^{(r)}$ (in particular, $w^{(r)}$ is a set of M real numbers in the case of a pure state set), our algorithm is expected to require lower computational time and memory than Ježek *et al.*'s algorithm at least if naively implemented. The other advantage is easy to accelerate the convergence speed by tuning the parameter d. The acceleration technique of our method is difficult to apply to Ježek *et al.*'s algorithm. Indeed, one may consider a formula such as

$$\Pi_m^{(r+1)} = \xi_m^{2d}([\Lambda^{(r)}]^{-1/2}\rho_m)^d \Pi_m^{(r)} (\rho_m [\Lambda^{(r)}]^{-1/2})^d \quad (37)$$

instead of Eq. (28). However, $\Pi^{(r+1)}$ given by Eq. (37) does not satisfy Eq. (1).

B. Example

We show an example of finding a minimum-error measurement using the proposed algorithm. We consider a binary pure state set $\rho = \{|\psi_0\rangle\langle\psi_0|, |\psi_1\rangle\langle\psi_1|\}$ satisfying $|\langle\psi_0|\psi_1\rangle| = 0.8$. Prior probabilities are set to $\xi_0 = 0.3$ and $\xi_1 = 0.7$. Since ρ is a pure state set, i.e., $R_0 = R_1 = 1$, $w_m^{(r)}$ and $Y_m^{(r)}$ are nonnegative real numbers. The acceleration parameter is set to $d = 1$.

Figures 1(a) and 1(b) respectively show $w_m^{(r)}$ and $\xi_m Y_m^{(r)}$. We set the initial weights to $w_0^{(0)} = \xi_0^2 = 0.09$ and $w_1^{(0)} = \xi_1^2 = 0.49$, which means that the POVM $\Pi^{(w^{(0)})}$ is the QWM. Note that the correct probability of the QWM is larger than that of the SRM [36]. In the first iteration, we compute $Y_m^{(0)}$ from Eq. (9). In this example, we have $\{\xi_m Y_m^{(0)}\} \approx \{0.789, 0.965\}$ ($\approx$ denotes approximately equal). Substituting this into Eq. (33) yields $w_m^{(1)} = [\xi_m Y_m^{(0)}]^2 w_m^{(0)} \approx \{0.0561, 0.456\}$. $\xi_0 Y_0^{(0)} < 1$ and $\xi_1 Y_1^{(0)} < 1$ hold, which gives $w_0^{(1)} < w_0^{(0)}$ and $w_1^{(1)} < w_1^{(0)}$. $\{\xi_m Y_m^{(r)}\}$ converges to $\{1, 1\}$ in this example (note that, from Eq. (33), if $w_m^{(r)}$ converges to a positive number, then $\xi_m Y_m^{(r)}$ converges to 1). The optimal weights w are $w \approx \{0.0519, 0.467\}$. We can see that weights $w^{(r)}$ converge to the optimal ones as r increases.

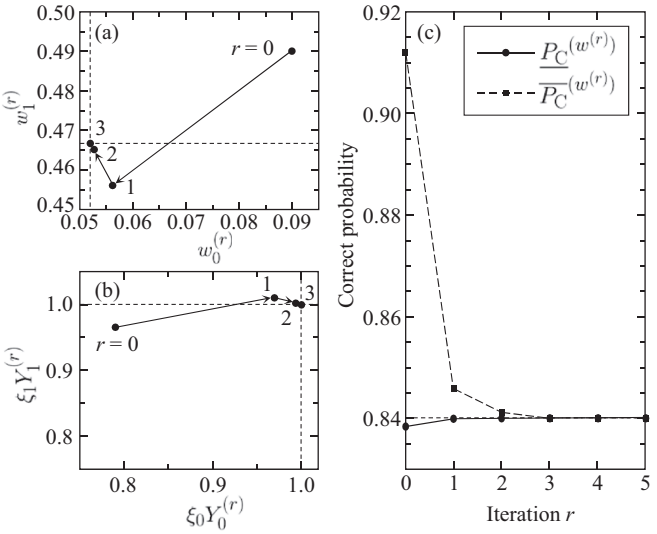

FIG. 1. Example of the result of our method for finding a minimum-error measurement in the case of a binary pure state set: (a) the transition of $w_m^{(r)}$; (b) the transition of $\xi_m Y_m^{(r)}$; (c) the lower and upper bounds for P_C^{opt} vs number of iterations.

Figure 1(c) shows the lower and upper bounds for P_C^{opt} (i.e., $\underline{P_C}^{(w^{(r)})}$ and $\overline{P_C}^{(w^{(r)})}$), which are discussed in Sec. IV C. P_C^{opt} can be obtained by [1]

$$P_C^{\text{opt}} = (1 + \sqrt{1 - 4\xi_0\xi_1 |\langle\psi_0|\psi_1\rangle|^2})/2 = 0.84. \quad (38)$$

We can see that both $\underline{P_C}^{(w^{(r)})}$ and $\overline{P_C}^{(w^{(r)})}$ converge to P_C^{opt}.

C. Estimation of accuracy of maximum correct probability

Obtaining the maximum correct probability P_C^{opt} is usually difficult; instead, we consider obtaining lower and upper bounds for P_C^{opt} based on the theory of SDP. More precisely, we exploit the fact that feasible solutions to the primal and dual problems for finding a minimum-error measurement always provide lower and upper bounds, respectively. In the proposed algorithm, these bounds are used as the stopping criterion.

Let $\underline{P_C}^{(w)}$ be the correct probability of the BWSRM $\Pi^{(w)}$, that is,

$$\underline{P_C}^{(w)} = P_C[\xi, \Pi^{(w)}] = \sum_{k=0}^{M-1} \xi_k \text{Tr}\left[Y_k^{(w)} w_k Y_k^{(w)}\right]. \quad (39)$$

It is obvious that $\underline{P_C}^{(w)}$ is a lower bound for P_C^{opt}.

The following proposition gives an upper bound.

Proposition 4. We consider a state set $\rho = \{\rho_m = \psi_m \psi_m^\dagger : m \in \mathcal{I}_M\}$ with prior probabilities $\xi = \{\xi_m : m \in \mathcal{I}_M\}$. Let $w \in \mathcal{W}$. Also let

$$\overline{P_C}^{(w)} = \text{Tr}[G^{(w)}]^{1/2} + \sum_{k=0}^{M-1} \left(\xi_k - \frac{1}{\lambda_0\left[Y_k^{(w)}\right]}\right)^+, \quad (40)$$

where $\lambda_0[Y_k^{(w)}]$ is the largest eigenvalue of $Y_k^{(w)}$, and $(x)^+$ is defined as x if $x > 0$ and zero otherwise. Then, $\overline{P_C}^{(w)} \geqslant P_C^{\text{opt}}$ holds for any $w \in \mathcal{W}$. Moreover, if w satisfies Eqs. (14) and (15) with $c = 1$ (note that the BWSRM $\Pi^{(w)}$ is a minimum-error measurement from Remark 1), then $\overline{P_C}^{(w)} = P_C^{\text{opt}}$ holds.

Proof. First, we prove $\overline{P_C}^{(w)} \geqslant P_C^{\text{opt}}$ for any $w \in \mathcal{W}$. Let

$$X^{(w)} = [G^{(w)}]^{1/2} + \sum_{k=0}^{M-1} \left(1 - c_k^{(w)}\right)^+ \xi_k \rho_k, \quad (41)$$

where $c_k^{(w)}$ is the largest real number such that

$$[G^{(w)}]^{1/2} \geqslant c_k^{(w)} \xi_k \rho_k. \quad (42)$$

From Eq. (41), we have that for any $m \in \mathcal{I}_M$,

$$\begin{aligned} X^{(w)} - \xi_m \rho_m &\geqslant \sum_{k=0}^{M-1} \left(1 - c_k^{(w)}\right)^+ \xi_k \rho_k - \left[1 - c_m^{(w)}\right] \xi_m \rho_m \\ &\geqslant \left(1 - c_m^{(w)}\right)^+ \xi_m \rho_m - \left[1 - c_m^{(w)}\right] \xi_m \rho_m \\ &\geqslant 0. \end{aligned} \quad (43)$$

In contrast, any positive semidefinite matrix X with $X \geqslant \xi_m \rho_m$ for any $m \in \mathcal{I}_M$, which means that X is a feasible solution

to the dual problem, satisfies $\mathrm{Tr}X \geqslant P_{\mathrm{C}}^{\mathrm{opt}}$ (see Theorem 1 of Ref. [18]). From Eq. (43), $X^{(w)}$ is a feasible solution of the dual problem, and thus $\mathrm{Tr}X^{(w)} \geqslant P_{\mathrm{C}}^{\mathrm{opt}}$ holds. Therefore, it suffices to show that $\mathrm{Tr}X^{(w)} = \overline{P_{\mathrm{C}}}^{(w)}$. Using Remark 11 in Appendix A (with $c = c_k^{(w)}\xi_k, \Phi = \psi_k$, and $B = [G^{(w)}]^{-1/2}$) and substituting $Y_k^{(w)} = \psi_k^\dagger [G^{(w)}]^{-1/2}\psi_k$, it follows that Eq. (42) is identical to $c_k^{(w)}\xi_k Y_k^{(w)} \leqslant I_m$. Since $c_k^{(w)}$ is the largest real number satisfying this inequality, $c_k^{(w)} = (\xi_k \lambda_0[Y_k^{(w)}])^{-1}$ holds. Substituting this into Eq. (41) and taking the trace, we obtain $\mathrm{Tr}X^{(w)} = \overline{P_{\mathrm{C}}}^{(w)}$.

