A Structure of Minimum Error Discrimination for Linearly Independent States

Tanmoy Singal, Eunsang Kim, and Sibasish Ghosh

1Department of Applied Mathematics, Hanyang University, Ansan Kyunggi-do, Korea
2Optics & Quantum Information Group, The Institute of Mathematical Sciences, CIT Campus, Taramani, Chennai, 600 113, India

In this paper we study the Minimum Error Discrimination problem (MED) for ensembles of linearly dependent (LI) states. We define a bijective map from the set of those ensembles to itself and we show that the Pretty Good Measurement (PGM) and the optimal measurement for the MED are related by the map. In particular, the fixed points of the map are those ensembles for which the PGM is the optimal measurement. Also, we simplify the optimality conditions for the measurement of an ensemble of LI states.

Keywords: minimum error discrimination, linearly independent states, mixed states pretty good measurement

I. INTRODUCTION

In quantum state discrimination, one wishes to optimally ascertain which of a collection of states has been provided. In general, two parties, Alice and Bob, are involved in this scenario. We may formulate the discrimination problem in the following way. Let $\mathcal{H}$ be a $d$-dimensional Hilbert space. Alice prepares a quantum state $\rho_i$, from an ensemble of quantum states $P = \{\rho_i, \rho_i\}_{i=1}^{m}$ with a priori probability $p_i$. Here the quantum states $\rho_i$ are density operators on $\mathcal{H}$ (i.e., $\rho_i \geq 0$, and $\text{Tr} \rho_i = 1$ for all $1 \leq i \leq m$), and the a priori probabilities $p_1, \ldots, p_m$ are such that $p_i > 0$ and $\sum_{i=1}^{m} p_i = 1$. We assume that the basis vectors of Range $\rho_i$ collectively span $\mathcal{H}$. Alice sends her state $\rho_i$ to Bob, without telling him what $i$ is. Instead, to find the value of $i$, Bob has to probe the state $\rho_i$ using an appropriate measurement. When the $\rho_i$’s are non-orthogonal, then they can’t be perfectly distinguished. The average probability of error in his inference of the value of $i$ is

$$E_i = \sum_{i,j=1}^{m} p_i \text{Tr} \rho_i E_j.$$ Bob’s objective is to obtain the positive operator valued measure (POVM), $\{E_i\}_{i=1}^{m}$, which maximizes the probability of success, i.e.,

$$p_s = \frac{\text{Max}}{\{E_i\}_{i=1}^{m}} \sum_{i=1}^{m} p_i \text{Tr} \rho_i E_i,$$

subject to the conditions $E_i \geq 0$ for all $1 \leq i \leq m$ and $\sum_{i=1}^{m} E_i = \text{Id}$, where the maximum is taken over the set of all $m$-element POVMs. This optimization problem is known as Minimum Error Discrimination (MED), or the quantum hypothesis testing problem [7, 1, 8, 4]. The POVM for which one obtains the maximum value is called the optimal POVM.

While there are many algorithms to iteratively solve the MED problem [10, 12, 21], there are only a few ensembles of states for which closed-form expressions for the optimal POVMs and success probability have been obtained. Some prominent examples for these are the two state ensemble [1, 4], ensembles of geometrically uniform states [1, 12, 13], various ensembles of states for $d = 2$ [14, 17, 18], etc. While most of the earlier results directly employ the optimality conditions (see Section II), to solve the problem, some of the later results use a variety of different structures of the problem to solve it, for instance, the geometric structure of the problem [3, 14, 13, 17], and an algebraic structure [18].

A structure of the MED problem was discovered by V. P. Belavkin [1]. He showed that for each distinct optimal POVM for the MED of some ensemble $P = \{\rho_i, \rho_i\}_{i=1}^{m}$ of quantum states, one can find another ensemble of quantum states $Q$, such that the pretty good measurement (PGM) of $Q$ is the optimal POVM of $P$. In [2] it was shown that in the case of linearly independent pure states, one can relate $P$ and $Q$ by a bijective mapping. In this work we prove that such a bijective mapping exists on sets of ensembles of LI mixed states as well. Using this map one may solve the MED problem for LI mixed state ensembles. However to construct the map we need the optimal POVM, and hence without knowing the optimal POVM we cannot construct this map. Our main objective in this paper is to construct the inverse map explicitly and by doing so we obtain some particular solutions for the MED problem. The fixed points of the map are ensembles whose pretty good measurements are their corresponding optimal POVMs. In this case, we find the necessary and sufficient condition for an ensemble to be a fixed point of this map. This is a generalisation of a result in [20]. Finally we show that the optimality conditions for the MED of LI mixed state ensembles is actually simpler than the well-known optimality conditions. This generalises a result obtained earlier [21, 22] for cases of LI pure state ensembles to LI mixed states.

