Minimum error discrimination between similarity-transformed quantum states

M. A. Jafarizadeh,1,2,3,† R. Sufiani,1,2,† and Y. Mazhari Khiavi1‡
1Department of Theoretical Physics and Astrophysics, University of Tabriz, Tabriz 51664, Iran
2Institute for Studies in Theoretical Physics and Mathematics, Tehran 19395-1795, Iran
3Research Institute for Fundamental Sciences, Tabriz 51664, Iran

(Received 5 August 2010; revised manuscript received 18 March 2011; published 1 July 2011)

Using the well-known necessary and sufficient conditions for minimum error discrimination (MED), we extract an equivalent form for the MED conditions. In fact, by replacing the inequalities corresponding to the MED conditions with an equivalent but more suitable and convenient identity, the problem of mixed state discrimination with optimal success probability is solved. Moreover, we show that the mentioned optimality conditions can be viewed as a Helstrom family of ensembles under some circumstances. Using the given identity, MED between N similarity transformed equiprobable quantum states is investigated. In the case that the unitary operators are generating a set of irreducible representation, the optimal set of measurements and corresponding maximum success probability of discrimination can be determined precisely. In particular, it is shown that for equiprobable pure states, the optimal measurement strategy is the square-root measurement (SRM), whereas for the mixed states, SRM is not optimal. In the case that the unitary operators are reducible, there is no closed-form formula in the general case, but the procedure can be applied in each case in accordance to that case. Finally, we give the maximum success probability of optimal discrimination for some important examples of mixed quantum states, such as generalized Bloch sphere m-qubit states, spin-j states, particular nonsymmetric qudit states, etc.

DOI: 10.1103/PhysRevA.84.012102 PACS number(s): 03.65.Ta, 03.67.Mn

I. INTRODUCTION

Although quantum state discrimination is an important and challenging problem in the field of quantum information theory, the optimal minimum error discrimination is known for a few cases. Both the necessary and sufficient conditions for optimal discrimination [1–5] and a Helstrom family of ensembles [6] are two known strategies for treating minimum error discrimination. However, by means of them, except for a few cases, many problems are solved difficulty. In recent years, particular attention has been paid to quantum states in which the states have symmetry property called geometric uniformity [7–11]. Here we have presented a technique (based on the necessary and sufficient conditions for optimal discrimination [5]) that comprises two previous strategies in which optimality conditions are given in equality form. The introduced optimality condition is powerful for solving problems of optimal discrimination between mixed quantum states that are not, in general, symmetric. In this paper, we first give a simplified reformulation for optimality conditions by introducing an operator denoted by \( \mathcal{M} \), and then we prove equivalence between optimality conditions and a Helstrom family of ensembles under some circumstances. Then, by using the introduced formulation, we study minimum error discrimination (MED) between \( N \) equiprobable quantum mixed states generated from a density operator \( \rho_1 \) as \( \rho_i = U_i \rho_1 U_i^{-1} \), \( i = 1, \ldots, N \), where \( U_i \)'s are unitary operators. In this work, the quantum states to be discriminated are not symmetric states but similarity transformed quantum states. In the case that the unitary operators \( U_i \) are irreducible representations of generators of a subgroup of unitary group \( U(d) \), it is proved that the operator \( \mathcal{M} \) is a multiple of the identity operator and the maximum success probability and optimal measurement operators (POM) are precisely derived. In the case that the generators \( U_i \) are reducible representations of the corresponding subgroup, it is shown that \( \mathcal{M} \) is diagonal and, although there is no closed-form formula in the general case, the procedure can be applied in each case in accordance with that case. As an another result, it is shown that, although the square root measurement (SRM) is an optimal measurement strategy for MED between \( N \) equiprobable “pure” states, it is not optimal for MED between “mixed” states. Finally, by using the introduced technique, we study MED between some important classes of mixed quantum states, such as generalized Bloch sphere m-qubit states, spin-j states, particular nonsymmetric qudit states, etc., in detail.

II. MINIMUM ERROR DISCRIMINATION BETWEEN QUANTUM STATES AND A HELSTROM FAMILY OF ENSEMBLES

In general, the measurement strategy is described in terms of a set of non-negative definite operators called the probability operator measure (POM). The measurement outcome labeled by \( i \) is associated with the element \( \Pi_i \) of POM that has all the eigenvalues either positive or zero. The POM elements must add up to the identity operator, i.e., \( \sum_i \Pi_i = I \). Assume that a quantum system can be described by a state from different states \( \rho_1, \rho_2, \ldots, \rho_N \) with the prior probabilities \( p_1, p_2, \ldots, p_N \), respectively (\( p_i \geq 0 \), \( \sum_i p_i = 1 \)). If the system has been prepared in the state \( \rho_i \), then the probability to infer the system being state \( \rho_i \) is \( p(i|j) = \text{Tr}(\rho_i \Pi_j) \) and the probability for correctly identifying states \( \rho_i \) is given by

\[
\rho_{\text{corr}} = 1 - p_{\text{err}} = \sum_{i=1}^{N} p_i \text{Tr}(\rho_i \Pi_i).
\]
where $p_{\text{err}}$ is the error probability. For a set of measurement operators to maximize the probability of correct detection, it must satisfy a known set of necessary and sufficient conditions \cite{4},

$$
\sum_{i=1}^{N} p_i \rho_i \Pi_i - p_j \rho_j \geq 0, \quad j = 1, \ldots, N. \tag{2}
$$

As shown in Ref. \cite{4}, the conditions (2) lead to

$$
\left( \sum_{i=1}^{N} p_i \rho_i \Pi_i - p_j \rho_j \right) \Pi_j = \Pi_j \left( \sum_{i=1}^{N} p_i \rho_i \Pi_i - p_j \rho_j \right) = 0
$$

for any $j = 1, \ldots, N$. In the following, we use the fact that the inequality (2) indicates

$$
\sum_{i=1}^{N} p_i \rho_i \Pi_i - p_j \rho_j = \alpha_j \tau_j, \quad j = 1, \ldots, N, \tag{4}
$$

where $\alpha_j \geq 0$ and $\tau_j$'s are positive operators with $\text{Tr}(\tau_j) = 1$. Now, by taking the trace of both sides of Eq. (4), we obtain

$$
\alpha_j = p_{\text{opt}} - p_j. \tag{5}
$$

If we postmultiply Eq. (4) by $\Pi_j$ and sum up over $j$, we get $\sum_{j=1}^{N}(p_{\text{opt}} - p_j)\tau_j \Pi_j = 0$. Because both $\tau_j$ and $\Pi_j$ are positive operators, it follows that $(p_{\text{opt}} - p_j)\tau_j \Pi_j = 0$ for every $j = 1, \ldots, N$. For $p_{\text{opt}} > p_j$, this indicates that