Next, we assume that w satisfies Eqs. (14) and (15) with $c = 1$ and prove $\overline{P_{\mathrm{C}}}^{(w)} = P_{\mathrm{C}}^{\mathrm{opt}}$. Let

$$X = \sum_{k=0}^{M-1} \xi_k \rho_k \Pi_k^{(w)}. \tag{44}$$

Since $\Pi^{(w)}$ is a minimum-error measurement, $X \geqslant \xi_m \rho_m$ holds for any $m \in \mathcal{I}_M$ [17]. In contrast, from Remark 12 in Appendix B, $[G^{(w)}]^{1/2} = X$ holds. Thus, we have $[G^{(w)}]^{1/2} \geqslant \xi_m \rho_m$, which means $c_k^{(w)} = 1$. Therefore, from Eqs. (41) and (44), we have

$$\overline{P_{\mathrm{C}}}^{(w)} = \mathrm{Tr}X^{(w)} = \mathrm{Tr}[G^{(w)}]^{1/2} = \mathrm{Tr}X = P_{\mathrm{C}}^{\mathrm{opt}}. \tag{45}$$

■

From Eq. (39) and Proposition 4, for any $w \in \mathcal{W}, P_{\mathrm{C}}^{\mathrm{opt}}$ satisfies

$$\underline{P_{\mathrm{C}}}^{(w)} = P_{\mathrm{C}}[\xi, \Pi^{(w)}] \leqslant P_{\mathrm{C}}^{\mathrm{opt}} \leqslant \overline{P_{\mathrm{C}}}^{(w)}. \tag{46}$$

In particular, if w satisfies Eqs. (14) and (15) with $c = 1$, then $\Pi^{(w)}$ is a minimum-error measurement and $\underline{P_{\mathrm{C}}}^{(w)} = P_{\mathrm{C}}^{\mathrm{opt}} = \overline{P_{\mathrm{C}}}^{(w)}$ holds. $\underline{P_{\mathrm{C}}}^{(w)}$ and $\overline{P_{\mathrm{C}}}^{(w)}$ are expected to be close to $P_{\mathrm{C}}^{\mathrm{opt}}$ under the condition that $\Pi^{(w^{(r)})}$ is close to a minimum-error measurement as the iteration proceeds.

Note that it is not guaranteed that $\overline{P_{\mathrm{C}}}^{(w^{(r)})}$ is monotonically decreasing with respect to r. Considering this, we can use $\min_{t \leqslant r} \overline{P_{\mathrm{C}}}^{(w^{(t)})}$ as an upper bound of $P_{\mathrm{C}}^{\mathrm{opt}}$ instead of $\overline{P_{\mathrm{C}}}^{(w^{(r)})}$ at each iteration in our proposed algorithm. In contrast, $\underline{P_{\mathrm{C}}}^{(w^{(r)})} = P_{\mathrm{C}}[\xi, \Pi^{(w^{(r)})}]$ is monotonically increasing with respect to r, as described in the next subsection.

D. Convergence of solution

Ježek *et al.* reported the numerical observation that their algorithm monotonically increases the correct probability at each iteration and converges to the maximum correct probability [24]. Later, Reimpell *et al.* proposed an algorithm for maximizing entanglement fidelity of quantum channels [26], which is a generalization of Ježek *et al.*'s algorithm (see [25]). Reimpell *et al.* also proved the monotonic increase of their algorithm [26,27], which means that Ježek *et al.*'s algorithm monotonically increases the correct probability. Moreover, Tyson gave a geometric interpretation of these algorithms by embedding the set of POVMs (or completely positive maps) in a semidefinite inner product space, and generalized Reimpell *et al.*'s monotonicity result to a more general iteration [25]. However, the convergence to the optimality has not been proved yet (see [25]). The following theorem shows that in the case of a linearly independent pure state set Ježek *et al.*'s algorithm converges to a minimum-error measurement (proof in Appendix C).

Theorem 5. We consider a linearly independent pure state set $\rho = \{|\psi_m\rangle\langle\psi_m| : m \in \mathcal{I}_M\}$ with prior probabilities $\xi = \{\xi_m : m \in \mathcal{I}_M\}$. Using the iterative formula of Eq. (29) [i.e., Eq. (33) with $d = 1$] starting from initial weights $w^{(0)} = \{w_m^{(0)} > 0 : m \in \mathcal{I}_M\}$, the BWSRM $\Pi^{(w^{(r)})}$ converges to a minimum-error measurement.

V. CALCULATION OF MINIMAX SOLUTION

We now consider a quantum state set $\rho = \{\rho_m = \psi_m \psi_m^\dagger : m \in \mathcal{I}_M\}$. We present our iterative algorithm for finding weights $w = \{w_m \in \mathcal{S}^{R_m} : m \in \mathcal{I}_M\}$ of the BWSRM $\Pi^{(w)} = \{\Pi_m^{(w)} : m \in \mathcal{I}_M\}$ so that $\Pi^{(w)}$ is as close as possible to a minimax measurement. Note again that, from Proposition 3, $\xi^{(w)}$ is minimax probabilities if $\Pi^{(w)}$ is a minimax measurement.

A. Iterative algorithm

In the previous section, we discussed the iterative formula of Eq. (33) by focusing on the fact that Eq. (15), which is a necessary condition for a minimum-error measurement, holds if and only if Eq. (35) holds for a positive real number d. In a similar manner, we construct an iterative algorithm for a minimax solution using a necessary condition for a minimax solution. To be more precise, we use the fact that $\Pi^{(w)}$ is a minimax measurement if and only if Eq. (22) holds, as given in Proposition 3.

First, we give an iterative formula for a pure state set. As stated in Proposition 3, for a pure state set, the QWM $\Pi^{([\xi^\star]^2)}$ is a minimax measurement for minimax probabilities $\xi^\star$. Here we assume that $w = \xi^2$ holds (i.e., $\Pi^{(w)}$ is the QWM) during each iteration and update prior probabilities ξ such that ξ converges to $\xi^\star$. We introduce the following formula:

$$\begin{aligned} \xi_m^{(r+1)} &= \Xi_m^{(d)}[\xi^{(r)}], \\ \Xi_m^{(d)}(\xi) &= \frac{\xi_m}{\left[\xi_m Y_m^{(\xi^2)}\right]^d} \left(\sum_{k=0}^{M-1} \frac{\xi_k}{[\xi_k Y_k^{(\xi^2)}]^d} \right)^{-1}, \end{aligned} \tag{47}$$

where $r \in \{0,1,2,\dots\}$. The basic idea is based on Remark 13 in Appendix B. Indeed, $\Xi_m^{(1)}(\xi)$ of Eq. (47) is identical to $\Xi_m(\xi^2)$ defined by Eq. (B3). From Remark 13, $\Pi^{([\xi^{(r)}]^2)}$ is a minimum-error measurement for ρ with $\xi^{(r+1)}$ if $d = 1$. Here, $d > 0$ is an acceleration parameter. It follows from Eq. (47) that ξ is a fixed point of $\Xi_m^{(d)}(\xi)$ if $\xi_m Y_m^{(\xi^2)}$ is a constant independent of $m \in \mathcal{I}_M$ (note that $c > 0$ exists such that $\xi_m Y_m^{(\xi^2)} = c$ in this case), i.e., ξ satisfies the necessary and sufficient condition of Eq. (22). Initial probabilities $\xi^{(0)}$ are chosen such that $\xi_m^{(0)} > 0$ and $\sum_{k=0}^{M-1} \xi_k^{(0)} = 1$ (for example, $\xi_m^{(0)} = 1/M$).

Next, we extend the iterative formula of Eq. (47) to a mixed state set. From Proposition 3, if optimal weights $w^\star$ are obtained, then $(\xi^{(w^\star)}, \Pi^{(w^\star)})$ is a minimax solution. To obtain an iterative formula with respect to w, let us consider the

following equation:

$$\left[\xi_m^{(w)}\right]^2 Y_m^{(w)} w_m Y_m^{(w)} = c^2 w_m, \tag{48}$$

which follows from Eq. (22). We normalize w such that $\sum_{k=0}^{M-1}\sqrt{\mathrm{Tr}w_k}=1$, which means $\xi_m^{(w)}=\sqrt{\mathrm{Tr}w_k}$ from Eq. (21). Then Eq. (48) can be rewritten as

$$\begin{aligned} w_m &= \left[\xi_m^{(w)}\right]^2 \frac{T_m}{\mathrm{Tr}T_m}, \\ T_m &= Y_m^{(w)} w_m Y_m^{(w)} \in \mathcal{S}^{R_m}. \end{aligned} \tag{49}$$

In consideration of this, we propose the following iterative formula:

$$\begin{aligned} w_m^{(r+1)} &= \left[\xi_m^{(r+1)}\right]^2 \frac{T_m^{(r)}}{\mathrm{Tr}T_m^{(r)}}, \\ \xi_m^{(r+1)} &= \frac{\sqrt{\mathrm{Tr}w_m^{(r)}}}{\left[\mathrm{Tr}T_m^{(r)}\right]^{d/2}} \left(\sum_{k=0}^{M-1} \frac{\sqrt{\mathrm{Tr}w_k^{(r)}}}{\left[\mathrm{Tr}T_k^{(r)}\right]^{d/2}}\right)^{-1}, \\ T_m^{(r)} &= Y_m^{(r)} w_m^{(r)} Y_m^{(r)}, \\ Y_m^{(r)} &= \psi_m^\dagger [G^{(r)}]^{-1/2} \psi_m, \\ G^{(r)} &= \sum_{k=0}^{M-1} \psi_k w_k^{(r)} \psi_k^\dagger, \end{aligned} \tag{50}$$

where $r \in \{0,1,2,\cdots\}$. Note that $Y_m^{(r)} = Y_m^{(w^{(r)})}$ and $G^{(r)} = G^{(w^{(r)})}$ hold. We can see that this formula is a generalization of Eq. (47). Indeed, in the case of a pure state set, the first line of Eq. (50) gives $w^{(r)} = [\xi^{(r)}]^2$ for any $r > 0$, and substituting this into the second line of Eq. (50) yields Eq. (47).

The proposed iterative algorithm can be summarized as follows (the stopping criterion will be shown in the next subsection):

Finding a minimax solution:
Input: quantum states $\rho = \{\rho_m = \psi_m \psi_m^\dagger\}$,
a constant for the stopping criteria δP_{C},
and an acceleration parameter d.
1. Initialize $w^{(0)}$ (e.g., $w_m^{(0)} = I_m/M^2$).
2. **for** $r = 0,1,\cdots$ **do**
/* Decide whether to stop */
3. Compute $\underline{P_{\mathrm{C}}^{\star(w^{(r)})}}$ and $\overline{P_{\mathrm{C}}^{\star(w^{(r)})}}$ from Eqs. (51) and (55), respectively.
4. **if** $\overline{P_{\mathrm{C}}^{\star(w^{(r)})}} - \underline{P_{\mathrm{C}}^{\star(w^{(r)})}} < \delta P_{\mathrm{C}}$ **then break**
/* Update the prior probabilities and weights */
5. Compute $w^{(r+1)}$ from Eq. (50) [Eq. (47) can also be used for a pure state set].
6. **end for**
7. Compute $\xi^{(w^{(r)})}$ and $\Pi^{(w^{(r)})}$ from Eqs. (21) and (5), respectively.
Output: the prior probabilities $\xi^{(w^{(r)})}$ and the POVM $\Pi^{(w^{(r)})}$.