This paper is organized as follows. In Section II we give a brief summary of the optimality conditions for MED. In Section III we describe a structure of the MED prob-
lem which was introduced by Belavkin \[1,3\]. In Section \[IV\] we build on this structure to prove the existence of a map on the set of LI ensembles, such that the PGM of the image (under the map) is the optimal POVM of the pre-image. Also, at the end of this Section we show that the optimality conditions for MED of LI mixed ensembles is actually simpler than for the well-known optimality conditions for general ensembles of states. In Section \[VII\] we prove that this map is bijective and explicitly construct its inverse. In Section \[VII\] we obtain necessary and sufficient conditions for the fixed points of this map. Section \[VII\] concludes the paper.

II. OPTIMALITY CONDITIONS

The set of \(m\)-element POVMs is a convex set. Thus MED is a convex optimization problem. Thus, one can formulate the dual problem as follows: for a given ensemble \(P = \{p_i, \rho_i\}_{i=1}^m\) of quantum states, find an operator \(Z\) which minimizes \(\text{Tr} Z\), subject to the condition \(Z \geq \rho_i\) for all \(1 \leq i \leq m\). For the MED problem there is no duality gap and the dual problem can be solved to obtain the optimal POVM \([2,3]\), i.e.,

\[
p_s = \min_{Z \geq \rho_i} \text{Tr} Z.
\]

We call the pair \(\{\Pi_i\}_{i=1}^m, Z\) an optimal dual pair when \(\{\Pi_i\}_{i=1}^m\) is an optimal POVM and \(Z\) satisfies the duality \([2]\). For an optimal dual pair \(\{\Pi_i\}_{i=1}^m, Z\) we have

\[
\sum_{i=1}^m p_i \text{Tr}(\Pi_i \rho_i) = \text{Tr} Z
\]

and hence it satisfies

\[
(Z - \rho_i) \Pi_i = \Pi_i (Z - \rho_i) = 0 \quad \text{for all} \quad i \in \{1, \cdots, m\}.
\]

Summing over \(i\) and using the relation \(\sum_{i=1}^m \Pi_i = \text{Id}\), we have

\[
Z = \sum_{i=1}^m p_i \Pi_i \rho_i = \sum_{i=1}^m p_i \Pi_i.
\]

Using equations \([4]\) and \([5]\) we get the following relations for all \(i, j \in \{1, \cdots, m\}\)

\[
\Pi_j (Z - p_j \rho_j) \Pi_i = \Pi_j (Z - p_j \rho_j) \Pi_i
\]

and this implies that

\[
\Pi_j (p_j \rho_j - \rho_i) \Pi_i = 0 \quad \text{for all} \quad i, j \in \{1, \cdots, m\}.
\]

One can also obtain equations \([4]\) from equations \([7]\) in the following way: \([7]\) implies that \(\sum_{j=1}^m p_j \rho_j \Pi_i = \sum_{j=1}^m p_j \Pi_j \rho_j\). Also, \([7]\) implies \([6]\), where \(Z := \sum_{j=1}^m p_j \rho_j \Pi_j\). In \([6]\), sum over \(j\) on both sides, and then we obtain equation \([4]\). Thus the conditions \([4]\) and \([7]\) are equivalent. Equation \([7]\) was derived in \([3]\) without using the dual problem. Equation \([7]\) is a constraint which the optimal POVM has to satisfy. The optimal POVM also has to satisfy the global maxima condition given by

\[
Z \geq p_i \rho_i \iff \sum_{j=1}^m p_j \rho_j \Pi_j - p_i \rho_i \geq 0, \text{for all} \quad 1 \leq i \leq m.
\]

Now we may summarize the optimality conditions in the following theorem as given in \([1,4]\) and \([7]\).

Theorem 1. A \(m\)-POVM \(\{\Pi_i\}_{i=1}^m\) for an ensemble \(P = \{p_i, \rho_i\}_{i=1}^m\) is optimal if and only if it satisfies the relations

1. \(\Pi_j (p_j \rho_j - \rho_i) \Pi_i = 0, \text{for all} \quad i, j \in \{1, \cdots, m\}\),
2. \(\sum_{j=1}^m p_j \rho_j \Pi_j \geq p_i \rho_i, \text{for all} \quad i \in \{1, \cdots, m\}\).