$$
\tau_j \Pi_j = 0, \quad \text{for all } j = 1, \ldots, N, \tag{6}
$$

which is equivalent to the orthogonality condition (3). In this case, in order that the optimal measurement operators $\Pi_j$ can be constructed, the states $\tau_j$ must possess at least one zero eigenvalue. We note that, with the assumption $p_{\text{opt}} > p_j$, we obtained Eq. (5), which is critical for the derivation of optimal $\Pi_j$ in the subsequent sections. However, in some cases, it is known that $p_{\text{opt}} = p_m$ for some $m$; see Ref. \cite{3} for such examples. In this case, the relation (4) implies that $\alpha_m = p_{\text{opt}} - p_m = 0$, so that we have

$$
\sum_{i=1}^{N} p_i \rho_i \Pi_i = p_m \rho_m. \tag{6}
$$

By rewriting Eq. (6) as

$$
\sum_{i=1, i \neq m}^{N} p_i \rho_i \Pi_i = p_m \rho_m (I - \Pi_m) = p_m \rho_m \sum_{i=1, i \neq m}^{N} \Pi_i,
$$

we obtain

$$
\sum_{i=1, i \neq m}^{N} (p_m \rho_m - p_i \rho_i) \Pi_i = 0. \tag{7}
$$

From Eqs. (2) and (6), we know that $p_m \rho_m - p_i \rho_i \geq 0$ and so Eq. (7) leads to

$$
(p_m \rho_m - p_i \rho_i) \Pi_i = 0 \quad \text{for all } i = 1, \ldots, N.
$$

For the case $p_m \rho_m - p_i \rho_i \neq 0$, we have $\Pi_i = 0$ for all $i \neq m$ and $\Pi_m = I$. If $p_m \rho_m - p_i \rho_i = 0$ for some $i \neq m$, we have $p_m = p_i$ (this is seen by taking the trace of both sides of the equality) and so $\rho_m = \rho_i$. Then, Eq. (6) gives $\Pi_m + \Pi_i = I$ and $\Pi_j = 0$ for all $j \neq m, i$. We will assume that all the states $\rho_i$ are distinct.

From Eq. (4), the necessary and sufficient conditions for realizing a minimum error discrimination can be rewritten as follows:

$$
\mathcal{M} = p_j \rho_j + (p_{\text{opt}} - p_j)\tau_j, \quad j = 1, \ldots, N, \tag{8}
$$

where

$$
\mathcal{M} = \sum_{i=1}^{N} p_i \rho_i \Pi_i. \tag{9}
$$

The main purpose of the rest of the paper is solving the optimality conditions (8), i.e., finding suitable positive $\tau_j$ with $\text{Tr}(\tau_j) = 1$ and corresponding optimal measurement operators $\Pi_j$, so that the equalities (8) and the constraint (5) are satisfied.

In the following section, we show that the conditions (8) together with the constraint (5) are necessary and sufficient conditions for having a so-called Helstrom family of ensembles.

A. Equivalence between optimality conditions and a Helstrom family of ensembles

First let us recall the definition of a Helstrom family of ensembles. A set of $N$ numbers $\{\tilde{\rho}_1, \rho_1; 1 - \tilde{\rho}_1, \tau_1\}_{i=1}^{N}$ is called a weak Helstrom family (of ensembles) if there exist $N$ numbers of binary probability discriminations $\{\tilde{\rho}_1, 1 - \tilde{\rho}_1\}_{i=1}^{N} (0 < \tilde{\rho}_1 \leq 1)$ and states $\{\tau_1, s\}_{i=1}^{N}$ satisfying

$$
p = \frac{p_1}{\tilde{\rho}_1} \leq 1 \tag{10}
$$

and

$$
p_1 \rho_1 + (p - p_1)\tau_1 = p_j \rho_j + (p - p_1)\tau_j, \tag{11}
$$

for any $i, j = 1, \ldots, N$, where $p$ and $\tau_1$ are called Helstrom ratio and conjugate state to $\rho_1$, respectively \cite{6}. We assume that a prior probability distribution satisfies $p_i \neq 0, 1$ in order to remove trivial cases. We have \cite{6}

$$
p_{\text{opt}} \leq p. \tag{12}
$$

The observables $\{\Pi_i\}_{i=1}^{N}$ satisfy $p_{\text{opt}} = p$ if $(p - p_1)\text{Tr}(\tau_1 \Pi_i) = 0$, $i = 1, \ldots, N$. In this case, the observables $\{\Pi_i\}_{i=1}^{N}$ give an optimal measurement to discrimination of states $\{\rho_i\}_{i=1}^{N}$, and we call the family $\{\tilde{\rho}_1, \rho_1; 1 - \tilde{\rho}_1, \tau_1\}_{i=1}^{N}$ Helstrom family of ensembles \cite{6}. Note that Eq. (11) with $p = p_1 = p_j, i \neq j$ implies $\rho_1 = \rho_j$, which contradicts the assumption that all of the states are different from each other. Therefore, we conclude that if prior probabilities are equal, i.e., $p_i = \frac{1}{N}, i = 1, \ldots, N$, then $p_{\text{opt}} > \frac{1}{N}$. Now we show the equivalence between optimality conditions (8) and a Helstrom family of ensembles. To this end, first we use the fact that the condition $\text{Tr}(\tau_1 \Pi_i) = 0$ is equivalent to $\tau_1 \Pi_i = 0$ (for a proof, see Appendix A). To see that a Helstrom family of ensembles is sufficient to minimize the error, let us postmultiply Eq. (11) by $\Pi_i$ to obtain

$$
p_1 \rho_1 \Pi_i + (p - p_1)\tau_1 \Pi_i = p_j \rho_j \Pi_i + (p - p_1)\tau_j \Pi_i. \tag{13}
$$

Then, by considering the optimality condition $\tau_1 \Pi_i = 0$ for all $i$, Eq. (13) implies that $p_1 \rho_1 \Pi_i = p_j \rho_j \Pi_i + (p_{\text{opt}} - p_j)\tau_j \Pi_i$.\n
012102-2
Summing up over $i$ and using Eq. (9) and the completeness condition $\sum_{i=1}^{N} \Pi_i = I$ for probability operators, we obtain
\[
\mathcal{M} = p_j \rho_j + (p_{\text{opt}} - p_j) \tau_j, \quad j = 1, \ldots, N,
\]
which is the same as condition (8). To show that a Helstrom family of ensembles is a necessary condition, it is enough to subtract Eq. (8) for indices $i$ and $j$, which immediately leads to Eq. (11).

