Mixed-quantum-state detection with inconclusive results

Yonina C. Eldar*
Technion-Israel Institute of Technology, Technion City, Haifa 32000, Israel
(Received 19 November 2002; revised manuscript received 27 January 2003; published 15 April 2003)

We consider the problem of designing an optimal quantum detector with a fixed rate of inconclusive results that maximizes the probability of correct detection, when distinguishing between a collection of mixed quantum states. We show that the design of the optimal detector can be formulated as a semidefinite programming problem, and derive a set of necessary and sufficient conditions for an optimal measurement. We then develop a sufficient condition for the *scaled inverse measurement* to maximize the probability of correct detection for the case in which the rate of inconclusive results exceeds a certain threshold. Using this condition we derive the optimal measurement for linearly independent pure-state sets, and for mixed-state sets with a broad class of symmetries. Specifically, we consider geometrically uniform (GU) state sets and compound geometrically uniform (CGU) state sets with generators that satisfy a certain constraint. We then show that the optimal measurements corresponding to GU and CGU state sets with arbitrary generators are also GU and CGU, respectively, with generators that can be computed very efficiently in polynomial time within any desired accuracy by solving a reduced size semidefinite programming problem.

DOI: 10.1103/PhysRevA.67.042309 PACS number(s): 03.67.Hk

I. INTRODUCTION

Quantum-information theory refers to the distinctive information processing properties of quantum systems, which arise when information is stored in or retrieved from quantum states. A fundamental aspect of quantum information theory is that nonorthogonal quantum states cannot be perfectly distinguished. Therefore, a central problem in quantum mechanics is to design measurements optimized to distinguish between a collection of nonorthogonal quantum states.

We consider a quantum-state ensemble consisting of m positive semidefinite Hermitian density operators $\{\rho_i,\ 1\leqslant i\leqslant m\}$ on an n-dimensional complex Hilbert-space $\mathcal{H}$, with prior probabilities $\{p_i>0, 1\leqslant i\leqslant m\}$. For our *measurement*, we consider general positive operator-valued measures [1,2], consisting of positive semidefinite Hermitian operators that form a resolution of the identity on $\mathcal{H}$.

Different approaches to distinguish between the density operators ρ_i have emerged. In one approach, the measurement consists of m measurement operators which are designed to maximize the probability of correct detection. Necessary and sufficient conditions for an optimum measurement maximizing the probability of correct detection have been developed [3–5]. Closed-form analytical expressions for the optimal measurement have been derived for several special cases [6–11]. In particular, the optimal measurement for pure and mixed-state ensembles with broad symmetry properties, referred to as geometrically uniform (GU) and compound GU (CGU) state sets, are considered in Refs. [10,11]. Iterative procedures maximizing the probability of correct detection have also been developed for cases in which the optimal measurement cannot be found explicitly [12,5].

More recently, a different approach to the problem has emerged, which in some cases may be more useful. This approach, referred to as unambiguous quantum-state discrimination, combines error-free discrimination with a certain fraction of inconclusive results [13–20]. The basic idea, pioneered by Ivanovic [13], is to design a measurement that with probability β returns an inconclusive result, but if the measurement returns an answer, then the answer is correct with probability 1. In this case, the measurement consists of $m+1$ measurement operators corresponding to $m+1$ outcomes, where m outcomes correspond to detection of each of the states and the additional outcome corresponds to an inconclusive result. Chefles [18] showed that a necessary and sufficient condition for the existence of unambiguous measurements for distinguishing between a collection of pure quantum states is that the states are linearly independent pure states. Necessary and sufficient conditions on the optimal measurement minimizing the probability β of an inconclusive result were derived in Ref. [20]. The optimal measurement when distinguishing between GU and CGU pure-state sets was also considered in Ref. [20], and was shown under certain conditions to be equal to the equal-probability measurement (EPM).

An interesting alternative approach for distinguishing between a collection of quantum-states, first considered by Chefles and Barnett [21] and Zhang, Li, and Guo [22] for pure-state ensembles, and then later extended by Fiurášek and Ježek [23] to mixed-state ensembles, is to allow for a certain probability of an inconclusive result, and then maximize the probability of correct detection when a conclusive result is obtained. Thus, in this approach, the measurement again consists of $m+1$ measurement outcomes. However, now the outcomes do not necessarily correspond to perfect detection of each of the states. Indeed, if, for example, the quantum states are linearly dependent pure states, then perfect detection of each of the states is not possible [18]. Nonetheless, by allowing for inconclusive results, a higher probability of correct detection can be obtained for the case in which a conclusive result is obtained in comparison with the

*Electronic address: yonina@ee.technion.ac.il

probability of correct detection attainable without inconclusive results.

Necessary conditions as well as a set of sufficient conditions on the optimal measurement operators maximizing the probability of correct detection subject to the constraint that the probability of an inconclusive result is equal to a constant β were derived in Ref. [23], using Lagrange multiplier theory. It was also pointed out in Ref. [23] that obtaining a closed form analytical solution to the optimal measurement operators directly from these conditions is a difficult problem.

In this paper, we extend the results of Ref. [23] in several ways. First, using principles of duality in vector space optimization, in Sec. III we show that the conditions derived in Ref. [23] are both necessary and sufficient. We also show that the Lagrange multipliers can be obtained by solving a reduced size semidefinite programming problem [24], which is a convex optimization problem. By exploiting the many well-known algorithms for solving semidefinite programs [24–27], the optimal measurement can be computed very efficiently in polynomial time within any desired accuracy. Furthermore, in contrast to the iterative algorithm proposed in Ref. [23], which is only guaranteed to converge to a local optimum, algorithms based on semidefinite programming are guaranteed to converge to the global optimum.

Second, we derive a general condition in Sec. IV under which the *scaled inverse measurement* (SIM) is optimal. This measurement consists of measurement operators that are proportional to the reciprocal states associated with the given state ensemble, and can be regarded as a generalization of the EPM to linearly dependent pure-state sets and mixed states.

Third, we develop the optimal measurement for state sets with broad symmetry properties. Specifically, in Sec. V we consider GU state sets defined over a finite group of unitary matrices. We obtain a convenient characterization of the SIM and show that the SIM operators have the same symmetries as the original state set. We then show that for a pure GU state set and for values of β exceeding a certain threshold, the SIM is optimal. For a mixed GU state set, under a certain constraint on the generator and for values of β exceeding a threshold, the SIM is again shown to be optimal. For arbitrary values of β, the optimal measurement operators corresponding to a pure or mixed GU state set are shown to be GU with the same generating group. As we show, the generators can be computed very efficiently in polynomial time within any desired accuracy.

In Sec. VI we consider CGU state sets [28], in which the states are generated by a group of unitary matrices using *multiple* generators. We obtain a convenient characterization of the SIM for CGU state sets, and show that the SIM vectors are themselves CGU. Under a certain condition on the generators and for values of β exceeding a threshold, the SIM is shown to be optimal. Finally we show that for arbitrary CGU state sets and for arbitrary values of β, the optimal measurement operators are also CGU, and we propose an efficient algorithm for computing the optimal generators.

It is interesting to note that a closed form analytical expression exists for the optimal measurement when distinguishing between GU and CGU (possibly mixed) state sets with generators that satisfy a certain constraint, under each of the three approaches outlined to quantum detection, where in the last approach we assume that β exceeds a given threshold. Furthermore, as shown in Refs. [11,20] and in Secs. V and VI, the optimal measurement operators corresponding to GU and CGU state sets are also GU and CGU, respectively, under each one of the three outlined optimality criteria.

Before proceeding to the detailed development, we provide in the following section a statement of our problem.

II. PROBLEM FORMULATION

Assume that a quantum channel is prepared in a quantum-state drawn from a collection of given states represented by density operators $\{\rho_i, 1 \leqslant i \leqslant m\}$ on an n-dimensional complex Hilbert space $\mathcal{H}$. We assume without loss of generality that the eigenvectors of $\rho_i, 1 \leqslant i \leqslant m$, collectively span [43] $\mathcal{H}$ so that $m \geqslant n$. Since ρ_i is Hermitian and positive semidefinite, we can express ρ_i as $\rho_i = \phi_i \phi_i^*$ for some matrix ϕ_i, e.g., via the Cholesky or eigendecomposition of ρ_i [29]. We refer to ϕ_i as a *factor* of ρ_i. The choice of ϕ_i is not unique; if ϕ_i is a factor of ρ_i, then any matrix of the form $\phi_i' = \phi_i Q_i$, where Q_i is an arbitrary matrix satisfying $Q_i Q_i^* = I$, is also a factor of ρ_i.

To detect the state of the system a measurement is constructed comprising $m+1$ measurement operators $\{\Pi_i, 0 \leqslant i \leqslant m\}$ that satisfy

$$\Pi_i \geqslant 0, \quad 0 \leqslant i \leqslant m,$$

$$\sum_{i=0}^{m} \Pi_i = I. \tag{1}$$

Each of the operators $\Pi_i, 1 \leqslant i \leqslant m$ correspond to detection of the corresponding states $\rho_i, 1 \leqslant i \leqslant m$, and Π_0 corresponds to an inconclusive result. We seek the measurement operators Π_i that maximize the probability P_D of correct detection, subject to the constraint that the probability P_I of an inconclusive result is equal to a constant $\beta < 1$.

Given that the transmitted state is ρ_j, the probability of correctly detecting the state using measurement operators $\{\Pi_i, 1 \leqslant i \leqslant m\}$ is $\mathrm{Tr}(\rho_j \Pi_j)$ and the probability of a detection error is $\Sigma_{i=1, i \neq j}^{m} \mathrm{Tr}(\rho_j \Pi_i)$. Therefore, the probability of correct detection is given by

$$P_D = \sum_{i=1}^{m} p_i \mathrm{Tr}(\rho_i \Pi_i), \tag{2}$$

where $p_i > 0$ is the prior probability of ρ_i with $\Sigma_i p_i = 1$, and the probability of a detection error is given by

$$P_E = \sum_{i=1}^{m} \sum_{j=1, j \neq i}^{m} p_i \mathrm{Tr}(\rho_i \Pi_j). \tag{3}$$

The probability of an inconclusive result is

$$P_I=\sum_{i=1}^{m} p_i \mathrm{Tr}(\rho_i \Pi_0)=\mathrm{Tr}(\Delta \Pi_0)=\beta, \tag{4}$$

where for brevity we denote

$$\Delta=\sum_{i=1}^{m} p_i \rho_i . \tag{5}$$

Our problem is to find the measurement operators $\{\Pi_i, 0 \leqslant i \leqslant m\}$ that maximize P_D of Eq. (2) subject to the constraints (1) and (4).

Note that since $\mathrm{Tr}(\rho_i)=1$ for all i,

$$P_D+P_E+P_I=\sum_{i=1}^{m} p_i \mathrm{Tr}(\rho_i)=1. \tag{6}$$

Each of the three approaches to quantum detection outlined in the introduction correspond to maximizing P_D subject to different constraints on P_I and P_E. The standard quantum detection problem is to maximize P_D subject to $P_I=0$, which implies that $\Pi_0=0$. Therefore, this approach is equivalent to seeking m measurement operators $\Pi_i \geqslant 0, 1 \leqslant i \leqslant m$ satisfying $\Sigma_{i=1}^{m} \Pi_i = I$ to maximize P_D, or equivalently, minimize P_E. In unambiguous quantum-state discrimination the problem is to choose $m+1$ measurement operators to maximize P_D, or equivalently minimize P_I, subject to $P_E=0$. As shown in [18], it is not always possible to choose measurement operators such that $P_E=0$. For example, if the state ensemble is a pure-state ensemble consisting of density operators ρ_i of the form $\rho_i=|\phi_i\rangle\langle\phi_i|$ for a set of linearly dependent vectors $|\phi_i\rangle$, then there is no measurement that will result in $P_E=0$. Nonetheless, we may seek the measurement operators that minimize P_E, or equivalently, maximize P_D, subject to $P_I=\beta$ for some $\beta<1$. By allowing for $\beta>0$ we can achieve a larger probability of correct detection when a conclusive result is obtained than that which can be achieved using the standard quantum detection approach in which we require that $\beta=0$.