For any ensemble an optimal POVM always exists \([3]\), hence there must always exist an \(m\)-POVM which satisfies the conditions in Theorem 1. Let \(P = \{p_i, \rho_i\}_{i=1}^m\) be an ensemble of quantum states and let \(\{\Pi_i\}_{i=1}^m, Z\) be the optimal dual pair for \(P\). Then

\[
\sum_{i=1}^m p_i \text{Tr}(\rho_i \Pi_i) = \text{Tr}(Z)
\]

and the pair satisfies the conditions \([4]\) and \([8]\). In \([8]\) it was established that the operator \(Z\) from the optimal dual pair is unique, whereas the optimal POVM \(\{\Pi_i\}_{i=1}^m\) may not be unique.

III. A STRUCTURE FOR THE MED PROBLEM

Let \(\{\Pi_i\}_{i=1}^m, Z\) be the optimal dual pair for an ensemble \(P = \{p_i, \rho_i\}_{i=1}^m\) of quantum states. Based on the work in \([3]\), we construct an ensemble of quantum states associated with \(\{\Pi_i\}_{i=1}^m, Z\). Let us define

\[
\sigma_i := \frac{Z \Pi_i Z}{\text{Tr}(Z^2 \Pi_i)}, \text{for all} \quad 1 \leq i \leq m,
\]

and

\[
q_i := \frac{\text{Tr}(Z^2 \Pi_i)}{\text{Tr}(Z^2)}, \text{for all} \quad 1 \leq i \leq m.
\]

Then by the property of POVM we have \(\sum_{i=1}^m q_i = 1\) and

\[
q_i \sigma_i = \frac{Z \Pi_i Z}{\text{Tr}(Z^2)}, \text{for all} \quad 1 \leq i \leq m.
\]

Lemma 1. Let \(P = \{p_i, \rho_i\}_{i=1}^m\) be an ensemble of quantum states and let \(\{\Pi_i\}_{i=1}^m, Z\) be an optimal dual pair for \(P\). Then the ensemble \(Q = \{q_i, \sigma_i\}_{i=1}^m\) of quantum states defined in \([9]\) and \([10]\) satisfy the following properties:
(a) \( \sigma_i \geq 0 \) for all \( i = 1, \ldots, m \).

(b) \( \text{Tr}(\sigma_i) = 1 \) for all \( i = 1, \ldots, m \).

(c) \( \text{Range } q_i \sigma_i \subseteq \text{Range } p_i \rho_i \) for all \( i = 1, \ldots, m \).

Proof. Conditions (a) and (b) follow directly from equations (19) and (10). By (8), the operator \( Z \) is invertible and by the definition of \( \sigma_i \), we get rank \( \sigma_i = \text{rank } \Pi_i \) and from equations (7) we obtain that \( Z \Pi_i Z = p_i \rho_i \Pi_i \rho_i \) for all \( 1 \leq i \leq m \). This implies that \( \text{Range } (q_i \sigma_i) \subseteq \text{Range } (p_i \rho_i) \) for all \( 1 \leq i \leq m \).

For any ensemble \( Q = \{q_i, \sigma_i\}_{i=1}^m \) of quantum states, PGM of \( Q \) is defined as follows: for each \( i \in \{1, \ldots, m\} \), let

\[
    E_i := \sigma^{-1/2}(q_i \sigma_i)\sigma^{-1/2},
\]

where \( \sigma = \sum_{i=1}^m q_i \sigma_i \). Then it is easy to see that for all \( i \in \{1, \ldots, m\} \), \( E_i \geq 0 \) and

\[
    \sum_{i=1}^m E_i = \sum_{i=1}^m \sigma^{-1/2}(q_i \sigma_i)\sigma^{-1/2} = \sigma^{-1/2} \sum_{i=1}^m q_i \sigma_i \sigma^{-1/2} = \sigma^{-1/2} \sigma^{-1/2} = \text{Id}.
\]

Thus we see that \( \{E_i\}_{i=1}^m \) is a POVM.

Let \( P = \{p_i, \rho_i\}_{i=1}^m \) be an ensemble of quantum states with an optimal dual pair \( \{\Pi_i\}_{i=1}^m \) as given in Section 1 and also let \( Q = \{q_i, \sigma_i\}_{i=1}^m \) be the ensemble constructed from the optimal dual pair using equations (19) and (10). By (11),

\[
    \sigma = \sum_{i=1}^m q_i \sigma_i = \frac{Z^2}{\text{Tr}(Z^2)}.
\]

So \( \sigma^{-1/2} = \sqrt{\text{Tr}(Z^2)} Z^{-1} \). Now from (11) and (12), we get for all \( 1 \leq i \leq m \)

\[
    E_i = \sigma^{-1/2}(q_i \sigma_i)\sigma^{-1/2} = \sqrt{\text{Tr}(Z^2)} Z^{-1} Z \Pi_i Z \text{Tr}(Z^2) Z^{-1} = \Pi_i.
\]

This shows that the PGM of \( Q \) is the optimal POVM for MED of \( P \). In particular, in the case of pure states we have a nice property which is proved in (11, 12).