III. MINIMUM ERROR DISCRIMINATION BETWEEN SIMILARITY TRANSFORMED QUANTUM STATES

Let $\{U_1 = I_d, U_2, \ldots, U_N\}$ be a set of $d \times d$ unitary operators. Now we consider the equiprobable states $\rho_i$ as
\[
\rho_i = U_i \rho_i U_i^{-1}, \quad i = 1, \ldots, N,
\]
which are called similarity transformed states throughout the paper. Then, we choose the corresponding conjugate states $\tau_i$ as
\[
\tau_i = U_i \tau_i U_i^{-1}, \quad i = 1, \ldots, N,
\]
where $\tau_i$ is a trace one positive operator, which is determined in such a way that the optimality condition (8) is satisfied (as illustrated in the following). Substituting Eqs. (15) and (16) in Eq. (8) gives
\[
\mathcal{M} = U_i \left[ \frac{1}{N} \rho_i + \left( p - \frac{1}{N} \right) \tau_i \right] U_i^{-1} = U_i \mathcal{M} U_i^{-1}, \quad i = 1, \ldots, N.
\]
This indicates that $\mathcal{M}$ commutes with $U_i$, and so commutes with all elements of the subgroup $G$ generated by $U_i$, i.e., we have
\[
U(g) \mathcal{M} = \mathcal{M} U(g), \quad \forall \ g \in G,
\]
where $U$ is an unitary representation of $G$. As shown in Appendix B, the commutativity relation (18) leads to the fact that $\mathcal{M}$ can be written in diagonal form by choosing an appropriate basis under which $U$ has block diagonal form; particularly, $\mathcal{M}$ is proportional to the identity matrix in the case that $U$ is an irreducible representation. We will assume that all $U(g)$ are already block-diagonal in the specific basis (so that $M$ is taken diagonal) and use these properties of $\mathcal{M}$ in solving the optimality condition (8).

In order to $\Pi_i$’s form a set of POM, we need to have the completeness condition $\sum_{i=1}^{N} \Pi_i = I$. Assume that $\Pi_i = \lambda_i \Pi_i$ with $\sum_{i=1}^{N} \lambda_i = 1$ for some non-negative numbers $\lambda_i$ and positive semidefinite operators $\Pi_i$. In order that $\Pi_i$’s satisfy the optimality condition $\lambda_i \lambda_i \text{Tr}(\tau_i \Pi_i) = 0$, the optimal operators $\Pi_i$ can be assumed to also be obtained from a positive operator $U_i$ via the same similarity transform that defines the states $\rho_i$ and the corresponding conjugate states $\tau_i$, i.e.,
\[
\Pi_i = U_i \Pi_i U_i^{-1}, \quad i = 1, 2, \ldots, N.
\]
Therefore, it is sufficient to choose $\Pi_i$ perpendicular to $\tau_i$ in order to have $\Pi_i \tau_i = 0$ for all $i$. Then, we must have
\[
\sum_{i=1}^{N} \Pi_i = \sum_{i=1}^{N} \lambda_i \Pi_i = I.
\]
That is, the convex hull of the operators $\Pi_i$, $i = 1, \ldots, N$ must conclude the identity operator. Also, the completeness condition (20) leads to
\[
\text{Tr}(\Pi_i) = d,
\]
where we have used the fact that $\text{Tr}(\Pi_i) = \text{Tr}(U_i \Pi_i U_i^{-1} - \text{Tr}(\Pi_i) = \text{Tr}(\tau_i \Pi_i)$. This result shows that if any set of suitable coefficients $\lambda_i$ are determined in a way that $\sum_{i=1}^{N} \lambda_i = 1$ and Eq. (20) are satisfied, we can determine the corresponding set of optimal measurement operators $\{\Pi_i = \lambda_i \Pi_i\}_{i=1}^{N}$. Therefore, in general, the optimal measurement set $\{\Pi_i\}$ satisfying the optimality condition (8) is not unique.

A. Irreducible case

Assume that $U_i$, $i = 1, 2, \ldots, N$ are generating a set of an irreducible representation of a subgroup of $U(d)$. Then, as a consequence of the Schur’s first lemma from representation theory (see Appendix B), the only operator which can be invariant under the action of representation $U_i$ is a multiple of the identity operator. Therefore, Eq. (18) implies that $\mathcal{M} = \mu I$, and so the relation (17) for $i = 1$ is written as the follows:
\[
\mathcal{M} = \mu I = \frac{1}{N} \rho_i + \left( p_{\text{opt}} - \frac{1}{N} \right) \tau_i, \quad i = 1, \ldots, N.
\]
Taking the trace of both sides of Eq. (22), we find $\mu = \frac{p_{\text{opt}}}{d}$. Now, in order to consider optimal discrimination between the states $\rho_i$ in Eq. (15), we assume that $\rho_i = \sum_{i=1}^{d} a_i |i^{(1)} \rangle \langle i^{(1)}|$ is mixed state. Then, the resolution of identity (22) implies that $\tau_1$ is also diagonal in the bases $|i^{(1)}\rangle$. The state $\tau_1$ can be written as $\tau_1 = \sum_{i=1}^{d} b_i |i^{(1)} \rangle \langle i^{(1)}|$. Then, by using Eq. (22), we obtain
\[
p_{\text{opt}} = \frac{d}{N} a_i.
\]
Since at least one of the coefficients $b_i$, say $b_1$, is zero, the above relation leads to the following result:
\[
p_{\text{opt}} = \frac{d}{N} a_1.
\]
On the other hand, assuming that the largest eigenvalue of $\rho_1$ is $a_k \equiv a_{\text{max}}$, we have
\[
p_{\text{opt}} = \frac{d}{N} (a_{\text{max}} + (p_{\text{opt}} - 1) b_k),
\]
where $b_k$ is the coefficient of $\tau_1$ corresponding to the eigenstate $|k^{(1)}\rangle$ with the largest eigenvalue. Comparing Eqs. (24) and (25) and using the facts that $p_{\text{opt}} \geq \frac{d}{N}$, $b_k \geq 0$, and $a_{\text{max}} \geq a_1$, we obtain $b_k = 0$ and so
\[
p_{\text{opt}} = \frac{d}{N} a_{\text{max}}.
\]
That is the eigenstate of $\rho_1$ with largest eigenvalue is the eigenstate of $\tau_1$ with zero eigenvalue. It should be noticed that, in the case that all $a_i$’s are distinct, only one of the coefficients $b_i$, say $b_1$, must be zero so that $\Pi_1$ will be
given as $\Pi'_i = d\rho^i \rho^i |l_i^{(1)} \rangle \langle l_i^{(1)}|$ and, consequently, the elements of measurement are pure, i.e.,

$$\Pi'_i = dU_i |l_i^{(1)} \rangle \langle l_i^{(1)}|U_i^{-1}. \quad (27)$$

Now, let there be $m$ independent eigenvectors of $\rho_1$ having the same maximum eigenvalue $a_{\text{max}}$. Then, $m$ eigenvalues of $\tau_1$ must be zero. Denoting these eigenvalues by $b_1, \ldots, b_m$, the operator $\Pi'_1$ will be written as

$$\Pi'_1 = \alpha_1 |l_1^{(1)} \rangle \langle l_1^{(1)}| + \cdots + \alpha_m |l_m^{(1)} \rangle \langle l_m^{(1)}| = \sum_{k=1}^m \alpha_k |l_k^{(1)} \rangle \langle l_k^{(1)}|,$$

(28)

where $\alpha_k$’s are arbitrary non-negative numbers with $\sum_{k=1}^m \alpha_k = d$.