Equipped with the standard operations of addition and multiplication by real numbers, the space $\mathcal{B}$ of all Hermitian $n \times n$ matrices is an n^2-dimensional *real* vector space. As noted in Ref. [23], by choosing an appropriate basis for $\mathcal{B}$, the problem of maximizing P_D subject to Eqs. (1) and (4) can be put in the form of a standard semidefinite programming problem, which is a convex optimization problem; for a detailed treatment of semidefinite programming problems see, e.g., Refs. [24–27]. Recently, methods based on semidefinite programming have been employed in a variety of different problems in quantum detection and quantum information [5,20,30–35]. By exploiting the many well-known algorithms for solving semidefinite programs [24], e.g., interior point methods [44] [25,27], the optimal measurement can be computed very efficiently in polynomial time within any desired accuracy.

The semidefinite programming formulation can also be used to derive necessary and sufficient conditions for optimality, which we discuss in the following section.

III. CONDITIONS FOR OPTIMALITY

Using Lagrange multipliers, it was shown in Ref. [23] that a set of measurement operators $\{\hat{\Pi}_i, 0 \leqslant i \leqslant m\}$ maximizes P_D subject to $P_I=\beta$ for a state set $\{\rho_i, 1 \leqslant i \leqslant m\}$ with prior probabilities $\{p_i, 1 \leqslant i \leqslant m\}$ if there exists an Hermitian $\hat{X}$ and a constant $\hat{\delta}$ satisfying

$$\hat{X} \geqslant p_i \rho_i, \quad 1 \leqslant i \leqslant m, \tag{7}$$

$$\hat{X} \geqslant \hat{\delta} \Delta, \tag{8}$$

such that

$$(\hat{X}-p_i \rho_i)\hat{\Pi}_i=0, \quad 1 \leqslant i \leqslant m, \tag{9}$$

$$(\hat{X}-\hat{\delta}\Delta)\hat{\Pi}_0=0. \tag{10}$$

It was also shown that Eqs. (9) and (10) are necessary conditions for optimality.

In Appendix A we use duality arguments similar to those used in Ref. [5] to show that Eqs. (7)–(10) are *necessary and sufficient* conditions for optimality, so that a set of measurement operators $\hat{\Pi}_i$ maximizes P_D subject to $P_I=\beta$ if and only if there exists an Hermitian $\hat{X}$ and a constant $\hat{\delta}$ satisfying Eqs. (7)–(10). Furthermore, we show that $\hat{X}$ and $\hat{\delta}$ can be determined as the solution to the following semidefinite programming problem:

$$\min_{X \in \mathcal{B}, \delta \in \mathcal{R}} \{\mathrm{Tr}(X)-\delta\beta\}, \tag{11}$$

where $\mathcal{B}$ is the set of $n \times n$ Hermitian operators and $\mathcal{R}$ denotes the reals, subject to

$$X \geqslant p_i \rho_i, \quad 1 \leqslant i \leqslant m,$$

$$X \geqslant \delta \Delta. \tag{12}$$

The problem of Eqs. (11) and (12) is referred to as the *dual problem.*

Note that the dual problem involves many fewer decision variables than the primal maximization problem. Specifically, in the dual problem we have n^2+1 real decision variables while the primal problem has $(m+1)n^2$ real decision variables. Therefore, it is advantageous to solve the dual problem and then use Eqs. (9) and (10) to determine the optimal measurement operators, rather than solving the primal problem directly. Once we determine $\hat{X}$ and $\hat{\delta}$, the optimal measurement operators $\hat{\Pi}_i$ can be computed using Eqs. (9) and (10), in a similar manner to that described in Ref. [5].

We summarize the results of Appendix A in the following theorem:

Theorem 1 (necessary and sufficient conditions). Let $\{\rho_i, 1 \leqslant i \leqslant m\}$ denote a collection of quantum states with

prior probabilities $\{p_i>0, 1\leq i\leq m\}$. Let Λ denote the set of all ordered sets of Hermitian measurement operators $\Pi=\{\Pi_i\}_{i=0}^m$ that satisfy $\Pi_i\geq 0$, $\Sigma_{i=0}^m\Pi_i=I$ and $\mathrm{Tr}(\Delta\Pi_0)=\beta$ where $\Delta=\Sigma_{i=1}^m p_i\rho_i$, and let Γ denote the set of Hermitian matrices X and scalars δ such that $X\geq p_i\rho_i, 1\leq i\leq m$ and $X\geq\delta\Delta$. Consider the problem $\max_{\Pi\in\Lambda}J(\Pi)$ and the dual problem $\min_{X,\delta\in\Gamma}T(X,\delta)$, where $J(\Pi)=\Sigma_{i=1}^m p_i\mathrm{Tr}(\rho_i\Pi_i)$ and $T(X,\delta)=\mathrm{Tr}(X)-\delta\beta$. Then

(1) For any $X,\delta\in\Gamma$ and $\Pi\in\Lambda$, $T(X,\delta)\geq J(\Pi)$;

(2) there is an optimal Π, denoted $\hat{\Pi}$, such that $\hat{J}=J(\hat{\Pi})\geq J(\Pi)$ for any $\Pi\in\Lambda$;

(3) there are an optimal X and δ, denoted $\hat{X}$ and $\hat{\delta}$, such that $\hat{T}=T(\hat{X},\hat{\delta})\leq T(X,\delta)$ for any $X,\delta\in\Gamma$;

(4) $\hat{T}=\hat{J}$;

(5) necessary and sufficient conditions on the optimal measurement operators $\hat{\Pi}_i$ are that there exists $X,\delta\in\Gamma$ such that

$$(X-p_i\rho_i)\hat{\Pi}_i=0,\quad 1\leq i\leq m,$$

$$(X-\delta\Delta)\hat{\Pi}_0=0;$$

(6) given $\hat{X}$ and $\hat{\delta}$, necessary and sufficient conditions on the optimal measurement operators $\hat{\Pi}_i$ are

$$(\hat{X}-p_i\rho_i)\hat{\Pi}_i=0,\quad 1\leq i\leq m,$$

$$(\hat{X}-\hat{\delta}\Delta)\hat{\Pi}_0=0.$$

In Ref. [5] the authors consider the standard quantum detection problem of choosing m measurement operators $\{\Pi_i, 1\leq i\leq m\}$ to maximize P_D subject to $P_I=0$, and develop a set of necessary and sufficient conditions for optimality on the measurement operators. Specifically, using an approach similar to that used in Appendix A they show that a set of measurement operators $\{\hat{\Pi}_i, 1\leq i\leq m\}$ maximizes P_D subject to $P_I=0$ for a state set $\{\rho_i, 1\leq i\leq m\}$ with prior probabilities $\{p_i, 1\leq i\leq m\}$ if there exists an Hermitian $\hat{X}$ satisfying

$$\hat{X}\geq p_i\rho_i,\quad 1\leq i\leq m,$$

$$(\hat{X}-p_i\rho_i)\hat{\Pi}_i=0,\quad 1\leq i\leq m. \tag{13}$$

Furthermore, they show that the conditions (13) imply that the rank of each optimal measurement operator $\hat{\Pi}_i, 1\leq i\leq m$ is no larger than the rank of the corresponding density operator $\rho_i, 1\leq i\leq m$. In particular, if the quantum state ensemble is a pure-state ensemble consisting of (not necessarily independent) rank-one density operators $\rho_i=|\phi_i\rangle\langle\phi_i|$, then the optimal measurement maximizing P_D subject to $P_I=0$ is a pure-state measurement consisting of rank-one measurement operators $\Pi_i=|\mu_i\rangle\langle\mu_i|$.

The conditions (13) for a set of measurement operators to maximize P_D subject to $P_I=0$ are equivalent to conditions (7) and (9) on a set of measurement operators to maximize P_D subject to $P_I=\beta$. Therefore, using the results in Ref. [5] we have the following proposition:

Proposition 2. Let $\{\rho_i, 1\leq i\leq m\}$ denote a collection of quantum states with prior probabilities $\{p_i, 1\leq i\leq m\}$. Then the optimal measurement that maximizes P_D subject to $P_I=\beta$ consists of measurement operators $\{\Pi_0, 1\leq i\leq m\}$ with $\mathrm{rank}(\Pi_i)\leq\mathrm{rank}(\rho_i), 1\leq i\leq m$. In particular if $\{\rho_i=|\phi_i\rangle\langle\phi_i|, 1\leq i\leq m\}$ is a pure-state quantum ensemble, then the optimal measurement is a pure-state measurement consisting of measurement operators of the form $\{\Pi_i=|\mu_i\rangle\langle\mu_i|, 1\leq i\leq m, \Pi_0=I-\Sigma_{i=1}^m|\mu_i\rangle\langle\mu_i|\}$.

As pointed out in Ref. [23], obtaining a closed-form analytical expression for the optimal measurement operators directly from the necessary and sufficient conditions for optimality of Theorem 1 is a difficult problem. Since Eq. (11) is a (convex) semidefinite programming [24,25,27] problem, there are very efficient methods for solving Eq. (11). In particular, the optimal matrix $\hat{X}$ and optimal scalar $\hat{\delta}$ minimizing $\mathrm{Tr}(X)-\delta\beta$ subject to Eq. (12) can be computed in Matlab using the linear matrix inequality (LMI) Toolbox (see Refs. [5,20] for further details).

A suboptimal measurement that has been suggested as a detection measurement for unambiguous quantum-state discrimination between linearly independent pure quantum-states is the EPM [18–20], in which the measurement vectors are proportional to the reciprocal states associated with the states to be distinguished. Specifically, the EPM corresponding to a set of state vectors $\{|\phi_i\rangle, 1\leq i\leq m\}$ that collectively span $\mathcal{H}$ with prior probabilities $\{p_i, 1\leq i\leq m\}$ consists of the measurement vectors $\{\mu_i, 1\leq i\leq m\}$, where [20]

$$\mu_i=\sqrt{\lambda_n}(\Psi\Psi^*)^{-1}\psi_i,\quad 1\leq i\leq m. \tag{14}$$

Here $|\psi_i\rangle=\sqrt{p_i}\phi_i$, Ψ is the matrix of columns $|\psi_i\rangle$, $\{\lambda_i, 1\leq i\leq n\}$ denote the eigenvalues of $\Psi\Psi^*$ and $\lambda_n=\min\lambda_i$. A general condition under which the EPM is optimal for unambiguous quantum-state discrimination between linearly independent pure quantum-states so that it maximizes P_D subject to $P_E=0$ was derived in Ref. [20]. It was also shown that for GU state sets and for CGU state sets with generators satisfying a certain constraint, the EPM maximizes P_D subject to $P_E=0$.

In the following section we consider a generalization of the EPM to linearly dependent pure states and mixed states, which we refer to as the *scaled inverse measurement* (SIM). We then use the necessary and sufficient conditions for optimality of Theorem 1 to derive a general condition under which the SIM maximizes P_D subject to $P_I=\beta$. In analogy to the EPM, in Secs. V and VI we show that for GU state sets and CGU state sets with generators satisfying a certain constraint, the SIM maximizes P_D subject to $P_I=\beta$ for values of β larger than a threshold, and derive explicit formulas for the optimal measurement operators.