Theorem 2. Let \( P = \{p_i, \rho_i\}_{i=1}^m \) be an ensemble of pure states on a d-dimensional Hilbert space \( \mathcal{H} \) and let \( \{E_i\}_{i=1}^m \) be the PGM of the pure state ensemble \( Q = \{q_i, |\psi_i\rangle\langle \psi_i|\}_{i=1}^m \). For all \( i = 1, \ldots, m \), if \( p_i \langle \psi_i| \rho_i^{-1/2} |\psi_i\rangle = C \), where \( \rho_q = \sum_{i=1}^m q_i |\psi_i\rangle\langle \psi_i| \) and \( C \) is a constant so that \( \sum_{i=1}^m p_i = 1 \), then \( \{E_i\}_{i=1}^m \) is the optimal POVM for \( P \).

IV. STRUCTURE FOR LINEARLY INDEPENDENT STATES OF THE MED PROBLEM

Consider an ensemble \( P = \{p_i, \rho_i\}_{i=1}^m \) of quantum states on an d-dimensional Hilbert space \( \mathcal{H} \). Assume that the eigenvectors of \( \rho_i \), \( 1 \leq i \leq m \) collectively span \( \mathcal{H} \). Since each density operator \( \rho_i \) is Hermitian and \( \rho_i \geq 0 \), it has the eigendecomposition as \( \rho_i = \phi_i \phi_i^* \), where \( \phi_i = [\phi_{i1} |\phi_{i2}| \cdots |\phi_{ir_i}] \) is a \( d \times r_i \) matrix and \( \text{Rank } \rho_i = r_i \) and \( |\phi_{ik}\rangle \) is an eigenvector of \( \rho_i \). The set of quantum states \( \{\rho_i\}_{i=1}^m \) is said to be linearly independent if the set of vectors \( \{|\phi_{ik}\rangle \mid 1 \leq k \leq r_i, 1 \leq i \leq m \} \) spans \( \mathcal{H} \), we have \( \sum_{i=1}^m r_i = d \). An ensemble \( P = \{p_i, \rho_i\}_{i=1}^m \) of quantum states is said to be a LI state ensemble if the set \( \{\rho_i\}_{i=1}^m \) of density operators form a linearly independent set.

Define \( E(r_1, \ldots, r_m) \) to be the set of all LI state ensembles \( P = \{p_i, \rho_i\}_{i=1}^m \) such that \( \text{Rank } \rho_i = r_i \) for all \( 1 \leq i \leq m \). In (10), it was shown that for each element in \( E(r_1, \ldots, r_m) \), the optimal POVM is a projective measurement. More explicitly, we have

Theorem 3. Let \( P = \{p_i, \rho_i\}_{i=1}^m \in E(r_1, \ldots, r_m) \). Then the optimal POVM \( \{\Pi_i\}_{i=1}^m \) is a projective measurement. In other words, it satisfies \( \Pi_i \Pi_j = \delta_{ij} \Pi_i \) for all \( i, j \in \{1, \ldots, m\} \), and \( \sum_{i=1}^m \Pi_i = \text{Id} \).

Furthermore we have

Theorem 4. Let \( P = \{p_i, \rho_i\}_{i=1}^m \in E(r_1, \ldots, r_m) \) and let \( \{\Pi_i\}_{i=1}^m \) be an optimal POVM for \( P \). Then \( \text{Rank } \rho_i = \text{Rank } \Pi_i \) for all \( i = 1, \ldots, m \) and the optimal POVM \( \{\Pi_i\}_{i=1}^m \) for \( P \) is unique.

Proof. Let \( t_i = \text{Rank } (\Pi_i) \). Then by equation (19) and Lemma (1) we have \( t_i \leq r_i \) for all \( i = 1, \ldots, m \). By the assumption of LI states, \( \sum_{i=1}^m r_i = d \). Since \( \sum_{i=1}^m \Pi_i = \text{Id} \) on a d-dimensional Hilbert space,

\[
    \sum_{i=1}^m t_i = \sum_{i=1}^m \text{Rank } (\Pi_i) = d.
\]

Thus by combining with the condition \( \sum_{i=1}^m r_i = d \), we have \( t_i = r_i \) for all \( i = 1, \ldots, m \).