B. Estimation of accuracy of a minimax value

We can calculate the upper and lower bounds for the minimax value $P_{\mathrm{C}}^\star$ at each iteration. Let

$$\begin{aligned} \underline{P_{\mathrm{C}}^{\star(w)}} &= \min_{m\in\mathcal{I}_M} \mathrm{Tr}\left[\rho_m \Pi_m^{(w)}\right] \\ &= \min_{m\in\mathcal{I}_M} \mathrm{Tr}\left[Y_m^{(w)} w_m Y_m^{(w)}\right]. \end{aligned} \tag{51}$$

$\underline{P_{\mathrm{C}}^{\star(w)}}$ is a lower bound for $P_{\mathrm{C}}^\star$ since $\underline{P_{\mathrm{C}}^{\star(w)}}$ satisfies

$$\underline{P_{\mathrm{C}}^{\star(w)}} = \min_{\xi} P_{\mathrm{C}}[\xi, \Pi^{(w)}] \leqslant P_{\mathrm{C}}[\xi^\star, \Pi^{(w)}] \leqslant P_{\mathrm{C}}^\star. \tag{52}$$

For a pure state set, it is known that Remark 13 in Appendix B holds, which yields, for any $w \in \mathcal{W}$,

$$P_{\mathrm{C}}[\Xi(w), \Pi^{(w)}] = P_{\mathrm{C}}^{\mathrm{opt}}[\Xi(w)] \geqslant \min_{\xi} P_{\mathrm{C}}^{\mathrm{opt}}(\xi) = P_{\mathrm{C}}^\star, \tag{53}$$

where $\Xi(w)$ is defined by Eq. (B3). Thus, $P_{\mathrm{C}}[\Xi(w), \Pi^{(w)}]$ (denoted as $\overline{P_{\mathrm{C}^{\mathrm{pure}(w)}}^\star}$) is an upper bound for $P_{\mathrm{C}}^\star$. From Eq. (8), $\overline{P_{\mathrm{C}^{\mathrm{pure}}}^{\star(w)}}$ is expressed as

$$\overline{P_{\mathrm{C}^{\mathrm{pure}}}^{\star(w)}} = \left(\sum_{k=0}^{M-1} \frac{1}{Y_k^{(w)}}\right)^{-1} \sum_{m=0}^{M-1} w_m Y_m^{(w)}. \tag{54}$$

It is obvious that if $\Pi^{(w)}$ is a minimax measurement, then $\underline{P_{\mathrm{C}}^{\star(w)}} = P_{\mathrm{C}}^\star = \overline{P_{\mathrm{C}^{\mathrm{pure}}}^{\star(w)}}$ holds. Although Remark 13 cannot be applied to a mixed state set, the following proposition gives an alternative upper bound.

Proposition 6. We consider a state set $\rho = \{\rho_m = \psi_m \psi_m^\dagger : m \in \mathcal{I}_M\}$. Let $w \in \mathcal{W}$. Also let

$$\overline{P_{\mathrm{C}}^{\star(w)}} = \left(\sum_{k=0}^{M-1} \frac{1}{\lambda_0\left[Y_k^{(w)}\right]}\right)^{-1} \mathrm{Tr}[G^{(w)}]^{1/2}, \tag{55}$$

where $\lambda_0[Y_k^{(w)}]$ is the largest eigenvalue of $Y_k^{(w)}$. Then, $\overline{P_{\mathrm{C}}^{\star(w)}} \geqslant P_{\mathrm{C}}^\star$ holds for any $w \in \mathcal{W}$, where $P_{\mathrm{C}}^\star$ is the minimax value. Moreover, if $\Pi^{(w)}$ is a minimax measurement, then $\overline{P_{\mathrm{C}}^{\star(w)}} = P_{\mathrm{C}}^\star$ holds.

Note that in the case of a pure state set $\overline{P_{\mathrm{C}}^{\star(w)}}$ is equivalent to $\overline{P_{\mathrm{C}^{\mathrm{pure}}}^{\star(w)}}$ since $\overline{P_{\mathrm{C}}^{\star(w)}}$ can also be expressed as

$$\overline{P_{\mathrm{C}}^{\star(w)}} = \left(\sum_{k=0}^{M-1} \frac{1}{\lambda_0\left[Y_k^{(w)}\right]}\right)^{-1} \mathrm{Tr} \sum_{m=0}^{M-1} w_m Y_m^{(w)}. \tag{56}$$

Indeed, this is given by

$$\begin{aligned} \mathrm{Tr} \sum_{m=0}^{M-1} w_m Y_m^{(w)} &= \mathrm{Tr} \sum_{m=0}^{M-1} w_m \psi_m^\dagger [G^{(w)}]^{-1/2} \psi_m \\ &= \mathrm{Tr} \sum_{m=0}^{M-1} \psi_m w_m \psi_m^\dagger [G^{(w)}]^{-1/2} \\ &= \mathrm{Tr}[G^{(w)}]^{1/2}, \end{aligned} \tag{57}$$

where the first and last lines follow from Eqs. (9) and (6), respectively.

Proof. First, we prove that $\overline{P_{\rm C}^{\star}}^{(w)} \geqslant P_{\rm C}^{\star}$ holds for any $w \in \mathcal{W}$. Let $c_m^{(w)}$ be the largest real number such that $[G^{(w)}]^{1/2} \geqslant c_m^{(w)}\rho_m$. Also let

$$\begin{aligned} \xi'_m &= C c_m^{(w)}, \\ X &= C[G^{(w)}]^{1/2}. \end{aligned} \tag{58}$$

where $C = [\sum_{k=0}^{M-1} c_k^{(w)}]^{-1}$. Then, we have that for any $m \in \mathcal{I}_M$,

$$X - \xi'_m\rho_m = C\left([G^{(w)}]^{1/2} - c_m^{(w)}\rho_m\right) \geqslant 0. \tag{59}$$

In contrast, from the theory of the SDP, if a positive semidefinite matrix X and prior probabilities ξ' satisfy $X - \xi'_m\rho_m \geqslant 0$, then $\mathrm{Tr}X \geqslant P_{\rm C}^{\rm opt}(\xi')$ holds (see Theorem 1 of Ref. [18]). Thus, we obtain

$$\mathrm{Tr}X \geqslant P_{\rm C}^{\rm opt}(\xi') \geqslant P_{\rm C}^{\star}. \tag{60}$$

Therefore, it suffices to show that $c_m^{(w)} = (\lambda_0[Y_k^{(w)}])^{-1}$, since, from Eqs. (55) and (58), $\overline{P_{\rm C}^{\star}}^{(w)} = \mathrm{Tr}X$ holds in this case.

Using Remark 11 in Appendix A (with $c = c_m^{(w)}, \Phi = \psi_m$, and $B = [G^{(w)}]^{-1/2}$) and substituting Eq. (9), it follows that $[G^{(w)}]^{1/2} \geqslant c_m^{(w)}\rho_m$ is equivalent to $c_m^{(w)}Y_m^{(w)} \leqslant I_m$. Therefore, $c_m^{(w)} = (\lambda_0[Y_m^{(w)}])^{-1}$ holds.

Next, we assume that $\Pi^{(w)}$ is a minimax measurement, and prove $\overline{P_{\rm C}^{\star}}^{(w)} = P_{\rm C}^{\star}$. From Proposition 3, Eq. (22) holds, which indicates that Eq. (16) with $\xi_m = \xi_m^{(w)}$ holds. Thus, the largest eigenvalue of $Y_k^{(w)}$ can be expressed by $\lambda_0[Y_k^{(w)}] = c[\xi_m^{(w)}]^{-1}$. Substituting this into Eq. (55) yields $\overline{P_{\rm C}^{\star}}^{(w)} = c\mathrm{Tr}[G^{(w)}]^{1/2}$. In contrast, from Remark 12 in Appendix B, we have

$$c[G^{(w)}]^{1/2} = \sum_{k=0}^{M-1} \xi_k^{(w)}\rho_k\Pi_k^{(w)}. \tag{61}$$

Taking the trace on both sides gives $c\mathrm{Tr}[G^{(w)}]^{1/2} = P_{\rm C}^{\star}$, which yields $\overline{P_{\rm C}^{\star}}^{(w)} = P_{\rm C}^{\star}$. ■

C. Convergence of solution

Here we show the monotonically decreasing property of $P_{\rm C}^{\rm opt}[\xi^{(r)}]$ for a pure state set when the acceleration parameter is $d = 1$. As described in Sec. V A, $\xi_m^{(r+1)} = \Xi_m^{(1)}[\xi^{(r)}] = \Xi_m([\xi^{(r)}]^2)$ holds for any $r \geqslant 0$ in the case of $d = 1$, which yields $\overline{P_{\rm C}^{\star}}^{(w)} = P_{\rm C}^{\rm opt}[\xi^{(r+1)}]$ from Eq. (53). Theorem 7 claims that the upper bound $\overline{P_{\rm C}^{\star}}^{(w)}$ is monotonically decreasing. We also show the condition of convergence toward an optimal solution in Theorem 8.

Theorem 7. We consider a pure state set $\rho = \{|\psi_m\rangle\langle\psi_m| : m \in \mathcal{I}_M\}$. Using the iterative formula of Eq. (47) with $d = 1$ starting from initial prior probabilities $\xi^{(0)}$ satisfying $\xi_m^{(0)} > 0$ for any $m \in \mathcal{I}_M$, $P_{\rm C}^{\rm opt}[\xi^{(r)}]$ is monotonically decreasing.