IV. THE SIM AND THE OPTIMAL MEASUREMENT

A. The SIM

The SIM corresponding to a set of density operators $\{\rho_i=\phi_i\phi_i^*, 1\leq i\leq m\}$ with eigenvectors that collectively span $\mathcal{H}$

and prior probabilities $\{p_i, 1 \leq i \leq m\}$ consists of the measurement operators $\{\Sigma_i = \mu_i \mu_i^*, 0 \leq i \leq m\}$, where

$$\mu_i = \gamma (\Psi \Psi^*)^{-1} \psi_i = \gamma \Delta^{-1} \psi_i, \quad 1 \leq i \leq m, \tag{15}$$

and $\Sigma_0 = I - \Sigma_{i=1}^m \mu_i \mu_i^*$. Here Ψ is the matrix of (block) columns $\psi_i = \sqrt{p_i}\phi_i$ and γ is chosen such $P_I = \beta$ and $\Sigma_{i=0}^m \Pi_i = I$. Note that since the eigenvectors of the $\{\rho_i\}$ collectively span $\mathcal{H}$, the columns of the $\{\psi_i\}$ also together span $\mathcal{H}$, so $\Psi \Psi^*$ is invertible.

From Eq. (15)

$$\sum_{i=1}^m \mu_i \mu_i^* = \gamma^2 \Delta^{-1} \left(\sum_{i=1}^m \psi_i \psi_i^* \right) \Delta^{-1} = \gamma^2 \Delta^{-1}, \tag{16}$$

so that

$$\Sigma_0 = I - \sum_{i=1}^m \mu_i \mu_i^* = I - \gamma^2 \Delta^{-1}. \tag{17}$$

It then follows that the probability of an inconclusive result using the SIM is

$$P_I = \mathrm{Tr}(\Delta \Sigma_0) = \mathrm{Tr}(\Delta) - \gamma^2 \mathrm{Tr}(I) = 1 - n\gamma^2. \tag{18}$$

Therefore to satisfy $P_I = \beta$,

$$\gamma = \sqrt{\frac{1-\beta}{n}}. \tag{19}$$

In addition, from Eq. (17) the SIM operators satisfy $\Sigma_{i=0}^m \Pi_i = I$ if and only if $\gamma^2 \leq \lambda_n$ where $\{\lambda_i, 1 \leq i \leq n\}$ denote the eigenvalues of $\Delta = \Psi \Psi^*$ and $\lambda_n = \min \lambda_i$. From Eq. (19) it follows that the SIM is defined only for values of β satisfying

$$\beta \geq 1 - n\lambda_n \stackrel{\triangle}{=} \beta_{\min}. \tag{20}$$

Since the factors ϕ_i are not unique, the SIM factors μ_i are also not unique. If μ_i are the SIM factors corresponding to ϕ_i, then the SIM factors corresponding to $\phi_i' = \phi_i Q_i$ with $Q_i Q_i^* = I$ are $\mu_i' = \mu_i Q_i$. Therefore, although the SIM factors are not unique, the SIM operators $\Sigma_i = \mu_i \mu_i^*$ are unique.

In the case in which the prior probabilities are all equal to a constant p,

$$\mu_i = \frac{\gamma}{\sqrt{p}} (\Phi \Phi^*)^{-1} \phi_i, \quad 1 \leq i \leq m, \tag{21}$$

where Φ is the matrix of (block) columns ϕ_i.

The SIM corresponding to a pure-state ensemble $|\phi_i\rangle$ consists of the measurement vectors $|\mu_i\rangle = \gamma \Delta^{-1} |\psi_i\rangle$, where $|\psi_i\rangle = \sqrt{p_i}|\phi_i\rangle$. From Eq. (14) it follows that in the special case in which the vectors $|\phi_i\rangle$ are linearly independent and $\gamma = \sqrt{\lambda_n}$, the SIM vectors are equal to the EPM vectors. From Eqs. (19) and (20) we then have that $\beta_{\min}$ is equal to the probability of an inconclusive result when using the EPM.

B. Optimality of the SIM

For linearly independent pure quantum-states it was shown in Ref. [20] that the SIM with $\gamma = \sqrt{\lambda_n}$, or equivalently the EPM, maximizes P_D subject to $P_E = 0$ for state sets with equal prior probabilities and strong symmetry properties. The smallest possible probability of an inconclusive result in this case is $\beta = \beta_{\min}$ given by Eq. (20). Theorem 3 below asserts that for a large class of state sets, including those discussed in Ref. [20], the SIM maximizes P_D subject to $P_I = \beta$ for $\beta \geq \beta_{\min}$.

From the necessary and sufficient conditions of Theorem 1 it follows that the SIM is optimal if and only if the measurement operators $\hat{\Pi}_i = \mu_i \mu_i^*, 1 \leq i \leq m$ and $\hat{\Pi}_0 = \Sigma_0$ defined by Eqs. (15) and (17) satisfy Eqs. (9) and (10) for some Hermitian $\hat{X}$ and constant $\hat{\delta}$ satisfying Eqs. (7) and (8). A sufficient condition for optimality of the SIM is given in the following theorem, the proof of which is provided in Appendix B.

Theorem 3 (optimality of the SIM). Let $\{\rho_i = \phi_i \phi_i^*, 1 \leq i \leq m\}$ denote a collection of quantum-states with prior probabilities $\{p_i, 1 \leq i \leq m\}$. Let $\{\Sigma_i = \mu_i \mu_i^*, 1 \leq i \leq m\}$ with $\{\mu_i = \gamma \Delta^{-1} \psi_i, 1 \leq i \leq m\}$ and $\Sigma_0 = I - \Sigma_{i=1}^m \Sigma_i$ denote the scaled inverse measurement (SIM) operators corresponding to $\{\psi_i = \sqrt{p_i}\phi_i, 1 \leq i \leq m\}$, where $\gamma^2 = (1-\beta)/n$, $\Delta = \Psi \Psi^*$ and Ψ is the matrix with block columns ψ_i, and let $\lambda_n = \min \lambda_i$, where λ_i are the eigenvalues of Δ. Then the SIM maximizes P_D subject to $P_I = \beta$ for $\beta \geq \beta_{\min}$ with $\beta_{\min} = 1 - n\lambda_n$ if for each $1 \leq i \leq m$,

$$(1/\gamma) \mu_i^* \psi_i = \psi_i^* \Delta^{-1} \psi_i = \alpha I,$$

where α is a constant independent of i.

It is interesting to note that the condition of Theorem 3 for the SIM to maximize P_D subject to $P_I = \beta$ for $\beta \geq \beta_{\min}$ is identical to the condition given in Theorem 1 of Ref. [11] for the least-squares measurement [10,34], or the square-root measurement [9,35–39], to maximize P_D subject to $P_I = 0$. A similar result for the special case in which the density operators ρ_i are rank-one operators of the form $\rho_i = |\phi_i\rangle\langle\phi_i|$ and the vectors $|\phi_i\rangle$ are linearly independent was derived in Ref. [37]. The least-squares measurement is in general a suboptimal measurement that has been employed as a detection measurement in many applications (see e.g., Refs. [37–39]) and has many desirable properties. Its construction is relatively simple; it can be determined directly from the given collection of states; it maximizes P_D subject to $P_I = 0$ for pure-state ensembles [9,10] and mixed-state ensembles [11] that exhibit certain symmetries; it is ''pretty good'' when the states to be distinguished are equally likely and almost orthogonal [35]; and it is asymptotically optimal [34,36]. Although the sufficient conditions given by Theorem 3 and Theorem 1 of Ref. [11] are the same, the optimal measurements in both cases are different: the SRM consists of m measurement operators $\Pi_i = \mu_i \mu_i^*, 1 \leq i \leq m$ with measurement factors $\mu_i = (\Psi \Psi^*)^{-1/2} \psi_i, 1 \leq i \leq m$, and the SIM consists of $m+1$ measurement operators $\Pi_i = \mu_i \mu_i^*, 1 \leq i \leq m$ and $\Pi_0 = I - \Sigma_{i=1}^m \Pi_i$ with measurement factors $\mu_i = \gamma (\Psi \Psi^*)^{-1} \psi_i, 1 \leq i \leq m$, where $\gamma^2 = (1-\beta)/n$.

We now try to gain some physical insight into the sufficient condition of Theorem 3. First we note that, as we expect, the condition $\psi_i^*\Delta^{-1}\psi_i=\alpha I$ does not depend on the choice of factors ϕ_i. Indeed, if $\phi_i'=\phi_i Q_i$ is another factor of ρ_i with Q_i satisfying $Q_iQ_i^*=I$, and if Ψ' is the matrix of block columns $\psi_i'=\sqrt{p_i}\phi_i'=\sqrt{p_i}\phi_i Q_i$, then it is easy to see that $(\psi_i')^*(\Psi'\Psi'^*)^{-1}\psi_i'=\alpha I$ if and only if $\psi_i^*\Delta^{-1}\psi_i=\alpha I$.

Now, if the state $\rho_i=\phi_i\phi_i^*$ is transmitted with prior probability p_i, then the probability of correctly detecting the state using measurement operators $\Sigma_i=\mu_i\mu_i^*$ is $p_i\mathrm{Tr}(\mu_i^*\phi_i\phi_i^*\mu_i)=\mathrm{Tr}(\mu_i^*\psi_i\psi_i^*\mu_i)$. It follows that if the condition for optimality of Theorem 3 is met, then the probability of correctly detecting each of the states ρ_i using the SIM is the same.

For a pure-state ensemble consisting of density operators $\rho_i=|\phi_i\rangle\langle\phi_i|$ with prior probabilities p_i, the probability of correct detection of the ith state is given by $|\langle\mu_i|\psi_i\rangle|^2$. Since $\langle\mu_i|\psi_i\rangle=\gamma\langle\psi_i|\Delta^{-1}|\psi_i\rangle\geq 0$ for any set of weighted vectors $|\psi_i\rangle$, $\langle\mu_i|\psi_i\rangle$ is constant for all i if and only if $|\langle\mu_i|\psi_i\rangle|^2$ is constant for all i. Therefore, we may interpret the condition in Theorem 3 for pure-state ensembles as follows. The SIM is optimal for a set of states $|\phi_i\rangle$ with prior probabilities p_i and for $\beta\geq\beta_{\min}$ if the probability of correctly detecting each one of the states using the SIM vectors is the same, regardless of the specific state chosen.

For pure-state ensembles with linearly independent state vectors we have already seen that the SIM with $\gamma=\sqrt{\lambda_n}$ is equal to the EPM, and results in $P_E=0$. Furthermore, for this choice of measurement vectors, $P_I=\beta_{\min}$, where $\beta_{\min}$ is given by Eq. (20). Since $P_I+P_D+P_E=1$, it follows that the EPM, or equivalently the SIM with $\gamma=\sqrt{\lambda_n}$, maximizes P_D from all measurements that result in $P_I=\beta_{\min}$. We can now use Theorem 3 to generalize this result to values of β that are larger than $\beta_{\min}$. Specifically, we have the following Corollary to Theorem 3.

Corollary 4. Let $\{\rho_i=|\phi_i\rangle\langle\phi_i|,1\leq i\leq m\}$ denote a pure-state ensemble with linearly independent vectors $|\phi_i\rangle$ and prior probabilities $\{p_i,1\leq i\leq m\}$. Then the scaled-inverse measurement maximizes P_D subject to $P_I=\beta$ for any $\beta\geq\beta_{\min}$.