For the uniqueness, let \( \{\Pi_i\}_{i=1}^m \) be another optimal POVM for \( P = \{p_i, \rho_i\}_{i=1}^m \). Then

\[
    \text{Rank } (\Pi_i) = r_i = \text{Rank } (\Pi_i) \quad \text{for all } i = 1, \ldots, m.
\]

By (5),

\[
    Z = \sum_{i=1}^m p_i \Pi_i \rho_i = \sum_{i=1}^m p_i \Pi_i \rho_i.
\]

By the result in (3), the operator \( Z \) is unique. Since \( \rho_i \) are linearly independent we have \( \Pi_i = \Pi_i \) for all \( i = 1, \ldots, m \).
Let $\mathcal{P}(r_1, \ldots, r_m)$ be the set of all $m$-element projective measurements $\{\Pi_i\}_{i=1}^{m}$ such that $\text{Rank}(\Pi_i) = r_i$ for all $1 \leq i \leq m$.

Define a map, which will be referred to the optimal POVM map

$$\text{OP} : \mathcal{E}(r_1, \ldots, r_m) \rightarrow \mathcal{P}(r_1, \ldots, r_m)$$

as follows: for each $P = \{p_i, \rho_i\}_{i=1}^{m} \in \mathcal{E}(r_1, \ldots, r_m)$,

$$\text{OP}(P) = \{\Pi_i\}_{i=1}^{m} \in \mathcal{P}(r_1, r_2, \ldots, r_m),$$

where $\{\Pi_i\}_{i=1}^{m}$ is the optimal POVM for the ensemble $P$. Then by Theorem 3 and Theorem 4, the map $\text{OP}$ is well-defined.

We also define a map

$$\mathcal{R} : \mathcal{E}(r_1, \ldots, r_m) \rightarrow \mathcal{E}(r_1, \ldots, r_m)$$

as follows: for each $P = \{p_i, \rho_i\}_{i=1}^{m} \in \mathcal{E}(r_1, \ldots, r_m)$, let $\text{OP}(P) = \{\Pi_i\}_{i=1}^{m}$. Then as constructed in Section 2, we have $\mathcal{R}(P) = Q = \{q_i, \sigma_i\}_{i=1}^{m}$, where $Q = \{q_i, \sigma_i\}_{i=1}^{m}$ and $q_i, \sigma_i = \frac{\Pi_i Z}{\text{Tr}(Z^2)}$. The map $\mathcal{R}$ is well-defined.

Using the pretty good measurement one can also define PGM as a function

$$\text{PGM} : \mathcal{E}(r_1, r_2, \ldots, r_m) \rightarrow \mathcal{P}(r_1, \ldots, r_m)$$

such that

$$\text{PGM}(Q) = \{\Pi_i\}_{i=1}^{m}, \text{ where } \Pi_i = \sigma^{-1/2}(q_i, \sigma_i) \sigma^{-1/2}.$$

We have defined two functions $\text{OP}$ and $\text{PGM}$ from the set $\mathcal{E}(r_1, r_2, \ldots, r_m)$ to a set $\mathcal{P}(r_1, \ldots, r_m)$ and $\mathcal{R}$ maps from $\mathcal{E}(r_1, r_2, \ldots, r_m)$ to itself. The relation between these three functions are given in the following theorem, which is obvious from the definitions of the functions.

**Theorem 5.** Let $\mathcal{E}(r_1, \ldots, r_m)$ be the set of LI states ensemble whose $i$-th state is of rank $r_i$ and let $\mathcal{P}(r_1, \ldots, r_m)$ be the set of projective POVMs such that $\text{rank}(\Pi_i) = r_i$ for all $i = 1 \cdots m$. Then we have the following relation

$$\text{OP} = \text{PGM} \circ \mathcal{R}. \quad (14)$$

Moreover, we can show that the map $\mathcal{R}$ is bijective. For this we first explicitly construct another function $\mathcal{R}'$ on $\mathcal{E}(r_1, \ldots, r_m)$, and later show that $\mathcal{R}'$ is the left and right inverse of $\mathcal{R}$, i.e., we show that $\mathcal{R}^{-1}$ exists and it is equal to $\mathcal{R}'$.

**V. BIJECTIVITY OF $\mathcal{R}$**

In order to show that the map $\mathcal{R} : \mathcal{E}(r_1, \ldots, r_m) \rightarrow \mathcal{E}(r_1, \ldots, r_m)$ is bijective, we construct the inverse of the map.