In order to check that the $M = \frac{dN}{d} I$ in Eq. (22) is the same as $M$ in Eq. (9), we calculate $\sum_{i=1}^N p_i \rho_i \Pi_i$ explicitly. By using Eq. (28) and the fact that $\Pi'_1$ is the projection operator to the eigenspace of $\rho_1$ corresponding to the maximum eigenvalue $a_{\text{max}}$, we have

$$\rho_1 \Pi'_1 = a_{\text{max}} \Pi'_1.$$

Then, by substituting $\Pi_i = \lambda_i \rho_i \Pi'_i \rho_i^*$, we obtain

$$\sum_{i=1}^N p_i \rho_i \Pi_i = \frac{1}{N} \sum_{i=1}^N \lambda_i \rho_i \Pi'_i \rho_i^* = \frac{a_{\text{max}}}{N} \sum_{i=1}^N \lambda_i \Pi'_i = \frac{a_{\text{max}}}{N} I = \frac{\text{opt} \ d}{d} I = M,$$

where, in the last equality, we have used Eq. (26).

1. **Group covariant quantum states**

In the special case in which the states $\rho_i$ are given by $\rho_i = U(g) \rho_0 U(g)^{-1}$ and $U(N,g) = 1$, where $U$ is an irreducible unitary representation of a group $G$ with order $N$, the set of optimal POM $\{\Pi'_{g}, g \in G\}$ is determined in terms of $\{\Pi'_g = U(g) \Pi'_1 U(g)^{-1}, g \in G\}$. These states are sometimes called symmetric states or group covariant quantum states. By means of Schur’s first lemma, it can result that the irreducible representation associated with the group elements belonging to the center is multiple of identity matrix; that is, we have $U(g) = e^{i \Phi(g)} I$ for all $g \in Z$. Therefore, one can consider the groups $G$ with trivial center $Z = \{e\}$ or, in the case of groups with nontrivial center, the quotient group $G/Z$ instead of $G$ and parametrize the initial states $\rho_q$ and POM set $\Pi'_q$, with elements of $G/Z$.

One should notice that, in this case, a natural choice for the coefficients $\lambda_k \equiv \lambda_i$ in the resolution of identity (20) is $\lambda_1 = \frac{1}{|G|}$, so that $\Pi'_1$ can be chosen as $\Pi'_q = \frac{1}{|G|} \Pi'_q$. To see this, we denote $\sum_{g \in G} \Pi'_g$ by $\Pi$. Then, one can write

$$U(g') \Pi U(g'^{-1})^{-1} = \sum_{g \in G} U(g') \Pi(g) U(g'^{-1})^{-1} = \sum_{g \in G} U(g') \Pi'_{g} U(g'^{-1})^{-1} \quad (29)$$

Due to the ineducability of group representation $U$, Schur’s lemma implies that $\Pi$ must be multiple of identity operator, i.e., we have

$$\Pi = \sum_{g \in G} \Pi'_g = \alpha I.$$

Taking the trace of both sides and using Eq. (21), we obtain

$$\alpha = |G|. \quad (30)$$

Therefore, $\frac{1}{|G|} \sum_{g \in G} \Pi'_g = I$, so that we can choose $\Pi'_g = \frac{1}{|G|} U(g) \Pi'_1 U(g)^{-1}$, where $\Pi'_1$ is determined as illustrated in the previous section [Eq. (28)].

B. Reducible case

In the case that $U_i, i = 1, \ldots, N$ are generating a set of a reducible representation, invariance of $M$ under the action of $U_i$ [i.e., Eq. (18)] implies that $M$ is diagonal but not, in general, proportional to the identity matrix, i.e., we have $M = \text{diag}(M_1, \ldots, M_d)$ (see Appendix B).

Then, we can choose the basis on which the state $\rho_1$ be diagonal, so that the identity (14) will lead to the fact that $\tau_1$ is also diagonal. In this case, we do not give a general solution for optimal discrimination between similarity transformed states (15), but as will be seen in some important examples of particular $n$-qubit and qudit mixed states, MED can be achieved in each case via the same technique illustrated above.

IV. **Optimality of the Square-Root Measurement (SRM) for Pure States**

A measurement that has been employed as a detection measurement in many applications is SRM [13–15], as follows:

$$\Pi_j = \frac{1}{N} \Phi^{-1} \rho_j \Phi^{-1} \quad (32)$$

where $\Phi = \frac{1}{N} \sum_{i=1}^N \rho_i$. The SRM operators $\Pi_j$ are positive and sum to the identity. For pure and mixed state ensembles that exhibit certain symmetries, SRM minimizes the probability of a detection error. We now show that a measurement is SRM if the equiprobable stats (15) are pure. We first show that $\Phi$ commutes with each of the operators $U_j$. We have that, for all $j$,

$$\Phi U_j = \sum_{i=1}^N p_i U_i \rho_i U_i^{-1} U_j = U_j \sum_{i=1}^N p_i U_i^{-1} U_i \rho_i U_i^{-1} U_j \quad (33)$$

$$= U_j \sum_{k=1}^N p_i U_k \rho_i U_k^{-1} = U_j \Phi.$$

012102-4
Because $\Phi$ commutes with $U_j$, Schur’s lemma implies that $\Phi = \delta I$. Thus $\delta = \frac{1}{d}$ and
\begin{equation}
\Pi_j = \frac{d}{N} \rho_j.
\end{equation}

The probability of correct detection is
\begin{align}
p &= \frac{d}{N^2} N \sum_{i=1}^{N} \text{Tr}(\rho_i \Pi_i) = \frac{d}{N^2} N \sum_{i=1}^{N} \text{Tr}(\rho_i^2) \\
&= \frac{d}{N^2} N \sum_{i=1}^{N} \text{Tr}(U_i \rho_i U_i^*-U_i \rho_i U_i^{-1}) = \frac{d}{N} \text{Tr}(\rho_i^2).
\end{align}

Since $\frac{1}{N} [\delta I - \text{Tr}(\rho_i^2)] \geq 0$, it follows that if the states $\rho_i$ are nonpure, SRM is not optimal. In the case where the states are pure, the two relations (26) and (35) coincide because we have $\alpha_{\text{max}} = 1$ and $\text{Tr}(\rho_i^2) = 1$. Then this implies that, for pure-state ensembles, the SRM is optimal.

We now consider equiprobable pure states that have the form
\begin{equation}
|\psi_{jk}\rangle = U^{k-1} V^{j-1} |\psi_{11}\rangle, \quad |\psi_{11}\rangle = \begin{pmatrix} \cos \theta & 0 \\ 0 & \sin \theta \end{pmatrix},
\end{equation}
where $j = 1, \ldots, n$, $k = 1, \ldots, m$, and $V^n = U^m = I$, so that the number of states is $N = nm$ and we have $\Phi = \frac{1}{nm} \sum_{j=1}^{n} \sum_{k=1}^{m} |\psi_{jk}\rangle \langle \psi_{jk}|$. Then, the SRM elements for realizing a minimum error measurement are given by [4]
\begin{equation}
\Pi_{jk} = \frac{1}{nm} \Phi^{-\frac{1}{2}} |\psi_{jk}\rangle \langle \psi_{jk}| \Phi^{-\frac{1}{2}}.
\end{equation}