Proof. Let $F(r,r') = P_{\rm C}[\xi^{(r')}, \Pi^{([\xi^{(r)}]^2)}]$. $P_{\rm C}^{\rm opt}[\xi^{(r)}] \geqslant F(r,r)$ is obvious. Since, from Remark 13, $\xi^{(r+1)} = \Xi([\xi^{(r)}]^2)$ are prior probabilities such that $\Pi^{([\xi^{(r)}]^2)}$ is a minimum-error measurement, $P_{\rm C}^{\rm opt}[\xi^{(r+1)}] = F(r,r+1)$ holds. Thus, it is sufficient to show $F(r,r) \geqslant F(r,r+1)$ since in this case we obtain

$$P_{\rm C}^{\rm opt}[\xi^{(r)}] \geqslant F(r,r) \geqslant F(r,r+1) = P_{\rm C}^{\rm opt}[\xi^{(r+1)}]. \tag{62}$$

Let $K^{(r)} = \sum_{k=0}^{M-1}[Y_k^{(r)}]^{-1}$. Since

$$\xi_m^{(r+1)} = \Xi_m([\xi^{(r)}]^2) = \frac{1}{K^{(r)}Y_m^{([\xi^{(r)}]^2)}}, \tag{63}$$

i.e., $Y_m^{([\xi^{(r)}]^2)} = 1/[K^{(r)}\xi_m^{(r+1)}]$,

$$\begin{aligned} F(r,r') &= \sum_{m=0}^{M-1} \xi_m^{(r')}\left[\xi_m^{(r)}Y_m^{([\xi^{(r)}]^2)}\right]^2 \\ &= \frac{1}{[K^{(r)}]^2}\sum_{m=0}^{M-1} \xi_m^{(r')}\left[\frac{\xi_m^{(r)}}{\xi_m^{(r+1)}}\right]^2, \end{aligned} \tag{64}$$

where the first line follows from Eq. (8) with $w = [\xi^{(r)}]^2$. Thus,

$$\begin{aligned} &F(r,r) - F(r,r+1) \\ &= \frac{1}{[K^{(r)}]^2}\sum_{m=0}^{M-1} \xi_m^{(r)}\left\{\left[\frac{\xi_m^{(r)}}{\xi_m^{(r+1)}}\right]^2 - \frac{\xi_m^{(r)}}{\xi_m^{(r+1)}}\right\} \\ &\geqslant \frac{1}{[K^{(r)}]^2}\sum_{m=0}^{M-1} \xi_m^{(r)}\ln\left[\frac{\xi_m^{(r)}}{\xi_m^{(r+1)}}\right] \\ &= \frac{D_{\rm KL}[\xi^{(r)}\|\xi^{(r+1)}]}{[K^{(r)}]^2} \geqslant 0, \end{aligned} \tag{65}$$

where $D_{\rm KL}$ is the Kullback-Leibler divergence. The first inequality of Eq. (65) follows from $x^2 - x \geqslant \ln x$ for any $x > 0$. The second inequality follows from the fact that the Kullback-Leibler divergence is non-negative. ■

Theorem 8. We consider a state set $\rho = \{\psi_m\psi_m^\dagger : m \in \mathcal{I}_M\}$. Assume that $w_m^{(r)} \in \mathcal{S}^{R_m}$ in Eq. (50) converges to a strictly positive matrix (denoted by w_m) for each $m \in \mathcal{I}_M$. Then, $(\xi^{(w)}, \Pi^{(w)})$ is a minimax solution, where $w = \{w_m : m \in \mathcal{I}_M\}$.

Proof. From Eq. (50), $\xi_m^{(r)}, T_m^{(r)}$, and $Y_m^{(r)}$ converge. We denote their limits by ξ_m, T_m, and Y_m, respectively. Taking the trace on both sides of the first equality of Eq. (50) yields $[\xi_m^{(r+1)}]^2 = \mathrm{Tr}w_m^{(r+1)}$. Thus, from Eq. (21), $\xi^{(r)} = \xi^{(w^{(r)})}$ holds for any $r \geqslant 1$, which indicates $\xi = \xi^{(w)}$. From Eq. (50), Y_m equals to $Y_m^{(w)}$ of Eq. (9). Hence, from Proposition 3, it is sufficient to show that there exists a positive real number c satisfying Eq. (22).

Equation (50) yields

$$w_m = \left[\xi_m^{(w)}\right]^2\frac{Y_m^{(w)}w_mY_m^{(w)}}{\mathrm{Tr}T_m}. \tag{66}$$

Note that $\mathrm{Tr}T_m > 0$ since $w_m > 0$. Let $A = \xi_m^{(w)}Y_m^{(w)}/\sqrt{\mathrm{Tr}T_m}$ and $B = w_m$, then Eq. (66) is expressed as $B = ABA$. Since $A \geqslant 0$ and $B > 0$, from Remark 10 in Appendix A, $A = I_m$ holds. Therefore, $\xi_m^{(w)}Y_m^{(w)} = cI_m$ holds with $c = \sqrt{\mathrm{Tr}T_m}$, and thus $\xi_m^{(w)}Y_m^{(w)}w_m = cw_m$ holds, which means that Eq. (22) holds with $c = \sqrt{\mathrm{Tr}T_m}$. ■

D. Group covariant state sets

We consider a state set $\rho = \{\rho_m = \psi_m\psi_m^\dagger : m \in \mathcal{I}_M\}$ that is invariant under the action of a group $\mathcal{G}$ ($|\mathcal{G}| \geqslant 2$) in which each element $g \in \mathcal{G}$ corresponds to an N-dimensional unitary matrix U_g. Such ρ is called as a group covariant state set with respect to $\mathcal{G}$ [42]. For any $m \in \mathcal{I}_M$ and $g \in \mathcal{G}$, there exists

$k \in \mathcal{I}_M$ such that

$$U_g \rho_m U_g^\dagger = \rho_k. \tag{67}$$

Assume without loss of generality that Eq. (67) is equivalent to $U_g \psi_m = \psi_k$ (indeed, this is true if $\{\psi_m : m \in \mathcal{I}_M\}$ is appropriately set). To simplify the notation, we denote such k as $g \circ m$. Similarly, we denote $U_g A U_g^\dagger$ as $g \circ A$ for any matrix A. Equation (67) can be rewritten as

$$g \circ \rho_m = \rho_{g \circ m}. \tag{68}$$

$\mathcal{G}$ may include an antiunitary matrix (see [16,42] for details). A group covariant state set is a generalization of a so-called geometrically uniform state set [40], in which for any $m \in \mathcal{I}_M$ and $k \in \mathcal{I}_M$, there exists $g \in \mathcal{G}$ such that $k = g \circ m$.

For a group covariant state set ρ, a minimax solution $(\xi^\star, \Pi^\star)$ exists such that, for any $m \in \mathcal{I}_M$ and $g \in \mathcal{G}$ [16],

$$\begin{aligned} \xi_m^\star &= \xi_{g \circ m}^\star, \\ g \circ \Pi_m^\star &= \Pi_{g \circ m}^\star. \end{aligned} \tag{69}$$

Thus, a minimax solution is expected to be computed efficiently for a group covariant state set. Note that a computationally efficient numerical solution using SDP for quantum states satisfying geometrical symmetry in minimal measuring strategy has been proposed [43,44].

The following proposition holds.

Proposition 9. Let $\mathcal{G}$ be a group with $|\mathcal{G}| \geqslant 2$ in which each element $g \in \mathcal{G}$ corresponds to an N-dimensional unitary or antiunitary matrix U_g. We consider a group covariant state set $\rho = \{\rho_m = \psi_m \psi_m^\dagger : m \in \mathcal{I}_M\}$ satisfying $U_g \psi_m = \psi_{g \circ m}$ for any $m \in \mathcal{I}_M$ and $g \in \mathcal{G}$. We consider computing a minimax solution for ρ using the iterative formula of Eq. (50), where initial weights $w^{(0)}$ are chosen such that $w_{g \circ m}^{(0)} = w_m^{(0)}$ for any $m \in \mathcal{I}_M$ and $g \in \mathcal{G}$. Then, we have that for any $r \geqslant 0, m \in \mathcal{I}_M$, and $g \in \mathcal{G}$,

$$w_{g \circ m}^{(r)} = w_m^{(r)}. \tag{70}$$

This proposition implies that we need to calculate only one weight for each orbit; we do not need to calculate M weights, where the orbit of m is the set $\{g \circ m : g \in \mathcal{G}\}$. In the case of a geometrically uniform state set [40], which has just one orbit, we need only calculate one weight $w_0^{(r)}$ for each r. Note that a similar result can be easily obtained in the case of the iterative formula for finding a minimum-error measurement.

Proof. We prove this proposition by induction on r. By assumption, it is obvious for $r = 0$. Assume that, for a certain $t \geqslant 0, w_{g \circ m}^{(t)} = w_m^{(t)}$ for any $m \in \mathcal{I}_M$ and $g \in \mathcal{G}$. We have, for any $g \in \mathcal{G}$,

$$\begin{aligned} g \circ G^{(w^{(t)})} &= \sum_{m=0}^{M-1} g \circ \left[\psi_m w_m^{(t)} \psi_m^\dagger\right] \\ &= \sum_{m=0}^{M-1} \psi_{g \circ m} w_m^{(t)} \psi_{g \circ m}^\dagger \\ &= \sum_{m'=0}^{M-1} \psi_{m'} w_{m'}^{(t)} \psi_{m'}^\dagger \\ &= G^{(w^{(t)})}, \end{aligned} \tag{71}$$

where $m' = g \circ m$. Thus, $Y_m^{(t)}$ satisfies

$$\begin{aligned} Y_{g \circ m}^{(t)} &= \psi_{g \circ m}^\dagger [G^{(w^{(t)})}]^{-1/2} \psi_{g \circ m} \\ &= \psi_m^\dagger \{g^{-1} \circ [G^{(w^{(t)})}]^{-1/2}\} \psi_m \\ &= \psi_m^\dagger [G^{(w^{(t)})}]^{-1/2} \psi_m \\ &= Y_m^{(t)}. \end{aligned} \tag{72}$$

Therefore, from Eq. (50), $w_{g \circ m}^{(t+1)} = w_m^{(t+1)}$ holds for any $m \in \mathcal{I}_M$ and $g \in \mathcal{G}$. ■

VI. NUMERICAL EXAMPLES OF PROPOSED ALGORITHMS

In this section, we explain the performance of proposed iterative algorithms. First, we investigate the convergence properties for pure states using an M-ary optical amplitude shift keyed (ASK) coherent state set $\rho = \{|k\alpha\rangle \langle k\alpha| : k \in \{0, \pm 1, \ldots, \pm \lfloor M/2 \rfloor\}\}$, where $|k\alpha\rangle$ is the eigenvector of the photon annihilation operator corresponding to the eigenvalue $k\alpha$. Let $|\alpha|^2 = 0.1$ and M be odd. Next, we investigate the performance for mixed states using randomly selected states.