Proof. From Theorem 3 the SIM maximizes P_D subject to $P_I=\beta$ for any $\beta\geq\beta_{\min}$ if $\langle\psi_i|(\Psi\Psi^*)^{-1}|\psi_i\rangle$ is independent of i, where $|\psi_i\rangle=\sqrt{p_i}|\phi_i\rangle$ and Ψ is the matrix of columns $|\psi_i\rangle$. Now, if the vectors $|\phi_i\rangle$ are linearly independent, then $m=n$ so that Ψ is invertible and $\Psi^*(\Psi\Psi^*)^{-1}\Psi=I$. Since $\langle\psi_i|(\Psi\Psi^*)|\psi_i\rangle$ is the ith diagonal element of $\Psi^*(\Psi\Psi^*)^{-1}\Psi$, $\langle\psi_i|(\Psi\Psi^*)^{-1}|\psi_i\rangle=1$ for all i, and the SIM is optimal. ■

In the remainder of the paper, we use Theorem 3 to derive the optimal measurement for mixed and (not necessarily independent) pure-state sets with certain symmetry properties. The symmetry properties we consider are quite general, and include many cases of practical interest.

V. GEOMETRICALLY UNIFORM STATE SETS

In this section we consider *geometrically uniform* (GU) [40] state sets in which the density operators ρ_i are defined over a group of unitary matrices and are generated by a single generating matrix. We first obtain a convenient characterization of the SIM for GU state sets, and show that under a certain constraint on the generator the SIM is optimal when $\beta\geq\beta_{\min}$. In particular, for (not necessarily independent) pure-state ensembles the SIM is optimal. We then show that for arbitrary GU state sets and arbitrary values of β, the optimal measurement is also GU, and we develop an efficient computational method for finding the optimal generators.

Let $\mathcal{G}=\{U_i,1\leq i\leq m\}$ be a finite group of m unitary matrices U_i. That is, $\mathcal{G}$ contains the identity matrix I, if $\mathcal{G}$ contains U_i, then it also contains its inverse $U_i^{-1}=U_i^*$, and the product U_iU_j of any two elements of $\mathcal{G}$ is in $\mathcal{G}$ [41].

A state set generated by $\mathcal{G}$ using a single generating operator ρ is a set $\mathcal{S}=\{\rho_i=U_i\rho U_i^*, U_i\in\mathcal{G}\}$. The group $\mathcal{G}$ is the *generating group* of $\mathcal{S}$. Such a state set has strong symmetry properties and is called GU. For consistency with the symmetry of $\mathcal{S}$, we will assume equiprobable prior probabilities on $\mathcal{S}$.

If the state set $\{\rho_i,1\leq i\leq m\}$ is GU, then we can always choose factors ϕ_i of ρ_i such that $\{\phi_i=U_i\phi, U_i\in\mathcal{G}\}$, where ϕ is a factor of ρ, so that the factors ϕ_i are also GU with generator ϕ. In the remainder of this section we explicitly assume that the factors are chosen to be GU.

A. Optimality of the SIM for GU States

For a GU state set with equal prior probabilities $1/m$ and generating group $\mathcal{G}$, $\Phi\Phi^*$ commutes with each of the matrices $U_i\in\mathcal{G}$ [28,11]. Consequently, $T=(\Phi\Phi^*)^{-1}$ also commutes with U_i for all i, so that from Eq. (21),

$$\mu_i=\zeta T\phi_i=\zeta TU_i\phi=\zeta U_iT\phi=U_i\mu,\quad 1\leq i\leq m, \tag{22}$$

where $\zeta=\sqrt{m}\gamma$ and

$$\mu=\zeta(\Phi\Phi^*)^{-1}\phi. \tag{23}$$

It follows that the SIM factors μ_i are also GU with generating group $\mathcal{G}$ and generator μ given by Eq. (23). Therefore, to compute the SIM factors for a GU state set all we need is to compute the generator μ. The remaining measurement factors are then obtained by applying the group $\mathcal{G}$ to μ.

From Eq. (22) we have that

$$(1/\gamma)\mu_i^*\psi_i=\frac{1}{\gamma\sqrt{m}}\mu^*U_i^*U_i\phi=\frac{1}{\gamma\sqrt{m}}\mu^*\phi, \tag{24}$$

where ϕ and μ are the generators of the state factors and the SIM factors, respectively. Thus, the probability of correct detection of each one of the states ρ_i using the SIM is the same, regardless of the state transmitted. This then implies from Theorem 3 that for a (not necessarily independent) pure-state GU ensemble the SIM is optimal when $\beta\geq\beta_{\min}$. For a mixed-state ensemble, if the generator ϕ satisfies

$$\phi^*(\Phi\Phi^*)^{-1}\phi=\alpha I \tag{25}$$

for some α, then from Theorem 3 the SIM is again optimal.

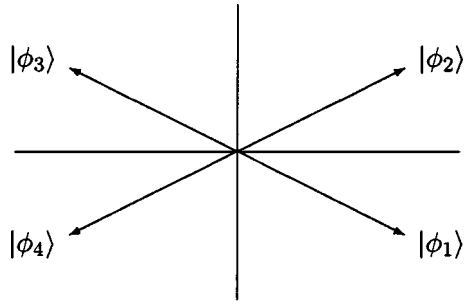

FIG. 1. Example of a GU state set.

B. Example of a GU state set

We now consider an example demonstrating the ideas of the preceding section. Consider the state set $\{\rho_i=|\phi_i\rangle\langle\phi_i|,1\leqslant i\leqslant 4\}$ with equal prior probabilities $\{p_i=1/4,1\leqslant i\leqslant 4\}$, where $\{|\phi_i\rangle=U_i|\phi\rangle,\mathcal{U}_i\in\mathcal{G}\}$. Here $\mathcal{G}=\{U_i,1\leqslant i\leqslant 4\}$ is an Abelian group of unitary matrices

$$U_1=I_2,\quad U_2=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix},$$

$$U_3=\begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix},\quad U_4=\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}, \tag{26}$$

and the generating vector is

$$|\phi\rangle=\frac{1}{2}\begin{bmatrix}\sqrt{3}\\ -1\end{bmatrix}. \tag{27}$$

The matrix U_2 represents a reflection about the x axis, U_3 represents a rotation by π, and U_4 represents a reflection about the y axis. Applying the group $\mathcal{G}$ to the generator $|\phi\rangle$, the GU state vectors are

$$|\phi_1\rangle=|\phi\rangle,\quad |\phi_2\rangle=\frac{1}{2}\begin{bmatrix}\sqrt{3}\\ 1\end{bmatrix},$$

$$|\phi_3\rangle=\frac{1}{2}\begin{bmatrix}-\sqrt{3}\\ 1\end{bmatrix},\quad |\phi_4\rangle=\frac{1}{2}\begin{bmatrix}-\sqrt{3}\\ -1\end{bmatrix}. \tag{28}$$

The GU state set of Eq. (28) is illustrated in Fig. 1.

From Eq. (22) it follows that the SIM vectors corresponding to the state set of Fig. 1 are also GU with generator $|\mu\rangle=2\gamma(\Phi\Phi^*)^{-1}|\phi\rangle$, where Φ is the matrix of columns $|\phi_i\rangle$ and $\gamma^2=(1-\beta)/2$. Here $\beta\geqslant\beta_{\min}$ is the desired rate of inconclusive results. Since

$$\Phi\Phi^*=\begin{bmatrix}3 & 0\\ 0 & 1\end{bmatrix}, \tag{29}$$

$|\mu\rangle$ is given by

$$|\mu\rangle=\gamma\begin{bmatrix}\frac{1}{\sqrt{3}}\\ -1\end{bmatrix}, \tag{30}$$

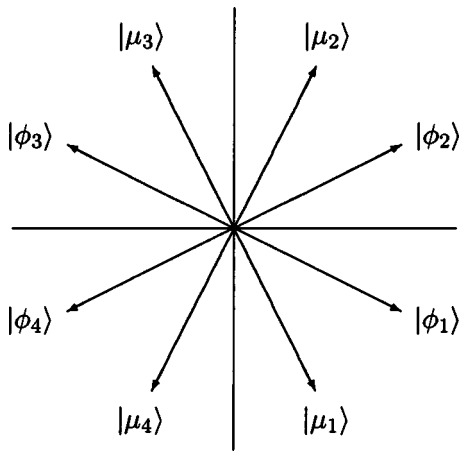

FIG. 2. Symmetry property of the state vectors $|\phi_i\rangle$ given by Eq. (28) and the scaled-inverse measurement (SIM) vectors $|\mu_i\rangle$ which are given by Eq. (31). The state vectors and the SIM vectors both have the same symmetry properties.

and the smallest eigenvalue of $\Psi\Psi^*=(1/4)\Phi\Phi^*$ is $\lambda_n=1/4$. From Eq. (20) it then follows that $\beta_{\min}=1/2$, so that the SIM is defined for values of β satisfying $\beta\geqslant 1/2$. Applying the group $\mathcal{G}$ to the generator $|\mu\rangle$, the SIM vectors are

$$|\mu_1\rangle=|\mu\rangle,\quad |\mu_2\rangle=\gamma\begin{bmatrix}\frac{1}{\sqrt{3}}\\ 1\end{bmatrix},$$

$$|\mu_3\rangle=\gamma\begin{bmatrix}-\frac{1}{\sqrt{3}}\\ 1\end{bmatrix},\quad |\mu_4\rangle=\gamma\begin{bmatrix}-\frac{1}{\sqrt{3}}\\ -1\end{bmatrix}. \tag{31}$$

Comparing Eq. (31) with Eq. (28) it is evident that the SIM vectors have the same symmetries as the original state vectors, as illustrated in Fig. 2.

The probability of correct detection using the SIM vectors, given that a conclusive result was obtained, is

$$\frac{1}{m}\sum_{i=1}^{m}|\langle\mu_i|\phi_i\rangle|^2=|\langle\mu|\phi\rangle|^2=1-\beta,\quad \beta\geqslant\frac{1}{2}. \tag{32}$$

Since, as we have seen in Sec. V A, for GU pure-state sets the SIM maximizes P_D subject to $P_I=\beta$ for any $\beta\geqslant 1/2$, it follows that the probability of correct detection P_D, given that a conclusive result was obtained, using any set of measurement operators $\{\Pi_i,0\leqslant i\leqslant m\}$ with $P_I=\beta$ satisfies

$$P_D\leqslant 1-\beta,\quad \beta\geqslant\frac{1}{2}. \tag{33}$$

Recall that the motivation for allowing for inconclusive results is to be able to increase the probability of correct detection with respect to the probability of correct detection attainable when $P_I=0$. Since the state set of Fig. 1 is GU, the measurement that maximizes P_D subject to $P_I=0$ is the least-squares measurement with measurement vectors $|\chi_i\rangle$

$=(\Phi\Phi^*)^{-1/2}|\phi_i\rangle$ [10,11]. The vectors $|\chi_i\rangle$ are also GU with generating group $\mathcal{G}$ and generating vector

$$|\chi\rangle=(\Phi\Phi^*)^{-1/2}|\phi\rangle=\frac{1}{2}\begin{bmatrix}1\\-1\end{bmatrix}. \tag{34}$$

The probability of correct detection using the least-squares measurement vectors $|\chi_i\rangle$ is

$$P_D=\frac{1}{m}\sum_{i=1}^{m}|\langle\chi_i|\phi_i\rangle|^2=|\langle\chi|\phi\rangle|^2=0.467. \tag{35}$$

Thus, the largest probability of correct detection attainable when $P_I=0$ is equal to 0.467.

If we allow for a probability $\beta=1/2$ of inconclusive results, then from Eq. (32) the largest possible probability of correct detection is equal to 1/2 which is larger than 0.467. Thus, as we expect, by allowing for a nonzero probability of an inconclusive result, we can increase the probability of correct detection.