The associated optimal success probability is then
\begin{align}
p_{\text{opt}} &= \frac{1}{nm} \sum_{j=1}^{n} \sum_{k=1}^{m} |\psi_{jk}\rangle \langle \psi_{jk}| \Pi_{jk} |\psi_{jk}\rangle \\
&= \frac{1}{(nm)^2} \sum_{j=1}^{n} \sum_{k=1}^{m} |\psi_{jk}\rangle \langle \psi_{jk}| \Phi^{-\frac{1}{2}} |\psi_{jk}\rangle \langle \psi_{jk}| \Phi^{-\frac{1}{2}}.
\end{align}

**Dihedral group $D_{2n}$**. For instance, consider the dihedral group $D_{2n}$ of order $2n$, which is defined as
\begin{equation}
D_{2n} = \{a, b : a^n = b^2 = 1, b^{-1}ab = a^{-1}\}.
\end{equation}

Then, $V$ and $U$ can be considered respectively as a two-dimensional irreducible representation of generators $a$ and $b$ of the group [16] as follows:
\begin{equation}
V = \begin{pmatrix} e^{\frac{2\pi i}{2n}} & 0 \\ 0 & e^{\frac{-2\pi i}{2n}} \end{pmatrix}, \quad U = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},
\end{equation}
where $l$ is integer with $1 \leq l < \frac{n-1}{2}$. Then, for states (36) with $U$ and $V$ given in Eq. (40), one can find $\Phi = \frac{1}{2} I_2$. Then, by using Eq. (38), it follows that
\begin{equation}
p_{\text{opt}} = \frac{2}{N} = \frac{1}{n},
\end{equation}
which coincides with the result (35), since for pure states $\text{Tr}(\rho_i^2) = 1$.

**Dicyclic group $T_{4n}$**. As another example of the equiprobable states in form (36), we consider the dicyclic group of order $4n$ defined by [16]
\begin{equation}
T_{4n} = \{a, b : a^{2n} = 1, a^n = b^2, b^{-1}ab = a^{-1}\},
\end{equation}
where, a two-dimensional irreducible representation of generators $a$ and $b$ is given by
\begin{equation}
V = \begin{pmatrix} e^{\frac{\pi i}{2n}} & 0 \\ 0 & e^{-\frac{\pi i}{2n}} \end{pmatrix}, \quad 1 \leq l \leq N, \quad U = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},
\end{equation}
respectively. Then, similarly, one can obtain $\Phi = \frac{1}{2} I_2$. By Eq. (38), it follows that the optimal success probability is given by $p_{\text{opt}} = \frac{1}{2} = \frac{1}{n} = \frac{1}{2}$. The probability of correct detection is
\begin{equation}
\text{Tr}(\rho_i^2) = \frac{1}{2}.
\end{equation}

**V. EXAMPLES OF MED BETWEEN SIMILARITY TRANSFORMED MIXED STATES**

**A. Generalized Bloch sphere $m$-qubit states (irreducible case)**

In this example, we consider particular $m$-qubit states in $d = 2^m$ dimensional Hilbert space, which possess properties similar to qubit density matrices represented in Bloch sphere, and so we call them generalized Bloch sphere states. Let
\begin{equation}
\rho_1 = \frac{1}{2^m} \left( I + a \sum_{i=1}^{2m+1} n_i \gamma_i \right) = \frac{1}{2^m} (I + a \hat{\gamma} \cdot \hat{\gamma}),
\end{equation}
where $\gamma_i$ for $i = 1, 2, \ldots, 2m + 1$ are generators of special orthogonal group $SO(2m + 1)$ and represented as traceless Hermitian matrices in a $2^m$-dimensional system. That is, $\gamma_i$ are a maximally anticommuting set of Hermitian traceless Dirac matrices, which satisfy
\begin{equation}
\{\gamma_i, \gamma_j\} = 2 \delta_{ij} I_d.
\end{equation}

For a brief review about Dirac matrices and an explicit construction of $\gamma_i$’s, one can refer to Ref. [17] or Appendix A of Ref. [18]. From the properties (45), one can easily see that $(\hat{\gamma} \cdot \hat{\gamma})^2 = I$ and so the eigenvalues of $\rho_1$ are $\frac{1 + \hat{\gamma} \cdot \hat{\gamma}}{2m}$. Therefore, the density matrix $\rho_1$ can be decomposed as
\begin{equation}
\rho_1 = \frac{1 + a}{2m} \left( I + \hat{\gamma} \cdot \hat{\gamma} \right) + \frac{1 - a}{2m} \left( I - \hat{\gamma} \cdot \hat{\gamma} \right),
\end{equation}
where $\frac{1 + \hat{\gamma} \cdot \hat{\gamma}}{2m}$ and $\frac{1 - \hat{\gamma} \cdot \hat{\gamma}}{2m}$ are respectively projection operators (idempotents) to the eigenstate subspaces corresponding to the eigenvalues $\frac{1 + \hat{\gamma} \cdot \hat{\gamma}}{2m}$ and $\frac{1 - \hat{\gamma} \cdot \hat{\gamma}}{2m}$. Then, according to the arguments of Sec III, we have
\begin{equation}
\tau_1 = \frac{1}{2m} (I - \hat{\gamma} \cdot \hat{\gamma})
\end{equation}
and so the optimal measurement operator $\Pi_1$, which is orthogonal to $\tau_1$, is given by
\begin{equation}
\Pi_1 = (I + \hat{\gamma} \cdot \hat{\gamma})
\end{equation}
and then
\begin{equation}
\Pi_1 = U_I (I + \hat{\gamma} \cdot \hat{\gamma}) U_I^{-1},
\end{equation}
where $U_I$’s are spinor representations of the group $SO(2m + 1)$,
namely \( U_i = e^{i \sum_{\gamma} \theta_i \gamma} \in \text{spin}(2m+1) \). It should be noted that \( U_i \)’s can be chosen also as spinor representation of any subgroup of \( \text{SO}(2m+1) \) of the same rank \( m \) (maximal subgroup), such as \( \text{SO}(i_1) \otimes \text{SO}(i_2) \otimes \cdots \otimes \text{SO}(i_t) \), where \( (i_1,i_2,\ldots,i_t) \) is an arbitrary partition of \( 2m+1 \) to \( t \) parts, i.e., \( 2m+1 = i_1 + i_2 + \cdots + i_t \).

It would also be noted that any measurement operator \( \Pi_i \) in Eq. (49) can be viewed as a point on a sphere with radius equal to 1 and so the resolution of identity \( \sum_{i=1}^{N} \lambda_i \Pi_i = I \) implies that the Bloch vectors corresponding to \( \Pi_i \) can not share the same hemisphere of the Bloch sphere, i.e., the Bloch vectors \( \hat{n}(i) \) associated with \( \Pi_i \) must be distributed in such a way that we have \( \sum_{i=1}^{N} \lambda_i \hat{n}(i) = 0 \).

Now, by using the result (26), we obtain

\[
p_{\text{opt}} = \frac{1 + a}{N},
\]

where the same result has been obtained in Ref. [19] via the Helstrom family of ensembles.