The convergence properties of the iterative algorithm for a minimum-error measurement, expressed by Eq. (33), for ρ with equal prior probabilities are shown in Fig. 2 for $d = 1.0$, 1.25, and 1.5. The initial weights are chosen to be $w_m^{(0)} = 1/M^2$ for each $m \in \mathcal{I}_M$. We plot the number of iterations r to achieve the difference between the upper and lower bounds for the maximum correct probability, i.e., $\overline{P_C}^{(w^{(r)})} - \underline{P_C}^{(w^{(r)})}$, less than 10^{-9} for different M. Note again that the result with $d = 1.0$ is the same as that of the iterative algorithm developed by Ježek *et al.* [24]. At least in the range of $1.0 \leqslant d \leqslant 1.5$, we can see that the number of iterations required decreases as d increases. We verified that solutions tend to diverge at about $d = 2.0$. d may change at each iteration, and our algorithm is expected to be accelerated if d is appropriately scheduled.

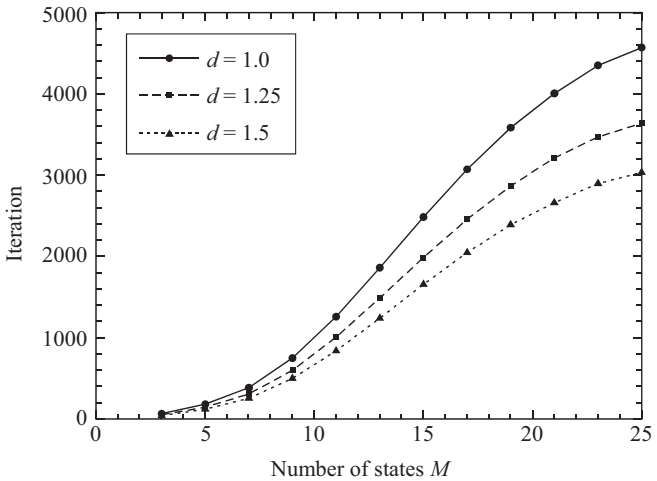

FIG. 2. Example of numerical calculations for minimum-error measurements for an M-ary optical ASK coherent state set. Number of iterations to achieve difference between upper and lower bounds for maximum correct probability, $\overline{P_C}^{(w^{(r)})} - \underline{P_C}^{(w^{(r)})}$, less than 10^{-9} is shown.

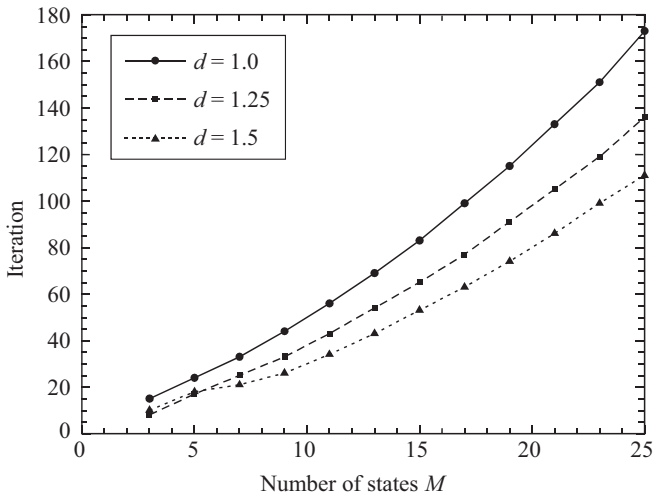

FIG. 3. Example of numerical calculations for minimax solutions for an M-ary optical ASK coherent state set. Number of iterations to achieve difference between upper and lower bounds for minimax value, $\overline{P_{\mathrm{C}}^{\star}}^{(w^{(r)})} - \underline{P_{\mathrm{C}}^{\star}}^{(w^{(r)})}$, less than 10^{-9} is shown.

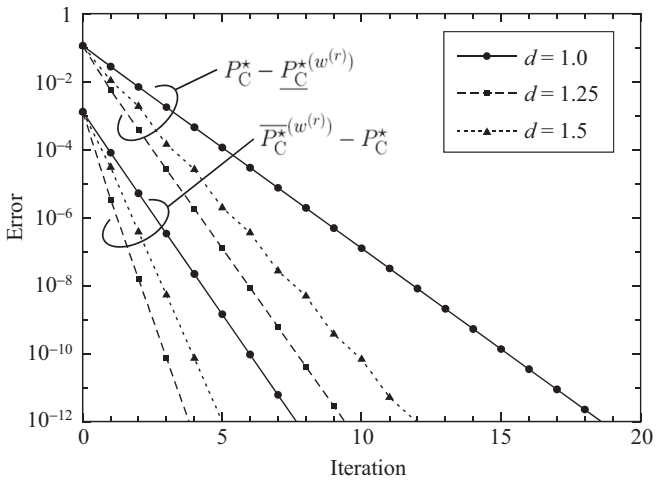

FIG. 4. Accuracy of minimax value vs number of iterations for a ternary ASK coherent state set.

The convergence properties of the iterative algorithm for a minimax solution, expressed by Eq. (47), for ρ are shown in Fig. 3. From Proposition 3, the QWM $\Pi^{([\xi^{\star}]^2)}$ with minimax probabilities $\xi^{\star}$ is a minimax measurement for a pure state set. In consideration of this, we set the initial weights as $w_m^{(0)} = 1/M^2$ for each $m \in \mathcal{I}_M$. In a similar manner to Fig. 2, we plot the number of iterations r to achieve $\overline{P_{\mathrm{C}}^{\star}}^{(w^{(r)})} - \underline{P_{\mathrm{C}}^{\star}}^{(w^{(r)})}$ less than 10^{-9} for different M. We can also see that the number of iterations required almost decreases as d increases in the range of $1.0 \leqslant d \leqslant 1.5$. Note that the number of iterations required for a minimax solution is less than that for a minimum-error measurement. However, we found that this is generally not the case; in many state sets, the algorithm for a minimum-error measurement has a faster rate of convergence than that for a minimax solution.

The values of $\overline{P_{\mathrm{C}}^{\star}}^{(w^{(r)})} - P_{\mathrm{C}}^{\star}$ and $P_{\mathrm{C}}^{\star} - \underline{P_{\mathrm{C}}^{\star}}^{(w^{(r)})}$ for 20 iterations when $M = 3$ are illustrated in Fig. 4. For $M = 3$, a closed-form analytical expression for the minimax value $P_{\mathrm{C}}^{\star}$ is known [28]. One can see that both $\overline{P_{\mathrm{C}}^{\star}}^{(w^{(r)})} - P_{\mathrm{C}}^{\star}$ and $P_{\mathrm{C}}^{\star} - \underline{P_{\mathrm{C}}^{\star}}^{(w^{(r)})}$ nearly exponentially decreases as the number of iterations increases. Note that in this case our algorithm with $d = 1.25$ has a faster rate of convergence than that with $d = 1.5$.

In Figs. 5 and 6, we show the dependence of the convergence properties of the proposed algorithms for a minimum-error measurement and a minimax solution, respectively, on the rank of density operators. We use 100 sets of randomly generated four quantum states $\{\rho_m : m \in \mathcal{I}_4\}$ whose supports are linearly independent. Prior probabilities are also randomly selected in the case of finding a minimum-error measurement. For any $m \in \mathcal{I}_4, \rho_m$ has a rank of R, i.e., $R_m = \mathrm{rank}\rho_m = R$, and thus, the dimension of the state space is $N = 4R$. We plot the average number of iterations r to achieve the difference between the upper and lower bounds for optimal value less than 10^{-9} for different R. In Figs. 5 and 6, the numbers of iterations required tend to increase and decrease as R increases, respectively; however, it is not significantly changed by R at least in the range of $R \leqslant 15$. Thus, since the proposed algorithms take $O(N^3) = O(R^3)$ time for a single iteration as explained in the next section, they require about R^3 times higher computational complexity than in the case of $R = 1$.

In Figs. 2–6, the proposed algorithms with $d = 1.25$ always take less average number of iterations than those with $d = 1.0$, which implies that $d \approx 1.25$ is expected to offer stable performance. Note that, using the property that the logarithm of the difference between the upper and lower bounds decreases approximately linearly with the number of iterations, as we can see in Fig. 4, we can obtain a rough estimate of the number of iterations required after a few iterations, and thus can choose an appropriate value of d.

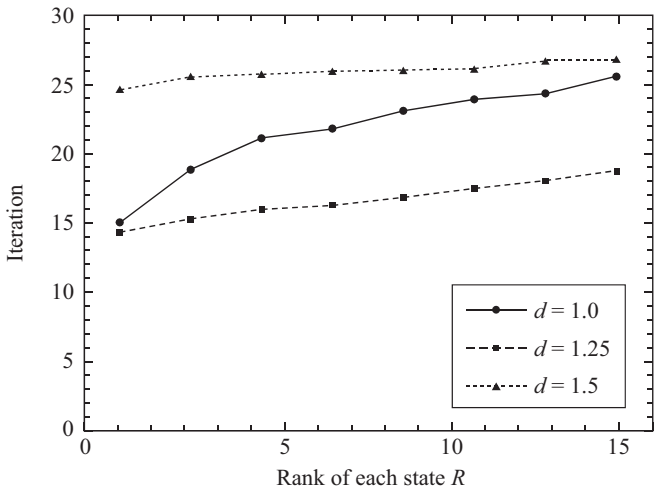

FIG. 5. Example of numerical calculations for minimum-error measurements for four pure ($R = 1$) and mixed ($R > 1$) states. Number of iterations to achieve difference between upper and lower bounds for maximum correct probability less than 10^{-9} is shown.