C. Optimal measurement for arbitrary GU states

If the generator ϕ does not satisfy Eq. (25), or if $\beta<\beta_{\min}$, then the SIM is no longer guaranteed to be optimal. Nonetheless, Proposition 5 below asserts that for a GU state set with generating group $\mathcal{G}$, the optimal measurement operators that maximize P_D subject to $P_I=\beta$ for any β are also GU with generating group $\mathcal{G}$. As we show, the optimal generator can be determined efficiently within any desired accuracy in polynomial time.

Proposition 5. Let $\mathcal{S}=\{\rho_i=U_i\rho U_i^*, U_i\in\mathcal{G}\}$ be a geometrically uniform (GU) state set with equal prior probabilities on an n-dimensional Hilbert space, generated by a finite group $\mathcal{G}$ of unitary matrices, where ρ is an arbitrary generator and let $\{\hat{\Pi}_i,0\leqslant i\leqslant m\}$ denote the measurement operators that maximize P_D subject to $P_I=\beta$ for any $\beta<1$. Then $\{\hat{\Pi}_i,1\leqslant i\leqslant m\}$ are also GU with generating group $\mathcal{G}$.

Proof. Suppose that the optimal measurement operators that maximize

$$J(\{\Pi_i\})=\sum_{i=1}^{m}\mathrm{Tr}(\rho_i\Pi_i), \tag{36}$$

subject to

$$P_I(\{\Pi_i\})=1-\frac{1}{m}\mathrm{Tr}\left(\sum_{i,j=1}^{m}\rho_i\Pi_j\right)=\beta, \tag{37}$$

are $\hat{\Pi}_i$ and let $\hat{J}=J(\{\hat{\Pi}_i\})$. Let $r(j,i)$ be the mapping from $\mathcal{I}\times\mathcal{I}$ to $\mathcal{I}$ with $\mathcal{I}=\{1,\ldots,m\}$, defined by $r(j,i)=k$ if $U_j^*U_i=U_k$. Then the measurement operators $\hat{\Pi}_i^{(j)}=U_j\hat{\Pi}_{r(j,i)}U_j^*,1\leqslant i\leqslant m$ and $\hat{\Pi}_0^{(j)}=I-\Sigma_{i=1}^{m}\hat{\Pi}_i^{(j)}$ for any $1\leqslant j\leqslant m$ are also optimal. Indeed, since $\hat{\Pi}_i\geqslant 0,1\leqslant i\leqslant m$ and $\Sigma_{i=1}^{m}\hat{\Pi}_i\leqslant I$, $\hat{\Pi}_i^{(j)}\geqslant 0,1\leqslant i\leqslant m$ and

$$\sum_{i=1}^{m}\hat{\Pi}_i^{(j)}=U_j\left(\sum_{i=1}^{m}\hat{\Pi}_i\right)U_j^*\leqslant U_jU_j^*=I. \tag{38}$$

Using the fact that $\rho_i=U_i\rho U_i^*$ for some generator ρ for any $1\leqslant j\leqslant m$,

$$\begin{aligned}J(\{\hat{\Pi}_i^{(j)}\})&=\sum_{i=1}^{m}\mathrm{Tr}(\rho U_i^*U_j\hat{\Pi}_{r(j,i)}U_j^*U_i)\\&=\sum_{k=1}^{m}\mathrm{Tr}(\rho U_k^*\hat{\Pi}_kU_k)\\&=\sum_{i=1}^{m}\mathrm{Tr}(\rho_i\hat{\Pi}_i)\\&=\hat{J}.\end{aligned} \tag{39}$$

Finally for any $1\leqslant j\leqslant m$,

$$\begin{aligned}\mathrm{Tr}\left(\sum_{i,s=1}^{m}\rho_i\hat{\Pi}_s^{(j)}\right)&=\mathrm{Tr}\left(\sum_{i,s=1}^{m}U_j^*U_i\rho U_i^*U_j\hat{\Pi}_{r(j,s)}\right)\\&=\mathrm{Tr}\left(\sum_{i,k=1}^{m}U_i\rho U_i^*\hat{\Pi}_k\right)\\&=\mathrm{Tr}\left(\sum_{i,k=1}^{m}\rho_i\hat{\Pi}_k\right),\end{aligned} \tag{40}$$

so that from Eq. (37), $P_I(\{\hat{\Pi}_i^{(j)}\})=P_I(\{\hat{\Pi}_i\})$.

Since the measurement operators $\hat{\Pi}_i^{(j)}$ are optimal for any j, it follows immediately that the measurement operators $\{\overline{\Pi}_i=(1/m)\Sigma_{j=1}^{m}\hat{\Pi}_i^{(j)},1\leqslant i\leqslant m\}$ and $\overline{\Pi}_0=I-\Sigma_{i=1}^{m}\overline{\Pi}_i$ are also optimal. Now, for any $1\leqslant i\leqslant m$,

$$\begin{aligned}\overline{\Pi}_i&=\frac{1}{m}\sum_{j=1}^{m}U_j\hat{\Pi}_{r(j,i)}U_j^*=\frac{1}{m}\sum_{k=1}^{m}U_iU_k^*\hat{\Pi}_kU_kU_i^*\\&=U_i\left(\frac{1}{m}\sum_{k=1}^{m}U_k^*\hat{\Pi}_kU_k\right)U_i^*=U_i\hat{\Pi}U_i^*,\end{aligned} \tag{41}$$

where $\hat{\Pi}=(1/m)\Sigma_{k=1}^{m}U_k^*\hat{\Pi}_kU_k$. We therefore conclude that the optimal measurement operators can always be chosen to be GU with generating group $\mathcal{G}$. ■

From Proposition 5 it follows that the optimal measurement operators satisfy $\Pi_i=U_i\Pi U_i^*,1\leqslant i\leqslant m$ for some generator Π. Thus, to find the optimal measurement operators all we need is to find the generator Π. The remaining operators are obtained by applying the group $\mathcal{G}$ to Π.

Since $\rho_i=U_i\rho U_i^*$, $\mathrm{Tr}(\rho_i\Pi_i)=\mathrm{Tr}(\rho\Pi)$, and the problem (2) reduces to the maximization problem

$$\max_{\Pi\in\mathcal{B}}\mathrm{Tr}(\rho\Pi), \tag{42}$$

where $\mathcal{B}$ is the set of $n\times n$ Hermitian operators, subject to the constraints

$$\Pi \geqslant 0,$$

$$\sum_{i=1}^{m} U_i \Pi U_i^* \leqslant I,$$

$$1 - \mathrm{Tr}\left(\sum_{i=1}^{m} U_i \rho U_i \Pi \right) = \beta. \tag{43}$$

The problem of Eqs. (42) and (43) is a (convex) semidefinite programming problem, and therefore the optimal Π can be computed very efficiently in polynomial time within any desired accuracy [24,25,27], for example, using the LMI toolbox on Matlab. Note that the problem of Eqs. (42) and (43) has n^2 real unknowns and 3 constraints, in contrast with the original maximization problem (2) subject to Eqs. (1) and (4) which has mn^2 real unknowns and $m+2$ constraints.

We summarize our results regarding GU state sets in the following theorem.

Theorem 6 (GU state sets). Let $\mathcal{S}=\{\rho_i = U_i \rho U_i^*, U_i \in \mathcal{G}\}$ be a geometrically uniform (GU) state set with equal prior probabilities on an n-dimensional Hilbert space, generated by a finite group $\mathcal{G}$ of m unitary matrices, where $\rho = \phi\phi^*$ is an arbitrary generator, and let Φ be the matrix of columns $\phi_i = U_i \phi$. Then the scaled inverse measurement (SIM) with $P_I = \beta$ is given by the measurement operators $\Sigma_i = \mu_i \mu_i^*, 0 \leqslant i \leqslant m$ with

$$\mu_i = U_i \mu, \quad 1 \leqslant i \leqslant m,$$

where

$$\mu = \sqrt{m}\, \gamma (\Phi \Phi^*)^{-1} \phi,$$

$\gamma^2 = (1-\beta)/n$, and $\Sigma_0 = I - \Sigma_{i=1}^m \mu_i \mu_i^*$. The SIM has the following properties.

(1) The measurement operators $\Sigma_i, 1 \leqslant i \leqslant m$ are GU with generating group $\mathcal{G}$;

(2) the probability of correctly detecting each of the states ρ_i using the SIM is the same;

(3) if $\phi^* (\Phi \Phi^*)^{-1} \phi = \alpha I$ for some α, then the SIM maximizes P_D subject to $P_I = \beta$ for $\beta \geqslant 1 - n\lambda_n$ where λ_n is the smallest eigenvalue of $(1/m)\Sigma_{i=1}^m \rho_i$; in particular, if $\phi = |\phi\rangle$ is a vector so that the state set is a pure-state ensemble, then the SIM maximizes P_D subject to $P_I = \beta$ for any $\beta \geqslant 1 - n\lambda_n$.

For an arbitrary generator ϕ the optimal measurement operators $\hat{\Pi}_i, 1 \leqslant i \leqslant m$ that maximize P_D subject to $P_I = \beta$ for any β are also GU with generating group $\mathcal{G}$ and generator Π that maximizes $\mathrm{Tr}(\rho\Pi)$ subject to $\Pi \geqslant 0, \Sigma_{i=1}^m U_i \Pi U_i^* \leqslant I$, and $\mathrm{Tr}(\Sigma_{i=1}^m U_i \rho U_i \Pi) = 1 - \beta$.

VI. COMPOUND GEOMETRICALLY UNIFORM STATE SETS

We now consider *compound geometrically uniform* (CGU) [28] state sets which consist of subsets that are GU. As we show, the SIM operators are also CGU so that they can be computed using a *set* of generators. Under a certain condition on the generators and for $\beta \geqslant \beta_{\min}$, we show that the optimal measurement associated with a CGU state set is equal to the SIM. For arbitrary CGU state sets and arbitrary values of β we show that the optimal measurement operators are CGU, and we derive an efficient computational method for finding the optimal generators.

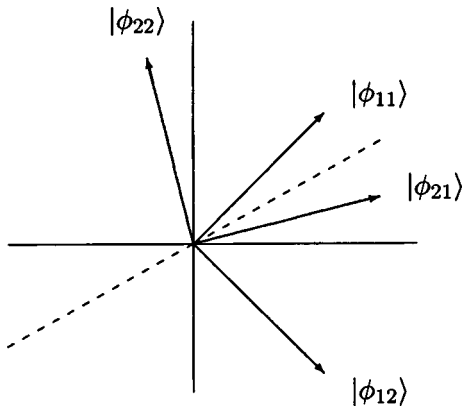

FIG. 3. A compound geometrically uniform pure-state set. The state sets $\mathcal{S}_1 = \{|\phi_{11}\rangle, |\phi_{21}\rangle\}$ and $\mathcal{S}_2 = \{|\phi_{12}\rangle, |\phi_{22}\rangle\}$ are both geometrically uniform (GU) with the same generating group; both sets are invariant under a reflection about the dashed line. However, the combined set $\mathcal{S} = \{|\phi_{11}\rangle, |\phi_{21}\rangle, |\phi_{12}\rangle, |\phi_{22}\rangle\}$ is no longer GU.

A CGU state set is defined as a set of density operators $\mathcal{S} = \{\rho_{ik} = \phi_{ik}\phi_{ik}^*, 1 \leqslant i \leqslant l, 1 \leqslant k \leqslant r\}$ such that $\rho_{ik} = U_i \rho_k U_i^*$, where the matrices $\{U_i, 1 \leqslant i \leqslant l\}$ are unitary and form a group $\mathcal{G}$, and the operators $\{\rho_k, 1 \leqslant k \leqslant r\}$ are the generators. We assume equiprobable prior probabilities on $\mathcal{S}$.