In the special case \( d = 2 \) \((m = 1)\), we have \( \rho_i = U_i \rho_1 U_i^{-1} \) with

\[
\rho_1 = \frac{1}{2}(I + a \hat{n} \cdot \hat{\sigma}) = \frac{1 + a}{2} \ket{+n} \bra{+n} + \frac{1 - a}{2} \ket{-n} \bra{-n},
\]

where \( \ket{+n} = \frac{1}{\sqrt{2}} \ket{\gamma_+} \) and \( \ket{-n} = \frac{1}{\sqrt{2}} \ket{\gamma_-} \) and the unitary operation \( U_i \) is a rotation operator with respect to the center of the Bloch sphere.

**B. Generalized Bloch sphere \( m \)-qubit states (reducible case)**

Now we consider the case that \( U_i \)’s are generating sets of a subgroup of \( \text{SO}(2m+1) \) or its maximal subgroup (any non-trivial subgroup with rank less than \( m \)). Then, we will have matrices other than the identity matrix, which can be invariant under the action of \( U_i \), and \( \mathcal{M} \) is not proportional to the identity matrix. In this case, we assume that the density matrix \( \rho_1 \) in Eq. (44) decomposes to two parts as follows:

\[
\rho_1 = \frac{1}{2m} \left( I + a \sum_{i \in S_I} n_i \gamma_i + a \sum_{i \in S_V} n_i \gamma_i \right),
\]

where

\[
S_I = \{i_1, \ldots, i_l\}, \quad S_V = \{i_{l+1}, \ldots, i_{2m+1}\}
\]

respectively denote the set of invariant and variant coefficients of the Bloch vector \( \hat{n} \).

Now, from the invariance of \( \mathcal{M} \) under the action of \( U_i \)’s [see Eq. (18)], we have

\[
\mathcal{M} = \alpha I_d + \sum_{k \in S_I} \beta_k \gamma_k.
\]

Then, we choose

\[
\tau_1 = \frac{1}{2m} \left( I - \sum_{i \in S_V} n_i' \gamma_i \right),
\]

and so

\[
\Pi_i' = U_i \left( I_d + \sum_{i \in S_V} n_i' \gamma_i \right) U_i^{-1}.
\]

Now, by using the optimality conditions (8), one can easily obtain

\[
\alpha = \frac{p_{\text{opt}}}{2m}, \quad \beta_k = \frac{aN_k}{N^2 2m}, \quad k \in S_I, \quad n_i' = \frac{aN_k}{1 - N p_{\text{opt}}}, \quad k \in S_V,
\]

and

\[
p_{\text{opt}} = \frac{1}{N} \left( 1 + a \sqrt{\sum_{k \in S_V} n_k'^2} \right).
\]

In the single qubit case \((m = 1)\), the result (58) leads to

\[
p_{\text{opt}} = \frac{1}{N} \left( 1 + a \sqrt{n_1^2 + n_2^2} \right) = \frac{1}{N} \left( 1 + a \sqrt{1 - n_1^2} \right)
\]

which is in agreement with the result derived in Ref. [19] for qubit states.

In order to check that the \( \mathcal{M} \) in Eq. (54) with \( \alpha, \beta \) given in Eq. (57) is the same as \( \mathcal{M} \) in Eq. (9), we calculate \( \sum_{i=1}^{N} \rho_i \Pi_i \) explicitly. To this end, we rewrite \( \rho_1 \) in Eq. (52) as

\[
\rho_1 = \frac{1}{2m} \left( I + a \sum_{i \in S_I} n_i \gamma_i + a \sum_{i \in S_V} n_i \gamma_i \right)
\]

\[
= \frac{1}{2m} \left( I + \tilde{\alpha} \sum_{i \in S_I} n_i' \gamma_i + \tilde{\beta} \sum_{i \in S_V} n_i \gamma_i \right)
\]

\[
= \frac{1 + \tilde{\alpha}}{2m} \left( I + \sum_{i \in S_I} n_i' \gamma_i \right) + \frac{1 - \tilde{\alpha}}{2m} \left( I - \sum_{i \in S_V} n_i' \gamma_i \right)
\]

\[
+ \frac{a}{2m} \sum_{i \in S_V} n_i \gamma_i,
\]

where

\[
\tilde{\alpha} = a \sqrt{\sum_{i \in S_I} n_i'^2} = N p_{\text{opt}} - 1, \quad n_i' = -\frac{n_i}{\sqrt{\sum_{i \in S_V} n_i'^2}}.
\]

Since we have

\[
\Pi_i' = I + \sum_{i \in S_V} n_i' \gamma_i,
\]

then

\[
\rho_1 \Pi_i' = \frac{1 + \tilde{\alpha}}{2m} \left( I + \sum_{i \in S_I} n_i' \gamma_i \right) + \frac{a}{2m} \left( \sum_{i \in S_I} n_i \gamma_i \right)
\]

\[
\times \left( I + \sum_{i \in S_V} n_i' \gamma_i \right)
\]

\[
= \frac{1}{2m} \left( 1 + \tilde{\alpha} \right) I + a \sum_{i \in S_I} n_i \gamma_i \Pi_i.
\]
Now we have
\[
\sum_{i=1}^{N} p_i \rho_i \Pi_i = \frac{1}{N} \sum_{i=1}^{N} \lambda_i U_i \rho_i \Pi'_i U_i^\dagger
\]
\[
= \frac{1}{N^{2m}} \left( (1 + \tilde{a}) I + a \sum_{i \in S_i} n_i \gamma_i \right) \sum_{j=1}^{N} \lambda_j \Pi'_j
\]
\[
= \frac{1}{N^{2m}} \left( (1 + \tilde{a}) I + a \sum_{i \in S_i} n_i \gamma_i \right)
\]
\[
= \frac{p_{\text{opt}}}{2m} I + \frac{1}{N^{2m}} a \sum_{i \in S_i} n_i \gamma_i = \mathcal{M},
\]
where, in the last line, we have used Eqs. (57), (58), and (60).