Let $Q = \{q_i, \sigma_i\}_{i=1}^{m}$ be any element in $\mathcal{E}(r_1, \ldots, r_m)$ and let $\text{PGM}(Q) = \{\Pi_i\}_{i=1}^{m} \in \mathcal{P}(r_1, \ldots, r_m)$, then for all $i = 1, \cdots, m$,

$$\Pi_i = \sigma^{-1/2}(q_i, \sigma_i) \sigma^{-1/2}, \quad \sigma = \sum_{i=1}^{m} q_i \sigma_i. \quad (15)$$

Consider the following decomposition of $\sigma^{1/2}$.

$$\sigma^{1/2} = (\text{Id} - \Pi_i + \Pi_i) \sigma^{1/2}/2 (\text{Id} - \Pi_i) + \Pi_i \sigma^{1/2}/2 (\text{Id} - \Pi_i) + (\text{Id} - \Pi_i) \sigma^{1/2}/2 \Pi_i. \quad (16)$$

Since $\Pi_i (\text{Id} - \Pi_i) = (\text{Id} - \Pi_i) \Pi_i = 0$, $\Pi_i$ and $\text{Id} - \Pi_i$ can be simultaneously diagonalizable in some orthonormal basis. In such a basis, $\sigma^{1/2}$ can be represented by the following matrix

$$\sigma^{1/2} \leftrightarrow \begin{pmatrix} A_i & B_i \\ B_i^\dagger & C_i \end{pmatrix}, \quad (17)$$

where

(a) $A_i$ is an $r_i \times r_i$ matrix which represents $\Pi_i \sigma^{1/2} \Pi_i$ in \((16)\), and hence $A_i > 0$ \([11]\).

(b) $C_i$ is a $(d - r_i) \times (d - r_i)$ matrix, which represents $(\text{Id} - \Pi_i) \sigma^{1/2} (\text{Id} - \Pi_i)$ in \((16)\), and hence $C_i > 0$.

(c) $B_i$ is an $r_i \times (d - r_i)$ matrix which represents $\Pi_i \sigma^{1/2} (\text{Id} - \Pi_i)$ in \((16)\).

Define

$$\Delta_i \equiv C_i - B_i^\dagger (A_i)^{-1} B_i. \quad (18)$$

Note that $\Delta_i$ is the Schur complement of $A_i$ in the matrix representation of $\sigma^{1/2}$ in \((17)\). Since $\sigma^{1/2} > 0$, $\Delta_i > 0$ \([11]\). Define $X_i$ to be an operator, which is represented by the following matrix using the same basis as in \((17)\)

$$X_i \leftrightarrow \begin{pmatrix} A_i & B_i \\ B_i^\dagger & C_i \end{pmatrix} = \begin{pmatrix} (A_i B_i) & (A_i B_i) \\ (B_i^\dagger A_i^{-1} B_i) & (A_i B_i) \end{pmatrix} \equiv \begin{pmatrix} \text{Id}_{r_i} & 0 \\ 0 & \text{Id}_{d-r_i} \end{pmatrix} \begin{pmatrix} A_i & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \text{Id}_{r_i} & 0 \\ 0 & \text{Id}_{d-r_i} \end{pmatrix}. \quad (19)$$

Thus we see that

$$\text{Rank} X_i = \text{Rank} A_i = r_i \quad \text{and} \quad X_i \geq 0 \quad (20)$$

Now define

$$p_i \equiv \frac{\text{Tr} X_i}{\sum_{j=1}^{m} \text{Tr} X_j} \quad \text{and} \quad \rho_i \equiv \frac{X_i}{\text{Tr} X_i}. \quad (21)$$

Thus we obtain the ensemble $P = \{p_i, \rho_i\}_{i=1}^{m}$ of quantum states and by \((20)\) $\text{Rank} \rho_i = r_i$, for all $i = 1, \cdots, m$. 

Theorem 6. For any $Q = \{q_i, \sigma_i\}_{i=1}^m \in \mathcal{E}(r_1, \ldots, r_m)$, define $\mathcal{R}'(Q) = P$, where $P = \{p_i, \rho_i\}_{i=1}^m$ is an ensemble of quantum states as given in (21). Then $P \in \mathcal{E}(r_1, \ldots, r_m)$ and $\mathcal{R}'$ defines a function on $\mathcal{E}(r_1, \ldots, r_m)$. Furthermore, PGM(Q) is the optimal PVM for MED of $P$.