If the state set $\{\rho_{ik}, 1 \leqslant i \leqslant l, 1 \leqslant k \leqslant r\}$ is CGU, then we can always choose factors ϕ_{ik} of ρ_{ik} such that $\{\phi_{ik} = U_i \phi_k, 1 \leqslant i \leqslant l\}$, where ϕ_k is a factor of ρ_k, so that the factors ϕ_{ik} are also CGU with generators $\{\phi_k, 1 \leqslant k \leqslant r\}$. In the remainder of this section we explicitly assume that the factors are chosen to be CGU.

A CGU state set is in general not GU. However, for every k, the matrices $\{\phi_{ik}, 1 \leqslant i \leqslant l\}$ and the operators $\{\rho_{ik}, 1 \leqslant i \leqslant l\}$ are GU with generating group $\mathcal{G}$.

An example of a CGU state set is illustrated in Fig. 3. In this example the state set is $\{\rho_{ik} = |\phi_{ik}\rangle\langle\phi_{ik}|, 1 \leqslant i, k \leqslant 2\}$, where $\{|\phi_{ik}\rangle = U_i |\phi_k\rangle, U_i \in \mathcal{G}\}$, $\mathcal{G} = \{I_2, U\}$ with

$$U = \frac{1}{2} \begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}, \tag{44}$$

and the generating vectors are

$$|\phi_1\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad |\phi_2\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \tag{45}$$

The matrix U represents a reflection about the dashed line in Fig. 3. Thus, the vector $|\phi_{21}\rangle$ is obtained by reflecting the generator $|\phi_{11}\rangle$ about this line, and similarly the vector $|\phi_{22}\rangle$ is obtained by reflecting the generator $|\phi_{12}\rangle$ about this line.

As can be seen from the figure, the state set is not GU. In particular, there is no isometry that transforms $|\phi_{11}\rangle$ into

$|\phi_{12}\rangle$ while leaving the set invariant. However, the sets $\mathcal{S}_1=\{|\phi_{11}\rangle,|\phi_{21}\rangle\}$ and $\mathcal{S}_2=\{|\phi_{12}\rangle,|\phi_{22}\rangle\}$ are both GU with generating group $\mathcal{G}$.

A. Optimality of the SIM for CGU state sets

With Φ denoting the matrix of (block) columns ϕ_{ik}, it was shown in Refs. [11,28] that $\Phi\Phi^*$, and consequently $T=(\Phi\Phi^*)^{-1}$, commutes with each of the matrices $U_i\in\mathcal{G}$. Thus, the SIM operators are $\Sigma_{ik}=\mu_{ik}\mu_{ik}^*, 1\leq i\leq l, 1\leq k\leq r$ with

$$\mu_{ik}=\zeta T\phi_{ik}=\zeta TU_i\phi_k=U_i\mu_k, \tag{46}$$

where $\zeta=\gamma\sqrt{lr}$ and

$$\mu_k=\zeta T\phi_k=\gamma\sqrt{lr}(\Phi\Phi^*)^{-1}\phi_k. \tag{47}$$

Therefore, the SIM factors are also CGU with generating group $\mathcal{G}$ and generators μ_k given by Eq. (47). To compute the SIM factors all we need is to compute the generators μ_k. The remaining measurement factors are then obtained by applying the group $\mathcal{G}$ to each of the generators.

From Eq. (46),

$$\mu_{ik}^*\phi_{ik}=\mu_k^*U_i^*U_i\phi_k=\mu_k^*\phi_k, \tag{48}$$

so that from Theorem 3 the SIM is optimal if

$$\mu_k^*\phi_k=\gamma\sqrt{lr}\phi_k^*T\phi_k=\alpha I, \quad 1\leq k\leq r \tag{49}$$

for some constant α.

B. CGU state sets with GU generators

A special class of CGU state sets is *CGU state sets with GU generators* in which the generators $\{\rho_k=\phi_k\phi_k^*, 1\leq k\leq r\}$ and the factors ϕ_k are themselves GU. Specifically, $\{\phi_k=V_k\phi\}$ for some generator ϕ, where the matrices $\{V_k, 1\leq k\leq r\}$ are unitary, and form a group $\mathcal{Q}$.

Suppose that U_i and V_k commute up to a phase factor for all i and k so that $U_iV_k=V_kU_ie^{j\theta(i,k)}$, where $\theta(i,k)$ is an arbitrary phase function that may depend on the indices i and k. In this case we say that $\mathcal{G}$ and $\mathcal{Q}$ commute up to a phase factor and that the corresponding state set is *CGU with commuting GU generators*. (In the special case in which $\theta=0$ so that $U_iV_k=V_kU_i$ for all i,k, the resulting state set is GU [28]). Then for all i,k, $\Phi\Phi^*$ commutes with U_iV_k [11], and the SIM factors μ_{ik} are given by

$$\mu_{ik}=\zeta T\phi_{ik}=\zeta TU_iV_k\phi=U_iV_k\bar{\mu}, \tag{50}$$

where $\bar{\mu}=\zeta T\phi$ and $\zeta=\gamma\sqrt{lr}$. Thus even though the state set is not in general GU, the SIM factors can be computed using a single generator.

Alternatively, we can express μ_{ik} as $\mu_{ik}=U_i\mu_k$, where the generators μ_k are given by

$$\mu_k=V_k\bar{\mu}. \tag{51}$$

From Eq. (51) it follows that the generators μ_k are GU with generating group $\mathcal{Q}=\{V_k, 1\leq k\leq r\}$ and generator $\bar{\mu}$. Then for all k,

$$\mu_k^*\phi_k=\bar{\mu}^*V_k^*V_k\phi=\bar{\mu}^*\phi. \tag{52}$$

If in addition,

$$\bar{\mu}^*\phi=\gamma\phi^*T\phi=\alpha I \tag{53}$$

for some α, then combining Eqs. (48), (52), and (53) with Theorem 3 we conclude that the SIM is optimal. In particular, for a pure-state ensemble, $\bar{\mu}^*\phi$ is a scalar so that Eq. (53) is always satisfied. Therefore, for a pure CGU state set with commuting GU generators, the SIM maximizes P_D subject to $P_I=\beta$ for $\beta\geq\beta_{\min}$.

C. Optimal measurement for arbitrary CGU states

If the generators ϕ_k do not satisfy Eq. (49), or if $\beta<\beta_{\min}$, then the SIM is no longer guaranteed to be optimal. Nonetheless, in a manner similar to that of Sec. V C, we show that the optimal measurement operators that maximize P_D subject to $P_I=\beta$ are CGU with generating group $\mathcal{G}$. The corresponding generators can be computed very efficiently in polynomial time within any desired accuracy.

Proposition 7. Let $\mathcal{S}=\{\rho_{ik}=U_i\rho_kU_i^*, 1\leq i\leq l, 1\leq k\leq r\}$ be a compound geometrically uniform (CGU) state set with equal prior probabilities on an n-dimensional Hilbert space, generated by a finite group $\mathcal{G}$ of unitary matrices and generators ρ_k and let $\{\hat{\Pi}_{ik}, 1\leq i\leq, 1\leq k\leq r\}$ and $\hat{\Pi}_0$ denote the measurement operators that maximize P_D subject to $P_I=\beta$ for any $\beta<1$. Then $\{\hat{\Pi}_{ik}, 1\leq i\leq l, 1\leq k\leq r\}$ are also CGU with generating group $\mathcal{G}$.

Proof. The proof is analogous to the proof of Proposition 5 and is given in Appendix C. ■

From Proposition 7 it follows that the optimal measurement operators satisfy $\Pi_{ik}=U_i\Pi_kU_i^*, 1\leq i\leq l, 1\leq k\leq r$ for some generators $\Pi_k, 1\leq k\leq r$. Thus, to find the optimal measurement operators all we need is to find the optimal generators $\Pi_k, 1\leq k\leq r$. The remaining operators are obtained by applying the group $\mathcal{G}$ to each of the generators.

Since $\rho_{ik}=U_i\rho_kU_i^*$, $\mathrm{Tr}(\rho_{ik}\Pi_{ik})=\mathrm{Tr}(\rho_k\Pi_k)$, and the problem (2) reduces to the maximization problem

$$\max_{\Pi_k\in\mathcal{B}}\sum_{k=1}^{r}\mathrm{Tr}(\rho_k\Pi_k), \tag{54}$$

subject to the constraints

$$\Pi_k\geq 0, \quad 1\leq k\leq r,$$

$$\sum_{i=1}^{l}\sum_{k=1}^{r}U_i\Pi_kU_i^*\leq I,$$

$$1-\frac{1}{r}\mathrm{Tr}\left(\sum_{i=1}^{l}\sum_{k,l=1}^{r}U_i\rho_kU_i\Pi_l\right)=\beta. \tag{55}$$

The problem of Eqs. (54) and (55) is a (convex) semidefinite programming problem, and therefore the optimal generators Π_k can be computed very efficiently in polynomial time within any desired accuracy [24,25,27], for example, using the LMI toolbox on Matlab. Note that the problem of Eqs. (54) and (55) has rn^2 real unknowns and $r+2$ constraints, in contrast with the original maximization (2) subject to Eqs. (1) and (4) which has lrn^2 real unknowns and $lr+2$ constraints.

We summarize our results regarding CGU state sets in the following theorem.

Theorem 8 (CGU state sets). Let $\mathcal{S}=\{\rho_{ik}=U_i\rho_k U_i^*, 1\leqslant i\leqslant l, 1\leqslant k\leqslant r\}$ be a compound geometrically uniform (CGU) state set with equal prior probabilities on an n-dimensional Hilbert space generated by a finite group $\mathcal{G}$ of unitary matrices and generators $\{\rho_k=\phi_k\phi_k^*, 1\leqslant k\leqslant r\}$, and let Φ be the matrix of columns $\phi_{ik}=U_i\phi_k$. Then the scaled inverse measurement (SIM) with $P_I=\beta$ is given by the measurement operators $\Sigma_{ik}=\mu_{ik}\mu_{ik}^*, 1\leqslant i\leqslant l, 1\leqslant k\leqslant r$ and $\Sigma_0=I-\Sigma_{i,k}\mu_{ik}\mu_{ik}^*$ with

$$\mu_{ik}=U_i\mu_k,$$

where

$$\mu_k=\sqrt{rl}\,\gamma(\Phi\Phi^*)^{-1}\phi_k,$$

and $\gamma^2=(1-\beta)/n$. The SIM has the following properties.

(1) The measurement operators $\Sigma_{ik}, 1\leqslant i\leqslant l, 1\leqslant k\leqslant r$ are CGU with generating group $\mathcal{G}$;

(2) the probability of correctly detecting each of the states ϕ_{ik} for fixed k using the SIM is the same;

(3) if $\phi_k^*(\Phi\Phi^*)^{-1}\phi_k=\alpha I$ for some α and for $1\leqslant k\leqslant r$, then the SIM maximizes P_D subject to $P_I=\beta$ with $\beta\geqslant 1-n\lambda_n$ where λ_n is the smallest eigenvalue of $(1/lr)\Sigma_{i,k}\rho_{ik}$.