### C. Spin-\(j\) quantum states

Consider the Abelian subset of SO(2) of the rotation group \SO(3), generated by a rotation operator as \exp(−i\(\phi\) \(J_z\)), that rotates a spin-\(j\) state by \(\phi\) with respect to the \(\hat{k}\) axis. We suppose the states \(\rho_i\) and \(\tau_j\) have the forms \(\rho_i = \frac{1}{\sqrt{d}} (I + 2a \hat{n} \cdot \hat{J})\) and \(\tau_j = \frac{1}{\sqrt{d}} (I + 2b \hat{n} \cdot \hat{J})\). We can write \(\tau_j\) in diagonal representation as \(\tau_j = \sum_{m=-j}^{+j} \frac{1}{d} (1 + 2bm) |m\rangle \langle m|\) in which \(|m\rangle\)'s are complete orthonormal bases. Also, it should be noticed that \(\tau_j\) is not full rank and so its minimum eigenvalue is zero, which indicates that \(b = \sqrt{\frac{N}{d}}\). Since \(\tau_j\) and \(\Pi_i\) are orthogonal, we have the result
\[
\Pi'_i = dU_i |j\rangle \langle -j| U_i^{-1}.\tag{61}
\]
Invariance of \(\mathcal{M}\) under rotations about the \(z\) axis implies that it has the following form:
\[
\mathcal{M} = \alpha I_d + \beta J_z,
\tag{62}
\]
where the constants \(\alpha\) and \(\beta\) must be calculated. Equaling Eqs. (62) and (8) for \(i = 1\) results in
\[
\alpha = \frac{p_{\text{opt}}}{d}, \quad \beta = \frac{2an_x + (Np_{\text{opt}} - 1)bn'_x}{Nd},
\tag{63}
\]
\[
an_x + (Np_{\text{opt}} - 1)bn'_x = an_x + (Np_{\text{opt}} - 1)bn'_y = 0.
\]
From the above equation, we have \(\frac{n_x'}{n_x} = \frac{n_y}{n_x}\) or \(\cot \phi' = \cot \phi\), so that \(|\phi' - \phi| = \pi\). Therefore, we have \(n_x' = \sin \theta' \cos \phi' = -\sin \theta' \cos \phi\) and Eq. (63) leads to
\[
p_{\text{opt}} = \frac{1}{N} \left( 1 + \frac{a \sin \theta}{b \sin \theta'} \right).
\tag{64}
\]
The Bloch vectors corresponding to \(\lambda_i, \Pi'_i\) must lay in the \(\theta' = \frac{\pi}{2}\) plane in order that there exists a convex combination of them as Eq. (20). Then, substituting \(\theta' = \frac{\pi}{2}\) and \(b = \frac{1}{\sqrt{d-1}}\) into Eq. (64) immediately gives
\[
p_{\text{opt}} = \frac{1}{N} \left[ 1 + a(d - 1) \sin \theta \right],\tag{65}
\]
which coincides with Eq. (59) for the special case \(d = 2\) (i.e., qubits).

### D. Qudit states based on polarization operators

For now, we consider the particular case of quantum states generated from the density operator,
\[
\rho_1 = \frac{I}{d} + b_{LL} T_{LL} + b_{LO} T_{L0} + b_{L(-L)} T_{L(-L)},
\tag{66}
\]
where \(T_{(2s)\langle 2s\rangle}, T_{(2s)\langle 0\rangle}\), and \(T_{(2s)\langle -2s\rangle}\) are the polarization operators in the Hilbert-Schmidt space of dimension \(d\) defined as (given in [20])
\[
T_{(2s)\langle M\rangle} = \sqrt{\frac{4s+1}{2s+1}} \sum_{k,l=1}^{d} C_{smk}^{M} |k\rangle \langle l|, \quad m_1 = s,
\tag{67}
\]
\[
m_2 = s - 1, \ldots, m_d = -s.
\]
By using Clebsch-Gordan coefficients displayed explicitly in tables, e.g., in Ref. [21], the polarization operators can be readily calculated as
\[
T_{(2s)\langle 2s\rangle} = (-1)^{2s} |1\rangle \langle d|, \quad T_{(2s)\langle -2s\rangle} = |1\rangle \langle d|.
\tag{68}
\]
Then, the formula in Eq. (22) can be applied in order to obtain
\[
\begin{pmatrix}
p_{\text{opt}} = \frac{1}{N} \\
M_1 - \frac{1}{N} \left( \frac{1}{d} + \frac{b_{LL}(2s)!}{\sqrt{(4s)!}} \right) \quad 0 \ldots 0 \\
0 \quad \ddots \quad \ddots \quad \ddots \\
0 \quad 0 \quad \ddots \quad \ddots \quad \ddots \\
-\frac{(-1)^2 b_{LL}}{N} \quad 0 \ldots 0 \quad M_d - \frac{1}{N} \left( \frac{1}{d} + \frac{b_{LL}(2s)!}{\sqrt{(4s)!}} \right)
\end{pmatrix},
\tag{69}
\]
where \(b_{LL}^* = (-1)^M b_{L(-M)}\). Note that the positivity condition of \(\rho_i\)'s are counted as inequalities,
\[
\frac{1}{d} + \frac{b_{LL}(2s)!}{\sqrt{(4s)!}} - |b_{LL}| \geq 0,\tag{70}
\]
\[
\frac{1}{d} + \frac{(-1)^{3s+m_2} b_{L(2s)!}}{\sqrt{(4s)!}(s + m_2)!} - |b_{L(-M)}| \geq 0, \quad k = 1, \ldots, d.\tag{71}
\]
In addition, the positivity condition on eigenvalues of \(p_{\text{opt}} - \frac{1}{N} \tau_1\) and the relation \(\Pi_1 \tau_1 = 0\) is satisfied if
\[
\Pi_1' = \begin{pmatrix}
1 & 0 & \cdots & 0 & \frac{(-1)^2 b_{LL}}{|b_{LL}|} \\
0 & 1 & \cdots & 0 & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \cdots & 0 & 1 & 0 \\
\frac{(-1)^2 b_{LL}}{|b_{LL}|} & 0 & \cdots & 0 & 1
\end{pmatrix} = \begin{pmatrix}
1 & 0 & \cdots & 0 & -\frac{1}{\sqrt{d}} \\
0 & 1 & \cdots & 0 & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \cdots & 0 & 1 & 0 \\
-\frac{1}{\sqrt{d}} & 0 & \cdots & 0 & 1
\end{pmatrix} = \Pi_1',
\tag{72}
\]
and
\[ M_1 = - \frac{1}{N} \left( \frac{1}{d} + \frac{|b_{LL}|}{\sqrt{(4s)^2}} \right) - \frac{|b_{LL}|}{N} = 0, \quad M_1 = M_d, \quad (73) \]
\[ M_k = - \frac{1}{N} \left( \frac{1}{d} + \frac{(-1)^{3s+m} b_{LL}(2s!)}{\sqrt{(4s)^2}} \right) = 0, \quad k = 2, \ldots, d - 1. \quad (74) \]

Substituting the obtained \( \Pi_i \) into Eq. (19) and considering the condition (20), one obtains the following requirement on the set of unitary operators:
\[ \sum_{k=1}^{N} \lambda_k e^{i(\theta_k - \theta_\alpha)} = 0. \quad (75) \]

It yields that the maximum of \( |\theta_k - \theta_\alpha| \) must be greater than or equal to \( \pi \). We note that since \( p_{\text{opt}} = \sum_{k=1}^{d} M_k \), summing up Eq. (74) over \( k \) from 2 to \( d - 1 \) and then adding to Eq. (73) yields \( p_{\text{opt}} = \frac{1}{N} (1 + 2|b_{LL}|) \). The success probability in discriminating among these \( N \) states is then \( p_{\text{opt}} = \frac{1}{N} (1 + b \sin \theta) \).