Proof. Let PGM(Q) = $\{\Pi_i\}_{i=1}^m \in \mathcal{P}(r_1, \ldots, r_m)$ and define

$$ Z \equiv \frac{\sigma^{1/2}}{\sum_{j=1}^m \text{Tr} X_j}. \quad (22) $$

Then for each $i = 1, \ldots, m$,

$$ Z - p_i \rho_i = \frac{\sigma^{1/2}}{\sum_{j=1}^m \text{Tr} X_j} \left( X_i - \sum_{j=1}^m \text{Tr} X_j \right) = \frac{1}{\sum_{j=1}^m \text{Tr} X_j} \left( \sigma^{1/2} - X_i \right). $$

In the matrix representation used earlier we see that

$$ \left( \sigma^{1/2} - X_i \right) \Pi_i \leftrightarrow \begin{pmatrix} 0 & 0 \\ 0 & \Delta_i \end{pmatrix} \begin{pmatrix} \text{Id}_{r_i} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \quad (23) $$

Thus $\{\Pi_i\}_{i=1}^m$ and $Z$ satisfy the equation (4) for the ensemble $\{p_i, \rho_i\}_{i=1}^m$. Also, since the matrix associated with $\sigma^{1/2} - X_i$ is a Schur complement in $\sigma^{1/2}$, $\sigma^{1/2} - X_i \geq 0$.

Thus $\{\Pi_i\}_{i=1}^m$ and $Z$ satisfy equation (8). By theorem 11 this shows that the pair $\{(\Pi_i\}_{i=1}^m, Z\}$ is an optimal dual pair for the MED of $P$.

Using the definition of $Z$ given in (22), and equations (4) and (11), we see that $p_i \rho_i$ should satisfy the following equation

$$ \frac{p_i \rho_i \Pi_i p_i \rho_i}{\text{Tr} Z^2} = \frac{Z \Pi_i Z}{\text{Tr} Z^2} = q_i \sigma_i, \quad i \in \{1, \ldots, m\}. \quad (24) $$

Hence $\text{Range} q_i \sigma_i \subseteq \text{Range} p_i \rho_i$, for all $i = 1, \ldots, m$. But since $\text{Rank} q_i \sigma_i = \text{Rank} p_i \rho_i = r_i$ for each $i = 1, \ldots, m$, we get $\text{Range} q_i \sigma_i = \text{Range} p_i \rho_i$. Since the $\sigma_i$’s are linearly independent states, the $\rho_i$’s are also linearly independent states. This shows that $P \in \mathcal{E}(r_1, \ldots, r_m)$.

From the construction, the $X_i$ are uniquely determined, and hence the map $\mathcal{R}' : \mathcal{E}(r_1, \ldots, r_m) \rightarrow \mathcal{E}(r_1, \ldots, r_m)$ is well-defined and this completes the proof. \hfill \Box

Hence in the theorem we show that

$$ \text{OP} (\mathcal{R}'(Q)) = \text{PGM} (Q). \quad (25) $$

We have shown that $\mathcal{R}'$ is a well-defined map, and we will show that this map is actually the inverse of $\mathcal{R}$. The map $\mathcal{R}$ was defined using equations (9) and (10). We see from equation (24) that $\mathcal{R}(P) = Q$, and hence we get that for each $Q$ in $\mathcal{E}(r_1, \ldots, r_m)$,

$$ \mathcal{R} \circ \mathcal{R}' (Q) = Q. \quad (26) $$

To establish that $\mathcal{R}'$ is the inverse of $\mathcal{R}$, it remains to show the following.

Theorem 7. $\mathcal{R} \circ \mathcal{R}' (P) = P$, for all $P \in \mathcal{E}(r_1, \ldots, r_m)$.

Proof. For any $P = \{p_i, \rho_i\}_{i=1}^m \in \mathcal{E}(r_1, \ldots, r_m)$, we obtain $Q = \mathcal{R}'(P) = \{q_i, \sigma_i\}_{i=1}^m$, using equation (11). Hence by Theorem 5 PGM(Q) = OP(P). Let OP(P) = $\{\Pi_i\}_{i=1}^m$. Thus by equation (5), $Z = \sum_{i=1}^m p_i \rho_i \Pi_i = \sum_{i=1}^m p_i \rho_i \rho_i$. By equation (13), we also have that $Z = \sqrt{\text{Tr} Z^2} \sigma^{1/2}$. Let $Z' = \sum_{i=1}^m \text{Tr} X_j \sigma^{1/2}$, where $Z'$ was introduced in equation (22). Thus $Z = cZ'$, where $c > 0$ is some constant. Let $\mathcal{R}'(Q) = P' = \{p_i', \rho_i'\}_{i=1}^m$, where $p_i'$ and $\rho_i'$ were defined in equation (21). Then we obtain the following conclusions.

$$ \Pi_i p_i \rho_i \Pi_i = \Pi_i Z \Pi_i = c \Pi_i Z' \Pi_i = c \Pi_i p_i' \rho_i' \Pi_i, \quad (27) $$

and

$$ \Pi_i p_i \rho_i (\text{Id} - \Pi_i) = \Pi_i Z (\text{Id} - \Pi_i) = c \Pi_i Z' (\text{Id} - \Pi_i) = c \Pi_i p_i' \rho_i' (\text{Id} - \Pi_i), \quad (28) $$