If in addition the generators $\{\phi_k=V_k\phi, 1\leqslant k\leqslant r\}$ are geometrically uniform with $U_iV_k=V_kU_ie^{j\theta(i,k)}$ for all i,k, then

(1) $\mu_{ik}=U_iV_k\overline{\mu}$ where $\overline{\mu}=\sqrt{rl}\,\gamma(\Phi\Phi^*)^{-1}\phi$ so that the SIM operators are CGU with geometrically uniform generators;

(2) the probability of correctly detecting each of the states ϕ_{ik} using the SIM is the same;

(3) if $\phi^*(\Phi\Phi^*)^{-1}\phi=\alpha I$ for some α, then the SIM maximizes P_D subject to $P_I=\beta$ with $\beta\geqslant 1-n\lambda_n$. In particular, if $\phi=|\phi\rangle$ is a vector so that the state set is a pure-state ensemble, then the SIM maximizes P_D subject to $P_I=\beta$ with $\beta\geqslant 1-n\lambda_n$.

For arbitrary CGU state sets the optimal measurement operators $\hat{\Pi}_{ik}, 1\leqslant i\leqslant l, 1\leqslant k\leqslant r$ that maximize P_D subject to $P_I=\beta$ for any β are CGU with generating group $\mathcal{G}$ and generators Π_k that maximize $\Sigma_{k=1}^r\mathrm{Tr}(\rho_k\Pi_k)$ subject to

$$\Pi_k\geqslant 0, 1\leqslant k\leqslant r, \quad \Sigma_{i,k}U_i\Pi_kU_i^*\leqslant I,$$

and

$$\mathrm{Tr}(\Sigma_{i=1}^l\Sigma_{k,l=1}^r U_i\rho_kU_i\Pi_l)=r(1-\beta).$$

VII. CONCLUSION

In this paper, we considered the optimal measurement operators that maximize the probability of correct detection given a fixed probability β of an inconclusive result, when distinguishing between a collection of *mixed* quantum states. We first derived a set of necessary and sufficient conditions for optimality by exploiting principles of duality theory in vector space optimization. Using these conditions, we derived a general condition under which the SIM is optimal. We then considered state sets with a broad class of symmetry properties for which the SIM is optimal. Specifically, we showed that for GU state sets and for CGU state sets with generators that satisfy certain constraints and for values of β exceeding a threshold, the SIM is optimal. We also showed that for arbitrary GU and CGU state sets and for arbitrary values of β, the optimal measurement operators have the same symmetries as the original state sets. Therefore, to compute the optimal measurement operators, we need only to compute the corresponding generators. As we showed, the generators can be computed very efficiently in polynomial time within any desired accuracy by solving a semidefinite programming problem.

ACKNOWLEDGMENTS

The author wishes to thank Professor A. Megretski and Professor G. C. Verghese for many valuable discussions and Professor J. Fiurášek for drawing her attention to Ref. [23]. The author is financially supported by the Taub foundation.

APPENDIX A: NECESSARY AND SUFFICIENT CONDITIONS FOR OPTIMALITY

Denote by Λ the set of all ordered sets

$$\Pi=\{\Pi_i\}_{i=0}^m, \Pi_i\in\mathcal{B}$$

satisfying Eq. (1) and (4) with $\beta<1$, and define $J(\Pi)=\Sigma_{i=1}^m p_i\mathrm{Tr}(\rho_i\Pi_i)$. Then our problem is

$$\max_{\Pi\in\Lambda} J(\Pi). \tag{A1}$$

We refer to this problem as the primal problem, and to any $\Pi\in\Lambda$ as a primal feasible point. The optimal value of $J(\Pi)$ is denoted by $\hat{J}$.

To derive necessary and sufficient conditions for optimality, we now formulate a *dual problem* whose optimal value serves as a certificate for $\hat{J}$. As described in Ref. [5], a general method for deriving a dual problem is to invoke the separating hyperplane theorem [42], which states that two disjoint convex sets [45] can always be separated by a hyperplane. We will take one convex set to be the point 0, and then carefully construct another convex set that does not contain 0, and that captures the equality constraints in the primal problem and the fact that for any primal feasible point, the value of the primal function is no larger than the optimal value. The dual variables will then emerge from the parameters of the separating hyperplane.

In our problem we have two equality constraints, $\Sigma_{i=0}^{m}\Pi_i=I$ and $\mathrm{Tr}(\Delta\Pi_0)=\beta$ and we know that $\hat{J}\geq J(\Pi)$. Our constructed convex set will accordingly consist of matrices of the form $-I+\Sigma_{i=0}^{m}\Pi_i$, where $\Pi_i\in\mathcal{B}$ and $\Pi_i\geq 0$, scalars of the form $\beta-\mathrm{Tr}(\Delta\Pi_0)$, and scalars of the form $r-J(\Pi)$, where $r>\hat{J}$. We thus consider the real vector space

$$\mathcal{L}=\mathcal{B}\times\mathcal{R}\times\mathcal{R}=\{(S,x,y):\ S\in\mathcal{B},\ x,y\in\mathcal{R}\},$$

where $\mathcal{R}$ denotes the reals, with inner product defined by

$$\langle(W,z,t),(S,x,y)\rangle=\mathrm{Tr}(WS)+zx+ty. \tag{A2}$$

We then define the subset Ω of $\mathcal{L}$ as points of the form

$$\Omega=\left(-I+\sum_{i=0}^{m}\Pi_i,\beta-\mathrm{Tr}(\Delta\Pi_0),r-\sum_{i=1}^{m}p_i\mathrm{Tr}(\Pi_i\rho_i)\right), \tag{A3}$$

where $\Pi_i\in\mathcal{B},\Pi_i\geq 0,r\in\mathcal{R}$ and $r>\hat{J}$.

It is easily verified that Ω is convex, and $0\notin\Omega$. Therefore, by the separating hyperplane theorem, there exists a nonzero vector $(Z,a,b)\in\mathcal{L}$ such that $\langle(Z,a,b),(Q,c,d)\rangle\geq 0$ for all $(Q,c,d)\in\Omega$, i.e.,

$$\mathrm{Tr}\left(Z\left(-I+\sum_{i=0}^{m}\Pi_i\right)\right)+b(\beta-\mathrm{Tr}(\Delta\Pi_0))+a\left(r-\sum_{i=1}^{m}p_i\mathrm{Tr}(\Pi_i\rho_i)\right)\geq 0 \tag{A4}$$

for all $\Pi_i\in\mathcal{B}$ and $r\in\mathcal{R}$ such that $\Pi_i\geq 0$, $r>\hat{J}$.

As we now show, the hyperplane parameters (Z,a,b) have to satisfy certain constraints, which lead to the formulation of the dual problem. Specifically, Eq. (A4) with $\Pi_i=0$, $r\to\hat{J}$ implies

$$a\hat{J}\geq\mathrm{Tr}(Z)-b\beta. \tag{A5}$$

Similarly, Eq. (A4) with $r=\hat{J}+1$, $\Pi_j=0$ for $j\neq i$, $\Pi_i=t|x\rangle\langle x|$ for one value $1\leq i\leq m$, where $|x\rangle\in\mathbb{C}^n$ is fixed and $t\to+\infty$ yields $\langle x|Z-ap_i\rho_i|x\rangle\geq 0$. Since $|x\rangle$ and i are arbitrary, this implies

$$Z\geq ap_i\rho_i,\quad 1\leq i\leq m. \tag{A6}$$

With $r=\hat{J}+1$, $\Pi_j=0$ for $j\neq 0$, $\Pi_0=t|x\rangle\langle x|$, where $|x\rangle\in\mathbb{C}^n$ is fixed and $t\to+\infty$, (A4) yields $\langle x|Z-b\Delta|x\rangle\geq 0$, which implies

$$Z\geq b\Delta. \tag{A7}$$

With $\Pi_i=0,0\leq i\leq m$, $r\to+\infty$, Eq. (A4) implies $a\geq 0$. If $a=0$, then Eq. (A5) yields $\mathrm{Tr}(Z)\leq b\beta<b$ and (A7) yields $\mathrm{Tr}(Z)\geq b$. Therefore we conclude that $a>0$, and define $\hat{X}=Z/a$, $\hat{\delta}=b/a$. Then Eq. (A5) implies that

$$T(\hat{X},\hat{\delta})\leq\hat{J}, \tag{A8}$$

where $T(X,\delta)=\mathrm{Tr}(X)-\delta\beta$, Eq. (A6) implies that $\hat{X}\geq p_i\rho_i$ for $1\leq i\leq m$, and Eq. (A7) implies that $\hat{X}\geq\hat{\delta}\Delta$.

Let Γ be the set of $X\in\mathcal{B}$, $\delta\in\mathcal{R}$ satisfying $X\geq p_i\rho_i,1\leq i\leq m$ and $X\geq\delta\Delta$. Then for any $X,\delta\in\Gamma$, $\Pi\in\Lambda$, we have

$$T(X,\delta)-J(\Pi)=\mathrm{Tr}\left(\sum_{i=1}^{m}\Pi_i(X-p_i\rho_i)\right)+\mathrm{Tr}(\Pi_0(X-\delta\Delta))\geq 0. \tag{A9}$$

Since $\hat{X}\in\Gamma$, from Eqs. (A8) and (A9) we conclude that $T(\hat{X},\hat{\delta})=\hat{J}$.

Thus we have proven that the dual problem associated with Eq. (A1) is

$$\min_{X\in\mathcal{B},\delta\in\mathcal{R}}\{\mathrm{Tr}(X)-\delta\beta\}, \tag{A10}$$

subject to

$$X\geq p_i\rho_i,\quad 1\leq i\leq m,$$

$$X\geq\delta\Delta. \tag{A11}$$

Furthermore, we have shown that there exists an optimal $\hat{X},\hat{\delta}\in\Gamma$ and an optimal value $\hat{T}=T(\hat{X},\hat{\delta})$ such that $\hat{T}=\hat{J}$.

Let $\hat{\Pi}_i$ denote the optimal measurement operators. Then combining Eq. (A9) with $\hat{T}=\hat{J}$, we conclude that

$$(\hat{X}-p_i\rho_i)\hat{\Pi}_i=0,\quad 1\leq i\leq m,$$

$$(\hat{X}-\hat{\delta}\Delta)\hat{\Pi}_0=0. \tag{A12}$$

Once we find the optimal $\hat{X}$ and $\hat{\delta}$ that minimize the dual problem Eq. (A10), the constraints Eq. (A12) are necessary and sufficient conditions on the optimal measurement operators $\hat{\Pi}_i$. We have already seen that these conditions are necessary. To show that they are sufficient, we note that if a set of feasible measurement operators Π_i satisfies (A12), then $\Sigma_{i=1}^{m}\mathrm{Tr}(\Pi_i(\hat{X}-p_i\rho_i))=0$ and $\mathrm{Tr}((\hat{X}-\hat{\delta}\Delta)\hat{\Pi}_0)=0$ so that from (A9), $J(\Pi)=T(\hat{X},\hat{\delta})=\hat{J}$.

APPENDIX B: PROOF OF THEOREM 3

In this appendix we prove Theorem 3. Specifically, we show that for a set of states $\rho_i=\phi_i\phi_i^*$ with prior probabilities p_i, if $(1/\gamma)\mu_i^*\psi_i=\alpha I,1\leq i\leq m$, where $\mu_i=\gamma(\Psi\Psi^*)^{-1}\psi_i=\gamma\Delta^{-1}\psi_i$ are the SIM factors and $\psi_i=\sqrt{p_i}\phi_i$, then there exists an Hermitian X and a constant δ such that

$$X\geq\psi_i\psi_i^*,\quad 1\leq i\leq m, \tag{B1}$$

$$X\geq\delta\Delta, \tag{B2}$$

$$(X-\psi_i\psi_i^*)\mu_i\mu_i^*=0,\quad 1\leq i\leq m, \tag{B3}$$

$$(X-\delta\Delta)(I-\gamma^2\Delta^{-1})=0. \tag{B4}$$

Let $X=\alpha\Delta$ and $\delta=\alpha$. Then Eqs. (B2) and (B4) are immediately satisfied. Next, since $\alpha I=\psi_i^*\Delta^{-1}\psi_i=\psi_i^*\Delta^{-1/2}\Delta^{-1/2}\psi_i$, it follows that

$$\alpha I\geqslant\Delta^{-1/2}\psi_i\psi_i^*\Delta^{-1/2}. \tag{B5}$$

Multiplying both sides of Eq. (B5) by $\Delta^{1/2}$ we have

$$\alpha\Delta\geqslant\psi_i\psi_i^*, \tag{B6}$$

which verifies that the conditions (B1) are satisfied.