E. Nonsymmetric qudit states

In this example, we generalize the optimal minimum-error discrimination among symmetric pure states reviewed in Ref. [8]. In a \( d \)-dimensional Hilbert space, the states \( \{ \rho_j \} \) for \( j = 0, \ldots, N - 1 \) that we wish to discriminate among are
\[ \rho_j = U_j \rho_0 U_j^{-1}, \quad \rho_0 = |\psi_1 \rangle \langle \psi_1| + |\psi_2 \rangle \langle \psi_2|, \quad \eta_1 + \eta_2 = 1, \quad |\psi_1 \rangle \langle \psi_1| = |\psi_2 \rangle \langle \psi_2| = 1 \quad (76) \]
where
\[ |\psi_1 \rangle = \sum_{k=1}^{m} c_k |k \rangle, \quad |\psi_2 \rangle = \sum_{k=d+1}^{d} c_k |k \rangle, \quad (77) \]
\[ U_j = \sum_{k=1}^{d} e^{i\theta_j} |k \rangle \langle k|, \]
with orthonormal basis \(|k \rangle\). The operators of the measurement can be written as follows:
\[ \Pi_k = |\mu_k \rangle \langle \mu_k| = U_j |\mu_0 \rangle \langle \mu_0| U_j^{-1}, \quad |\mu_0 \rangle = \sum_{k=1}^{d} \mu_{0k} |k \rangle. \quad (78) \]

Therefore, noting that \( p = \sum_{k=1}^{d} M_k \), we can express our problem as minimization of \( p \) subject to the constraints
\[ \text{Tr} \left[ \left( \mathcal{M} - \frac{1}{N} \rho_0 \right) |\mu_0 \rangle \langle \mu_0| \right] = 0, \quad \sum_{k=1}^{d} |\mu_{0k}|^2 = d. \quad (79) \]

The Lagrangian is \( L = \sum_{k=1}^{d} M_k + \lambda \left[ \sum_{k=1}^{d} M_k |\mu_{0k}|^2 - \frac{1}{N} (\eta_1 \sum_{k=1}^{m} c_k^2 |\mu_{0k}|^2 + \eta_2 \sum_{k=d+1}^{d} c_k^2 |\mu_{0k}|^2 + \nu (\sum_{k=1}^{d} |\mu_{0k}|^2 - d) \right] \). The vanishing gradient of \( L \) with respect to \( M_k \) and \( \mu_{0k}^* \) determines optimal \( \mu_{0k} \). We find that optimal coefficients \( \mu_{0k} \) do not depend on \( k \) and are proportional to \( c_k \). Thus
\[ \mu_{0k} = \frac{c_k}{|c_k|} = e^{i\phi_k}, \quad |\mu_j \rangle = \sum_{k=1}^{d} e^{i(\theta_j + \phi_k)} |k \rangle. \quad (80) \]

Since we have \( \Pi_j = \lambda_j \Pi_j \) and \( \sum_{j=1}^{N} \lambda_j = 1 \), it is straightforward to compute the maximum success probability as
\[ p_{\text{opt}} = \frac{1}{N} \sum_{j=0}^{N-1} \lambda_j \text{Tr}(\rho_j \Pi_j) = \frac{1}{N} \text{Tr}(\rho_0 \Pi_0) \]
\[ = \frac{1}{N} \left[ \eta_1 \left( \sum_{k=1}^{m} |c_k|^2 \right)^2 + \eta_2 \left( \sum_{k=m+1}^{d} |c_k|^2 \right)^2 \right]. \quad (81) \]

We can also generalize decomposition of generating state \( \rho_0 \) as a convex combination of more than two orthogonal rank-one states.

VI. CONCLUSION

Using the optimality conditions for quantum state discrimination and a Helstrom family of ensembles, a more suitable and convenient form for the MED conditions was extracted. By using this formulation, we studied MED among \( N \) equiprobable quantum states generated from a density operator by similarity transformations. In the case that the set of the unitary operators \( U_i \) are irreducible representations of the subgroup of \( U(d) \) generated by \( U_i \), we have precisely derived the maximum success probability and the corresponding optimal POM. For the case that \( U_i \)'s are reducible representations, we did not give a general formula but illustrated how the method can be applied to solve optimality conditions in each particular case. Moreover, it was shown that the SRM strategy is optimal only in the case that MED among pure states is considered and in the case of mixed states is not optimal. Finally, we discussed MED between some important classes of mixed quantum states, such as generalized Bloch sphere \( m \)-qubit states, spin-\( j \) states, etc. Although we have considered that the initial states have equal prior probability, the approach of the paper can be applied for optimal discrimination of a more general case in which the states are prepared with different prior probabilities. This case is under investigation. It seems that the presented approach can be used in investigating some other problems.

APPENDIX A

In this Appendix, we prove that the condition \( \text{Tr}(\tau_i \Pi_i) = 0 \) is equivalent to \( \tau_i \Pi_i = 0 \). To this end, we write the spectral decomposition of the operators \( \tau_i \) and \( \Pi_i \) as follows:
\[ \tau_i = \sum_{j=1}^{d} t_{ij} |\phi_j \rangle \langle \phi_j|, \quad \Pi_i = \sum_{k=1}^{d} \pi_{ik} |\psi_k \rangle \langle \psi_k|. \quad (A1) \]

Now it is obvious that
\[ \text{Tr}(\tau_i \Pi_i) = \sum_{j=1}^{d} \sum_{k=1}^{d} t_{ij} \pi_{ik} |\langle \phi_j| \psi_k \rangle|^2. \quad (A2) \]
Since $t_{ij}$ and $\pi_{ik}$ are non-negative, if $\text{Tr}(\tau_i \Pi_i) = 0$, then $t_{ij} \pi_{ik} \langle \phi_j | \psi_k \rangle = 0$, i.e., $\tau_i \Pi_i = 0$. Conversely, if $\tau_i \Pi_i = 0$, then $\text{Tr}(\tau_i \Pi_i) = 0$ holds trivially.

APPENDIX B

In this Appendix, we prove that the invariance property of $\mathcal{M}$, i.e., Eq. (18), induces that $\mathcal{M}$ is diagonal. To see this, we first consider that the representation $U$ is reducible, one can decompose it as irreducible components $V_i$ (see Ref. [23] for a proof). In other words, one can find an appropriate basis under which $U$ has block-diagonal form, i.e., $U = TU'T^{-1}$, with $U'$ block-diagonal and $T$ unitary. Substituting $U = TU'T^{-1}$ into Eq. (18), we have $TU'T^{-1}M = MTU'T^{-1}$, which leads to $U'(T^{-1}MT) = (T^{-1}MT)U'$. Now we show that $U'$ is a direct sum of two nonequivalent irreducible representations: $n_1 \times n_1$ matrix $V_1$ and $n_2 \times n_2$ matrix $V_2$. If we denote $\mathcal{M}' = \{A_{ij}, B_{ij}; C_{ij}, D_{ij}\}$, then from the commutativity relation (18), we can conclude $V_1(g)A = AV_1(g)$, $V_2(g)C = CV_1(g)$, $V_1(g)B = BV_2(g)$, and $V_2(g)D = DV_2(g)$ for all $g \in G$. Now, from Schur’s first lemma, we conclude that $A = \alpha_1 I_{n_1}$ and $D = \alpha_2 I_{n_2}$, with $\alpha_1$ and $\alpha_2$ belonging to the set of complex numbers. From the fact that $V_1$ and $V_2$ are irreducible representations, Schur’s second lemma implies that $B = C = 0$, where $0$ denotes zero matrix. The proof for $\mathcal{M}'$ diagonal in the general case, in which $U'$ is decomposed to more than two nonequivalent irreducible representations, is similar.