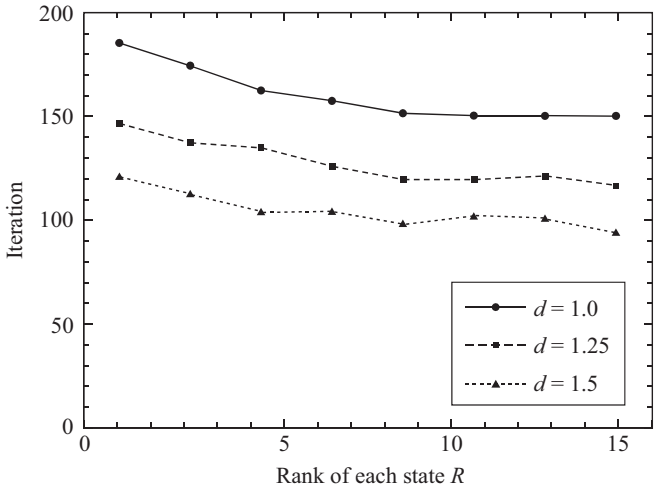

FIG. 6. Example of numerical calculations for minimax solutions for four pure ($R = 1$) and mixed ($R > 1$) states. Number of iterations to achieve difference between upper and lower bounds for minimax value less than 10^{-9} is shown.

VII. TIME AND SPACE COMPLEXITY

The proposed algorithms are expected to have lower computational complexity and memory size than classical approaches for solving an SDP. Here we will compare our algorithms with CSDP [45], which is a widely used SDP solver implementing a primal-dual interior point method. We assume that the spans of $\{\rho_m : m \in \mathcal{I}_M\}$ are linearly independent (i.e., $\sum_{m=0}^{M-1} R_m = N$), and N^2 is much larger than M, which is satisfied in many practical cases. We do not assume that ρ is symmetric.

In minimal measuring strategy, CSDP solves the problem of maximizing $P_{\mathrm{C}}(\xi,\Pi)$ under the constraint of Eq. (1). Since the number of scalar equality constraints is N^2, CSDP takes $O(N^6)$ time for a single iteration and requires $O(N^4)$ storage [45]. In contrast, in our algorithm the calculation of $[G^{(r)}]^{-1/2}$ of Eq. (9), which requires $O(N^3)$ time for a single iteration, is often much harder than the other operations. Our algorithm requires $O(N^2)$ storage for N-dimensional square matrices.

In minimax strategy, the problem can be expressed as the following SDP [16]:

$$\begin{aligned} &\underset{\Pi\in\mathcal{M}}{\text{argmax}} \ \ \mathrm{Tr}(\rho_0\Pi_0), \\ &\text{subjectto} \ \ \mathrm{Tr}(\rho_m\Pi_m) = \mathrm{Tr}(\rho_0\Pi_0), \forall m \in \mathcal{I}_M. \end{aligned} \tag{73}$$

Since Π must satisfy Eq. (1) (i.e., $\Pi \in \mathcal{M}$), the number of scalar equality constraints is also N^2. This means that CSDP has the same order of time and memory complexity as that in the minimal measuring strategy. Similarly, our algorithm takes $O(N^3)$ time for a single iteration since the major computational cost is to compute $[G^{(r)}]^{-1/2}$. Our algorithm also requires $O(N^2)$ storage. Table I summarizes the above arguments.

We should mention the total complexity. We here consider that the permissible error of the correct probability is 10^{-9}. We can see from numerical observations that the number of iterations required by CSDP is typically less than 30. In contrast, our algorithm sometimes needs thousands of iterations as shown in Fig. 2. However, since the time complexity required by our algorithm and CSDP for a single iteration are respectively $O(N^3)$ and $O(N^6)$, it is expected that the total time complexity of our algorithm is much less than that of CSDP. We did numerical experiments on an Intel Core i7-4600M machine with 4 GB memory to compare the processing time. Both CSDP and our algorithm are implemented in C language. In the case of ASK coherent state set with $|\alpha|^2 = 0.1$ (this is the same condition as in Figs. 2 and 3) and $M = 15$, the processing time for CSDP and our algorithm in the minimal measuring strategy was 50 and 1 s, and that in the minimax strategy was 80 and 0.03 s, respectively.

TABLE I. Time and memory complexity for finding an optimal POVM.

Strategy	Method	Time complexity per iteration	Storage requirements
Minimum	CSDP	$O(N^6)$	$O(N^4)$
measuring	Proposed	$O(N^3)$	$O(N^2)$
Minimax	CSDP	$O(N^6)$	$O(N^4)$
	Proposed	$O(N^3)$	$O(N^2)$

We now compare our algorithm in the minimal measuring strategy with that of Ježek *et al.*. First, we consider the case of $d = 1$. If we naively implement Ježek *et al.*'s algorithm, then it involves several multiplications of N-dimensional square matrices, which requires additional $O(N^3)$ time for a single iteration and $O(N^2)$ storage in practice. Thus, it requires much computation time and memory, although it has the same order of computational complexity as our algorithm. However, if we optimize the implementation of Ježek *et al.*'s algorithm by representing ρ_m and Π_m as $\rho_m = \psi_m\psi_m^\dagger$ and $\Pi_m = \pi_m\pi_m^\dagger$, respectively ($\psi_m$ and π_m are $N \times R_m$ matrices for each $m \in \mathcal{I}_M$), then it can have the almost same time and memory requirements as our method. Next, we consider the case of $d > 1$. As described in Sec. IV A, the acceleration technique of our method is difficult to apply to Ježek *et al.*'s algorithm. In contrast, our algorithm can accelerate with only a small additional computational cost. Note that the optimal acceleration parameter d is different for each state set. The issue of how to optimize d is remained for future study.

It is worth discussing the parallel computing. SDP with matrix size of $N \times N$ and L constraints can be solved in polylog(N)polylog(L) time (for a given permissible error), using poly(N)poly(L) processors [46], where poly(a) and polylog(b) respectively denote the polynomial of a and polylogarithm of b. This is because operations for Hermitian matrices such as matrix multiplication and eigendecomposition take polylog time using polynomial number of processors. For the same reason, our algorithm can obtain a solution in polylog(N)polylog(L) time using poly(N)poly(L) processors. Even if parallel processors are used, our method usually must achieve sufficient accuracy with lower time complexity than SDP-based algorithms since our method has a lower total computational complexity in many cases.

If ρ is symmetric, then using the symmetry we can often reduce the time complexity of both SDP-based and our algorithms. For example, we consider a cyclic state set ρ,

which can be expressed as $\rho = \{\rho_m = U^m \rho_0 U^{-m} : m \in \mathcal{I}_M\}$ with a unitary matrix U. Since there exists a matrix $X \geqslant 0$ satisfying Eqs. (18) and (19) such that X commutes with U, we can simplify the SDP expression [43]. Similarly, $G^{(r)}$ commutes with U, which indicates that we can reduce the computational cost of computing $[G^{(r)}]^{-1/2}$ in our algorithm.

VIII. CONCLUSION

We investigated the problem of computing a minimum-error measurement and a minimax solution by optimizing the weights of the BWSRM. We derived necessary and sufficient conditions for the BWSRM to be a minimax measurement, which indicates that the trace of each optimal weight of the BWSRM is proportional to the square of the corresponding minimax probability. We also showed an extension of the iterative algorithm developed by Ježek *et al.* [24] for a minimum-error measurement. For a linearly independent pure state set, the algorithm of Ježek *et al.* converges to an optimal measurement. Then, we proposed an iterative algorithm for a minimax solution. For a pure state set, our algorithm monotonically decreases an upper bound of minimax value.

ACKNOWLEDGMENTS

We are grateful to O. Hirota of Tamagawa University for support. We would also like to thank the anonymous reviewers for their helpful comments. T.S.U. was supported (in part) by JSPS KAKENHI (Grant No. 24360151).

APPENDIX A: REMARKS ON MATRICES

We give some remarks for providing our main results.

Remark 10. For any positive semidefinite matrices A and B and any positive real number d, $A^d B A^d = B$ and $AB = B$ are identical. In particular, if B is strictly positive, $A^d B A^d = B$ if and only if A is the identity matrix.

Proof. First, assume that $A^d B A^d = B$. Let a Schatten decomposition of A be $A = \sum_k a_k |a_k\rangle \langle a_k|$. Premultiplying and postmultiplying both sides of $A^d B A^d = B$ by $\langle a_k|$ and $|a_k\rangle$, respectively, yield

$$a_k^{2d} \langle a_k| B |a_k\rangle = \langle a_k| B |a_k\rangle . \tag{A1}$$

Let $\mathcal{K} = \{k : \langle a_k| B |a_k\rangle > 0\}$. From Eq. (A1), $a_k = 1$ holds for any $k \in \mathcal{K}$. Therefore, $AB = B$ holds.

Next, assume that $AB = B$. Each column of B is an eigenvector of A with eigenvalue 1, and also an eigenvector of A^d with eigenvalue 1. Therefore, $A^d B A^d = B A^d = (A^d B)^\dagger = B$ holds.

If B is strictly positive, i.e., B^{-1} exists, then it is obvious that $AB = B$ if and only if A is the identity matrix. ■

Remark 11. Let I'_t be the t-dimensional identity matrix. For any $n \times k$ matrix Φ, any n-dimensional strictly positive semidefinite matrix B, and any real number c, $c\Phi^\dagger B \Phi \leqslant I'_k$ and $B^{-1} \geqslant c\Phi\Phi^\dagger$ are identical.