Finally,

$$(X-\psi_i\psi_i^*)\mu_i=\alpha\gamma\Delta\Delta^{-1}\psi_i-\alpha\gamma\psi_i=0, \tag{B7}$$

so that the conditions (B3) are also satisfied.

APPENDIX C: PROOF OF PROPOSITION 7

Suppose that the optimal measurement operators that maximize

$$J(\{\Pi_{ik}\})=\sum_{i=1}^{l}\sum_{k=1}^{r}\mathrm{Tr}(\rho_{ik}\Pi_{ik}), \tag{C1}$$

subject to

$$P_I(\{\Pi_{ik}\})=1-\frac{1}{lr}\mathrm{Tr}\left(\sum_{i,j=1}^{l}\sum_{k,s=1}^{r}\rho_{ik}\Pi_{js}\right)=\beta, \tag{C2}$$

are $\hat{\Pi}_{ik}$, and let $\hat{J}=J(\{\hat{\Pi}_{ik}\})$. Let $r(j,i)$ be the mapping from $\mathcal{I}\times\mathcal{I}$ to $\mathcal{I}$ with $\mathcal{I}=\{1,\ldots,l\}$, defined by $r(j,i)=s$ if $U_j^*U_i=U_s$. Then the measurement operators $\hat{\Pi}_{ik}^{(j)}=U_j\hat{\Pi}_{r(j,i)k}U_j^*$ for any $1\leqslant j\leqslant l$ are also optimal. Indeed, since $\hat{\Pi}_{ik}\geqslant 0, 1\leqslant i\leqslant l, 1\leqslant k\leqslant r$ and $\Sigma_{i=1}^{l}\Sigma_{k=1}^{r}\hat{\Pi}_{ik}\leqslant I$, $\hat{\Pi}_{ik}^{(j)}\geqslant 0, 1\leqslant i\leqslant l, 1\leqslant k\leqslant r$ and

$$\sum_{i=1}^{l}\sum_{k=1}^{r}\hat{\Pi}_{ik}^{(j)}=U_j\left(\sum_{i=1}^{l}\sum_{k=1}^{r}\hat{\Pi}_{ik}\right)U_j^*\leqslant U_jU_j^*=I. \tag{C3}$$

Using the fact that $\rho_{ik}=U_i\rho_kU_i^*$ for some generators ρ_k,

$$\begin{aligned}J(\{\hat{\Pi}_{ik}^{(j)}\})&=\sum_{i=1}^{l}\sum_{k=1}^{r}\mathrm{Tr}(\rho_kU_i^*U_j\hat{\Pi}_{r(j,i)k}U_j^*U_i)\\&=\sum_{s=1}^{l}\sum_{k=1}^{r}\mathrm{Tr}(\rho_kU_s^*\hat{\Pi}_{sk}U_s)\\&=\sum_{i=1}^{l}\sum_{k=1}^{r}\mathrm{Tr}(\rho_{ik}\hat{\Pi}_{ik})=\hat{J}.\end{aligned} \tag{C4}$$

Finally,

$$\begin{aligned}\mathrm{Tr}\left(\sum_{i,s=1}^{l}\sum_{k,t=1}^{r}\rho_{ik}\hat{\Pi}_{st}^{(j)}\right)&=\mathrm{Tr}\left(\sum_{i,s=1}^{l}\sum_{k,t=1}^{r}U_j^*U_i\rho_kU_i^*U_j\hat{\Pi}_{st}\right)\\&=\mathrm{Tr}\left(\sum_{i,s=1}^{l}\sum_{k,t=1}^{r}U_i\rho_kU_i^*\hat{\Pi}_{st}\right)\\&=\mathrm{Tr}\left(\sum_{i,s=1}^{l}\sum_{k,t=1}^{r}\rho_{ik}\hat{\Pi}_{st}\right),\end{aligned} \tag{C5}$$

so that from Eq. (C2), $P_I(\{\hat{\Pi}_{ik}^{(j)}\})=P_I(\{\hat{\Pi}_{ik}\})$.

Since the measurement operators $\hat{\Pi}_{ik}^{(j)}$ are optimal for any j, it follows immediately that the measurement operators $\{\overline{\Pi}_{ik}=(1/l)\Sigma_{j=1}^{l}\hat{\Pi}_{ik}^{(j)}, 1\leqslant i\leqslant l, 1\leqslant k\leqslant r\}$ and $\overline{\Pi}_0=I-\Sigma_{i,k}\overline{\Pi}_{ik}$ are also optimal. Now, for any $1\leqslant i\leqslant l, 1\leqslant k\leqslant r$,

$$\begin{aligned}\overline{\Pi}_{ik}&=\frac{1}{l}\sum_{j=1}^{l}U_j\hat{\Pi}_{r(j,i)k}U_j^*=\frac{1}{l}\sum_{s=1}^{l}U_iU_s^*\hat{\Pi}_{sk}U_sU_i^*\\&=U_i\left(\frac{1}{l}\sum_{s=1}^{l}U_s^*\hat{\Pi}_{sk}U_s\right)U_i^*=U_i\hat{\Pi}_kU_i^*,\end{aligned} \tag{C6}$$

where $\hat{\Pi}_k=(1/l)\Sigma_{s=1}^{l}U_s^*\hat{\Pi}_{sk}U_s$.

We therefore conclude that the optimal measurement operators can always be chosen to be CGU with the same generating group $\mathcal{G}$ as the original state set. Thus, to find the optimal measurement operators all we need is to find the optimal generators $\{\hat{\Pi}_k, 1\leqslant k\leqslant r\}$. The remaining operators are obtained by applying the group $\mathcal{G}$ to each of the generators.

[1] A. Peres, Found. Phys. **20**, 1441 (1990).
[2] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, Boston, MA, 1995).
[3] A.S. Holevo, J. Multivariate Anal. **3**, 337 (1973).
[4] H.P. Yuen, R.S. Kennedy, and M. Lax, IEEE Trans. Inf. Theory **IT-21**, 125 (1975).
[5] Y.C. Eldar, A. Megretski, and G.C. Verghese, IEEE Trans. Inform. Theory (to be published).
[6] C.W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
[7] M. Charbit, C. Bendjaballah, and C.W. Helstrom, IEEE Trans. Inform. Theory **35**, 1131 (1989).
[8] M. Osaki, M. Ban, and O. Hirota, Phys. Rev. A **54**, 1691 (1996).
[9] M. Ban, K. Kurokawa, R. Momose, and O. Hirota, Int. J. Theor. Phys. **36**, 1269 (1997).
[10] Y.C. Eldar and G.D. Forney, Jr., IEEE Trans. Inform. Theory **47**, 858 (2001).
[11] Y.C. Eldar, A. Megretski, and G.C. Verghese, e-print quant-ph/0211111.
[12] C.W. Helstrom, IEEE Trans. Inform. Theory **28**, 359 (1982).
[13] I.D. Ivanovic, Phys. Lett. A **123**, 257 (1987).
[14] D. Dieks, Phys. Lett. A **126**, 303 (1988).
[15] A. Peres, Phys. Lett. A **128**, 19 (1988).
[16] G. Jaeger and A. Shimony, Phys. Lett. A **197**, 83 (1995).
[17] A. Peres and D.R. Terno, J. Phys. A **31**, 7105 (1998).

[18] A. Chefles, Phys. Lett. A **239**, 339 (1998).
[19] A. Chefles and S.M. Barnett, Phys. Lett. A **250**, 223 (1998).
[20] Y.C. Eldar, IEEE Trans. Inform. Theory **49**, 446 (2003).
[21] A. Chefles and S.M. Barnett, J. Mod. Opt. **45**, 1295 (1998).
[22] C.W. Zhang, C.F. Li, and G.C. Guo, Phys. Lett. A **261**, 25 (1999).
[23] J. Fiurášek and M. Ježek, Phys. Rev. A **67**, 012321 (2003).
[24] L. Vandenberghe and S. Boyd, SIAM Rev. **38**, 40 (1996).
[25] F. Alizadeh, Ph.D. thesis, University of Minnesota, Minneapolis, MN, 1991.
[26] F. Alizadeh, in *Advances in Optimization and Parallel Computing*, edited by P. Pardalos (North-Holland, Amsterdam, 1992).
[27] Y. Nesterov and A. Nemirovski, *Interior-Point Polynomial Algorithms in Convex Programming* (SIAM, Philadelphia, 1994).
[28] Y.C. Eldar and H. Bölcskei, IEEE Trans. Inform. Theory (to be published), e-print math.FA/0108096.
[29] G.H. Golub and C.F.V. Loan, *Matrix Computations*, 3rd ed. (Johns Hopkins University Press, Baltimore, MD, 1996).
[30] M. Ježek, J. Řeháček, and J. Fiurášek, Phys. Rev. A **65**, 060301 (2002).
[31] A.C. Doherty, P.A. Parrilo, and F.M. Spedalieri, Phys. Rev. Lett. **88**, 187904 (2002).
[32] E.M. Rains, IEEE Trans. Inform. Theory **47**, 2921 (2001).
[33] K. Audenaert and B.D. Moor, Phys. Rev. A **65**, 030302 (2002).
[34] A.S. Holevo, IEEE Trans. Inform. Theory **44**, 269 (1998).
[35] P. Hausladen and W.K. Wootters, J. Mod. Opt. **41**, 2385 (1994).
[36] P. Hausladen, R. Josza, B. Schumacher, M. Westmoreland, and W.K. Wootters, Phys. Rev. A **54**, 1869 (1996).
[37] M. Sasaki, K. Kato, M. Izutsu, and O. Hirota, Phys. Rev. A **58**, 146 (1998).
[38] M. Sasaki, T. Sasaki-Usuda, M. Izutsu, and O. Hirota, Phys. Rev. A **58**, 159 (1998).
[39] K. Kato, M. Osaki, M. Sasaki, and O. Hirota, IEEE Trans. Commun. **47**, 248 (1999).
[40] G.D. Forney, Jr., IEEE Trans. Inform. Theory **37**, 1241 (1991).
[41] M.A. Armstrong, *Groups and Symmetry* (Springer-Verlag, New York, 1988).
[42] D.G. Luenberger, *Optimization by Vector Space Methods* (Wiley, New York, 1968).
[43] Otherwise we can transform the problem to a problem equivalent to the one considered in this paper by reformulating the problem on the subspace spanned by the eigenvectors of $\{\rho_i, 1 \leq i \leq m\}$.
[44] Interior point methods are iterative algorithms that terminate once a prespecified accuracy has been reached. A worst-case analysis of interior point methods shows that the effort required to solve a semidefinite program to a given accuracy grows no faster than a polynomial of the problem size. In practice, the algorithms behave much better than predicted by the worst case analysis, and in fact in many cases the number of iterations is almost constant in the size of the problem.
[45] A set C is convex if for any $x, y \in C$, $\alpha x + (1-\alpha)y \in C$ for all $\alpha \in [0,1]$.