Proof. If $c \leqslant 0$, then this remark is obvious, since in this case $I'_k - c\Phi^\dagger B \Phi$ and $B^{-1} - c\Phi\Phi^\dagger$ are always positive semidefinite. Assume $c > 0$. Let $A = c^{1/2}\Phi^\dagger B^{1/2}$. Then, $c\Phi^\dagger B \Phi \leqslant I'_k$ is identical to $AA^\dagger \leqslant I'_k$. In contrast, $B^{-1} \geqslant c\Phi\Phi^\dagger$ is identical to

$$I'_n \geqslant cB^{1/2}\Phi\Phi^\dagger B^{1/2} = A^\dagger A. \tag{A2}$$

Now we will show that $AA^\dagger \leqslant I'_k$ and $A^\dagger A \leqslant I'_n$ are identical. $AA^\dagger \leqslant I'_k$ holds if and only if all eigenvalues of $AA^\dagger$ are less than or equal to 1. Since $AA^\dagger$ and $A^\dagger A$ have the same nonzero eigenvalues [47], this means that all eigenvalues of $A^\dagger A$ are less than or equal to 1, that is, $A^\dagger A \leqslant I'_n$. ■

APPENDIX B: REMARKS ON BWSRMs

Remark 12. We consider a state set $\rho = \{\rho_m = \psi_m \psi_m^\dagger : m \in \mathcal{I}_M\}$ with prior probabilities $\xi = \{\xi_m : m \in \mathcal{I}_M\}$. Let $w \in \mathcal{W}$.

If w satisfies Eqs. (14) and (15) with a certain c (i.e., the BWSRM $\Pi^{(w)}$ is a minimum-error measurement from Remark 1), then we have

$$c[G^{(w)}]^{1/2} = \sum_{k=0}^{M-1} \xi_k \rho_k \Pi_k^{(w)}. \tag{B1}$$

Proof. We have

$$\begin{aligned} c[G^{(w)}]^{1/2} &= cG^{(w)}[G^{(w)}]^{-1/2} \\ &= c\sum_{k=0}^{M-1} \psi_k w_k \psi_k^\dagger [G^{(w)}]^{-1/2} \\ &= c\sum_{k=0}^{M-1} \xi_k \psi_k Y_k^{(w)} w_k \psi_k^\dagger [G^{(w)}]^{-1/2} \\ &= \sum_{k=0}^{M-1} \xi_k \psi_k \psi_k^\dagger [G^{(w)}]^{-1/2} \psi_k w_k \psi_k^\dagger [G^{(w)}]^{-1/2} \\ &= \sum_{k=0}^{M-1} \xi_k \rho_k \Pi_k^{(w)}, \end{aligned} \tag{B2}$$

where the second to fifth lines follow from Eqs. (6), (15), (9), and (5), respectively. ■

Remark 13 (Theorem 5 of Ref. [35]). We consider a pure state set $\rho = \{|\psi_m\rangle \langle\psi_m| : m \in \mathcal{I}_M\}$. Let $\Xi(w) = \{\Xi_m(w) : m \in \mathcal{I}_M\}$, where $w \in \mathcal{W}$ and

$$\Xi_m(w) = \frac{1}{Y_m^{(w)}} \left(\sum_{k=0}^{M-1} \frac{1}{Y_k^{(w)}} \right)^{-1}. \tag{B3}$$

Note that $\Xi_m(w) > 0$ and $\sum_{m=0}^{M-1} \Xi_m(w) = 1$ hold, and thus $\Xi(w)$ can be interpreted as prior probabilities. For any $w \in \mathcal{W}$, the BWSRM $\Pi^{(w)}$ is a minimum-error measurement for ρ with prior probabilities $\Xi(w)$.

From Remark 13, we can obtain the prior probabilities ξ such that $\Pi^{(w)}$ is a minimum-error measurement for a given w, that is, $\xi = \Xi(w)$. However, it is often difficult to obtain w such that $\Pi^{(w)}$ is a minimum-error measurement for a given ξ.

Note that Remark 13 can be generalized to a mixed state set as follows.

Remark 14. We consider a state set $\rho = \{\rho_m = \psi_m \psi_m^\dagger : m \in \mathcal{I}_M\}$ with prior probabilities $\xi = \{\xi_m : m \in \mathcal{I}_M\}$. Let $R_m = \mathrm{rank}\rho_m$. We also consider a BWSRM $\Pi^{(w)} = \{\Pi_m^{(w)} : m \in \mathcal{I}_M\}$, where $w = \{w_m \in \mathcal{S}^{R_m} : m \in \mathcal{I}_M\} \in \mathcal{W}$. Assume

$w_m > 0$ for any $m \in \mathcal{I}_M$. If $\Pi^{(w)}$ is a minimum-error measurement for ρ with ξ, then we have that for any $m \in \mathcal{I}_M$,

$$\xi_m = \frac{R_m}{\mathrm{Tr}Y_m^{(w)}} \left(\sum_{k=0}^{M-1} \frac{R_k}{\mathrm{Tr}Y_k^{(w)}} \right)^{-1}. \tag{B4}$$

Proof. Assume that $\Pi^{(w)}$ is a minimum-error measurement for ρ with ξ. Postmultiplying both sides of Eq. (15) by w_m^{-1} yields $\xi_m Y_m^{(w)} = cI_m$. Taking the trace on both sides we have $\xi_m \mathrm{Tr}Y_m^{(w)} = cR_m$. Therefore, $\xi_m \propto R_m/\mathrm{Tr}Y_m^{(w)}$, i.e., Eq. (B4) holds. ■

APPENDIX C: PROOF OF THEOREM 5

1. Outline

First, we derive that $\xi_m Y_m^{(r)}$ converges to 1 as r tends to infinity. Next, using this property, we prove that the correct probability of the BWSRM $\Pi^{(w^{(r)})}$, denoted as $P_{\mathrm{C}}^{(r)}$, converges to the maximum correct probability $P_{\mathrm{C}}^{\mathrm{opt}}$. Finally, we show that $\Pi^{(w^{(r)})}$ converges to a minimum-error measurement, denoted as $\Pi^{\bullet}$.

2. Proof that $\xi_m Y_m^{(r)}$ converges to 1

Let $E_m^{(r)}$ be an N-dimensional square matrix satisfying $\Pi_m^{(w^{(r)})} = [E_m^{(r)}]^\dagger E_m^{(r)}$. Also let

$$S^{(r)} = \sum_{m=0}^{M-1} \mathrm{Tr}\left[E_m^{(r+1)} - E_m^{(r)}\right]^\dagger \left[E_m^{(r+1)} - E_m^{(r)}\right] \xi_m \rho_m. \tag{C1}$$

We exploit the fact that $S^{(r)}$ converges to 0 as r tends to infinity, which is proved by Tyson [see Eq. (16) of Ref. [25]]. $E_m^{(r)}$ can be expressed as

$$E_m^{(r)} = |0\rangle \langle \psi_m| \sqrt{w_m^{(r)}} [G^{(r)}]^{-1/2}, \tag{C2}$$

where $|0\rangle$ is any normal vector in $\mathcal{H}$. From Eq. (9), we have

$$E_m^{(r)} |\psi_m\rangle = |0\rangle \sqrt{w_m^{(r)}} Y_m^{(r)}. \tag{C3}$$

We obtain, for any $m \in \mathcal{I}_M$,

$$\begin{aligned} S^{(r)} &\geqslant \mathrm{Tr}\left[E_m^{(r+1)} - E_m^{(r)}\right]^\dagger \left[E_m^{(r+1)} - E_m^{(r)}\right] \xi_m \rho_m \\ &= \xi_m \left| \sqrt{w_m^{(r+1)}} Y_m^{(r+1)} - \sqrt{w_m^{(r)}} Y_m^{(r)} \right|^2 \\ &= \xi_m \left| \sqrt{w_m^{(r)}} Y_m^{(r)} [\xi_m Y_m^{(r+1)} - 1] \right|^2, \end{aligned} \tag{C4}$$

where the first to third lines follow from Eqs. (C1), (C3), and (29), respectively. Since $S^{(r)}$ converges to 0, the right-hand side of last line of Eq. (C4) also converges to 0. Therefore, to prove that $\xi_m Y_m^{(r)}$ converges to 1, it suffices to show that $\sqrt{w_m^{(r)}} Y_m^{(r)}$ has a positive lower bound.

Let $|\psi_m^\perp\rangle \in \mathcal{H}\,(m \in \mathcal{I}_M)$ be a unit vector such that $\langle \psi_m^\perp | \psi_k \rangle = 0$ for any $k \in \mathcal{I}_M$ with $k \neq m$. Then, we obtain a lower bound of $\sqrt{w_m^{(r)}} Y_m^{(r)}$ as follows:

$$\begin{aligned} \sqrt{w_m^{(r)}} Y_m^{(r)} &= \sqrt{w_m^{(r)}} \langle \psi_m | [G^{(r)}]^{-1/2} | \psi_m \rangle \\ &\geqslant \frac{\sqrt{w_m^{(r)}} |\langle \psi_m^\perp | \psi_m \rangle|^2}{\langle \psi_m^\perp | [G^{(r)}]^{1/2} | \psi_m^\perp \rangle} \\ &\geqslant \frac{\sqrt{w_m^{(r)}} |\langle \psi_m^\perp | \psi_m \rangle|^2}{\sqrt{\langle \psi_m^\perp | G^{(r)} | \psi_m^\perp \rangle}} \\ &= \frac{\sqrt{w_m^{(r)}} |\langle \psi_m^\perp | \psi_m \rangle|^2}{\sqrt{\langle \psi_m^\perp | \psi_m \rangle w_m^{(r)} \langle \psi_m | \psi_m^\perp \rangle}} \\ &= |\langle \psi_m^\perp | \psi_m \rangle|, \end{aligned} \tag{C5}$$

where the second line follows from Remark 11 (with $B = [G^{(r)}]^{1/2}, c = \langle \psi_m^\perp | [G^{(r)}]^{1/2} | \psi_m^\perp \rangle^{-1}$, and $\Phi = |\psi_m^\perp\rangle$). The third line follows from

$$\langle \psi_m^\perp | [G^{(r)}]^{1/2} | \psi_m^\perp \rangle^2 \leqslant \langle \psi_m^\perp | G^{(r)} | \psi_m^\perp \rangle, \tag{C6}$$

which is derived from the fact that $A^\dagger B A \leqslant A^\dagger A$ holds for any matrices A and B satisfying $B \leqslant I$ (in this case, $A = [G^{(r)}]^{1/2} |\psi_m^\perp\rangle$ and $B = |\psi_m^\perp\rangle \langle \psi_m^\perp|$). The fourth line follows from Eq. (6). Since $\{|\psi_m\rangle : m \in \mathcal{I}_M\}$ is linearly independent, $|\langle \psi_m^\perp | \psi_m \rangle|$ is positive.

3. Proof that $P_{\mathrm{C}}^{(r)}$ converges to $P_{\mathrm{C}}^{\mathrm{opt}}$

Since $\xi_m Y_m^{(r)}$ converges to 1, for any $\epsilon' > 0, r_0 > 0$ exists such that $\xi_m Y_m^{(r+1)} \leqslant 1 + \epsilon'$ for any $r > r_0$. Let $\epsilon = 1 - (1 + \epsilon')^{-1}$ (note that ϵ converges to 0 if ϵ' converges to 0). Then, we have

$$(1 - \epsilon) \xi_m Y_m^{(r+1)} \leqslant 1. \tag{C7}$$

Substituting Eq. (9) into Eq. (C7) and using Remark 11 [with $c = (1 - \epsilon)\xi_m, \Phi = |\psi_m\rangle$, and $B = [G^{(r+1)}]^{-1/2}$] yield

$$[G^{(r+1)}]^{1/2} \geqslant (1 - \epsilon) \xi_m |\psi_m\rangle \langle \psi_m|. \tag{C8}$$

Hence, we have

$$\begin{aligned} &\mathrm{Tr}[G^{(r+1)}]^{1/2} - (1 - \epsilon) P_{\mathrm{C}}^{\mathrm{opt}} \\ &= \mathrm{Tr} \sum_{m=0}^{M-1} \{[G^{(r+1)}]^{1/2} - (1 - \epsilon) \xi_m |\psi_m\rangle \langle \psi_m|\} \Pi_m^{\bullet} \\ &\geqslant 0, \end{aligned} \tag{C9}$$

where the last inequality follows from the fact that $\mathrm{Tr}AB \geqslant 0$ for any matrices $A \geqslant 0$ and $B \geqslant 0$. In contrast, we can obtain

$$\begin{aligned} P_{\mathrm{C}}^{(r)} &= \sum_{m=0}^{M-1} \xi_m \langle \psi_m | \Pi_m^{(w^{(r)})} | \psi_m \rangle \\ &= \sum_{m=0}^{M-1} \xi_m^{-1} w_m^{(r+1)} \\ &\geqslant (1 - \epsilon) \sum_{m=0}^{M-1} Y_m^{(r+1)} w_m^{(r+1)} \end{aligned}$$

$$= (1-\epsilon)\mathrm{Tr}\sum_{m=0}^{M-1}[G^{(r+1)}]^{-1/2} w_m^{(r+1)} |\psi_m\rangle\langle\psi_m|$$
$$= (1-\epsilon)\mathrm{Tr}[G^{(r+1)}]^{1/2}. \quad \text{(C10)}$$

The second line follows from

$$w_m^{(r+1)} = \xi_m^2 w_m^{(r)} \left[Y_m^{(r)}\right]^2 = \xi_m^2 \langle\psi_m| \Pi_m^{(w^{(r)})} |\psi_m\rangle , \quad \text{(C11)}$$

which is given by Eqs. (8) and (29). The third and fourth lines of Eq. (C10) follow from Eqs. (C7) and (9), respectively. From Eqs. (C9) and (C10), we obtain

$$P_{\mathrm{C}}^{(r)} \geqslant (1-\epsilon)^2 P_{\mathrm{C}}^{\mathrm{opt}}. \quad \text{(C12)}$$

Therefore, for any $\epsilon > 0$, $r_0 > 0$ exists such that $P_{\mathrm{C}}^{(r)} \geqslant (1-\epsilon)^2 P_{\mathrm{C}}^{\mathrm{opt}}$ for any $r > r_0$, which implies that $P_{\mathrm{C}}^{(r)}$ converges to $P_{\mathrm{C}}^{\mathrm{opt}}$ as r tends to infinity.

4. Proof that $\Pi^{(w^{(r)})}$ converges to $\Pi^{\bullet}$

The entire set of POVMs $\mathcal{M}$ can be regarded as a closed convex set, and $P_{\mathrm{C}}(\xi,\Pi)$ is a concave function with respect to Π. $P_{\mathrm{C}}^{\mathrm{opt}}$ is the maximum value of $P_{\mathrm{C}}(\xi,\Pi)$ over all $\Pi \in \mathcal{M}$. In contrast, the minimum-error measurement $\Pi^{\bullet}$ is uniquely determined for a linearly independent pure state set [48]. Thus, since $P_{\mathrm{C}}[\xi,\Pi^{(w^{(r)})}]$ converges to $P_{\mathrm{C}}^{\mathrm{opt}}$, it is obvious that $\Pi^{(w^{(r)})}$ converges to $\Pi^{\bullet}$. ■

[1] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).
[2] A. S. Holevo, J. Multivariate Anal. **3**, 337 (1973).
[3] H. P. Yuen, R. S. Kennedy, and M. Lax, IEEE Trans. Inf. Theory **21**, 125 (1975).
[4] V. P. Belavkin, Radio Eng. Electron. Phys. **20**, 39 (1975).
[5] M. Ban, K. Kurokawa, R. Momose, and O. Hirota, Int. J. Theor. Phys. **36**, 1269 (1997).
[6] T. S. Usuda, I. Takumi, M. Hata, and O. Hirota, Phys. Lett. A **256**, 104 (1999).
[7] S. M. Barnett, Phys. Rev. A **64**, 030303 (2001).
[8] Y. C. Eldar and G. D. Forney Jr., IEEE Trans. Inf. Theory **47**, 858 (2001).
[9] E. Andersson, S. M. Barnett, C. R. Gilson, and K. Hunter, Phys. Rev. A **65**, 052308 (2002).
[10] K. Kato and O. Hirota, IEEE Trans. Inf. Theory **49**, 3312 (2003).
[11] C. L. Chou and L. Y. Hsu, Phys. Rev. A **68**, 042305 (2003).
[12] O. Hirota and S. Ikehara, Trans. IECE Jpn. **E65**, 627 (1982).
[13] M. Osaki, M. Ban, and O. Hirota, Phys. Rev. A **54**, 1691 (1996).
[14] G. M. D'Ariano, M. F. Sacchi, and J. Kahn, Phys. Rev. A **72**, 032310 (2005).
[15] K. Kato, in *Proceedings of the IEEE International Symposium on Information Theory, Cambridge, New York* (IEEE, New York, 2012), pp. 1077–1081.
[16] K. Nakahira, K. Kato, and T. S. Usuda, Phys. Rev. A **88**, 032314 (2013).
[17] A. S. Holevo, Prob. Inf. Transmission **10**, 317 (1974).
[18] Y. C. Eldar, A. Megretski, and G. C. Verghese, IEEE Trans. Inf. Theory **49**, 1007 (2003).
[19] M. Zibulevsky and M. Elad, IEEE Signal Process. Mag. **27**, 76 (2010).
[20] I. Daubechies, M. Defrise, and C. De-Mol, Commun. Pure Appl. Math. **57**, 1413 (2004).
[21] Y. Boykov and V. Kolmogorov, IEEE Trans. Patt. Anal. Mach. Intell. **26**, 1124 (2004).
[22] T. Blumensath and M. E. Davies, J. Fourier Anal. Applicat. **14**, 629 (2008).
[23] A. Beck and M. Teboulle, SIAM J. Imaging Sci. **2**, 183 (2009).
[24] M. Ježek, J. Řeháček, and J. Fiurášek, Phys. Rev. A **65**, 060301 (2002).
[25] J. Tyson, J. Math. Phys. **51**, 092204 (2010).
[26] M. Reimpell and R. F. Werner, Phys. Rev. Lett. **94**, 080501 (2005).
[27] M. Reimpell, *Quantum Information and Convex Optimization*. Ph.D. thesis, Technische Universität, 2007.
[28] K. Kato, J. Phys.: Conf. Ser. **414**, 012039 (2012).
[29] P. Hausladen and W. K. Wootters, J. Mod. Opt. **41**, 2385 (1994).
[30] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W. K. Wootters, Phys. Rev. A **54**, 1869 (1996).
[31] M. Sasaki, K. Kato, M. Izutsu, and O. Hirota, Phys. Rev. A **58**, 146 (1998).
[32] M. Sasaki, T. Sasaki-Usuda, M. Izutsu, and O. Hirota, Phys. Rev. A **58**, 159 (1998).
[33] K. Kato, M. Osaki, M. Sasaki, and O. Hirota, IEEE Trans. Commun. **47**, 248 (1999).
[34] V. P. Belavkin and V. Maslov, in *Mathematical Aspects of Computer Engineering*, edited by V. Maslov (MIR, Moscow, 1987), pp. 146–237.
[35] C. Mochon, Phys. Rev. A **73**, 032328 (2006).
[36] J. Tyson, Phys. Rev. A **79**, 032343 (2009).
[37] J. Tyson, J. Math. Phys. **50**, 032106 (2009).
[38] I. Ekeland and R. Téman, *Convex Analysis and Variational Problems* (SIAM, North Holland, 1999).
[39] A. S. Holevo, Theory Probab. Appl. **23**, 411 (1979).
[40] Y. C. Eldar, A. Megretski, and G. C. Verghese, IEEE Trans. Inf. Theory **50**, 1198 (2004).
[41] Theorem 2 of Ref. [12] holds if all of the minimax probabilities are nonzero (see [15]). Note that necessary and sufficient conditions for any minimax probabilities are given in Refs. [15,16].
[42] K. Nakahira and T. S. Usuda, Phys. Rev. A **87**, 012308 (2013).
[43] A. Assalini, G. Cariolaro, and G. Pierobon, Phys. Rev. A **81**, 012315 (2010).
[44] Although Ref. [43] only considers the case of a cyclic group of unitary matrices, this approach can be used in more general cases.
[45] B. Borchers, Optim. Methods Softw. **11**, 613 (1999).
[46] R. Jain and P. Yao in *Proceedings of the IEEE Symposium on Foundations of Computer Science, New York* (IEEE Computer Society, Washington, DC, USA, 2011), pp. 463–471.
[47] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, New York, 1985).
[48] A. S. Holevo, Probab. Math. Stat. **3**, 113 (1982).