Using equations (19) and (21), we may represent $p_i \rho_i$ in the same orthonormal basis used in equation (17) as follows

$$ p_i \rho_i \leftrightarrow \frac{c}{\sum_{j=1}^m \text{Tr} X_j} \left( A_i \ B_i \ B_i^\dagger \ W_i \right), $$

where $W_i$ an $(d - r_i) \times (d - r_i)$ matrix, which should be positive semidefinite. $W_i - B_i^\dagger A_i^{-1} B_i$ is the Schur complement of $A_i$ in $\left( A_i \ B_i \ B_i^\dagger \ W_i \right)$. Using a result in [11],

$$ \text{Rank} p_i \rho_i = \text{Rank} A_i + \text{Rank} \left( W_i - B_i^\dagger A_i^{-1} B_i \right), \quad (30) $$

but since $\text{Rank} p_i \rho_i = \text{Rank} A_i = r_i$, we get that $\text{Rank} \left( W_i - B_i^\dagger A_i^{-1} B_i \right) = 0$. In other words, $W_i = B_i^\dagger A_i^{-1} B_i$, and thus

$$ p_i \rho_i \leftrightarrow \frac{c}{\sum_{j=1}^m \text{Tr} X_j} \left( A_i \ B_i \ B_i^\dagger \ A_i^{-1} B_i \right), \quad (31) $$

hence $p_i \rho_i = c p_i' \rho_i'$. But note that $\sum_{j=1}^m \text{Tr} p_i \rho_i = c = 1$. Hence we obtain that $P = P'$. Thus $\mathcal{R}'(Q) = P$, and hence $\mathcal{R}' \circ \mathcal{R} (P) = P$, for all $P \in \mathcal{E}(r_1, \ldots, r_m)$. \hfill \Box

Thus we have proved that $\mathcal{R}'$ is the left and right inverse of $\mathcal{R}$, which implies that $\mathcal{R}$ is a bijection. Also, note that we have explicitly constructed the mapping $\mathcal{R}^{-1}$. In the course of the proof of Theorem 7 we find a simplified condition for optimality which is given in Theorem...
Let \( Z = \sum_{i=1}^{m} p_i \rho_i \Pi_i \) using the same notations in Theorem 7. Then we see that the matrix representation of \( Z - p_i \rho_i \) is given by
\[
Z - p_i \rho_i \longleftrightarrow \frac{1}{\sum_{j=1}^{m} \text{Tr}X_j} \begin{pmatrix} 0 & 0 \\ 0 & C_i - B_i^† A_i^{-1} B_i \end{pmatrix}.
\] (32)

Note that \( C_i - B_i^† A_i^{-1} B_i \) is the Schur complement of \( A_i \) in \( \begin{pmatrix} A_i & B_i \\ B_i^† & C_i \end{pmatrix} \). If \( \begin{pmatrix} A_i & B_i \\ B_i^† & C_i \end{pmatrix} \geq 0 \), then \( C_i - B_i^† A_i^{-1} B_i \geq 0 \). This simplifies condition (2) in Theorem 11. This is summarized as follows.

**Theorem 8.** Let \( P = \{p_i, \rho_i\}_{i=1}^{m} \in \mathcal{E}(r_1, \cdots, r_m) \). Then \( \{\Pi_i\}_{i=1}^{m} \in \mathcal{P}(r_1, \cdots, r_m) \) is the optimal POVM for MED of \( P \) if
\begin{enumerate}
\item \( \{\Pi_i\}_{i=1}^{m} \) satisfies equation (4) and
\item \( \sum_{i=1}^{m} p_i \rho_i \Pi_i \geq 0 \).
\end{enumerate}

In fact, the optimality conditions can be simplified even further.

**Corollary 1.** Let \( P \in \mathcal{E}(r_1, \cdots, r_m) \) and \( \{\Pi_j\}_{j=1}^{m} \in \mathcal{P}(r_1, \cdots, r_m) \). Then \( \{\Pi_j\}_{j=1}^{m} \) is the optimal POVM for the MED of \( P \) if \( \sum_{i=1}^{m} p_i \rho_i \Pi_i \geq 0 \).

**Proof.** All we need to do is to prove that \( \{\Pi_i\}_{i=1}^{m} \) satisfy the condition (4). Since \( \sum_{i=1}^{m} p_i \rho_i \Pi_i > 0 \), \( \sum_{i=1}^{m} p_i \rho_i \Pi_i = \sum_{i=1}^{m} p_i \rho_i \Pi_i \). Thus we have that
\[
\Pi_j \left( \sum_{i=1}^{m} p_i \rho_i \Pi_i - \sum_{i=1}^{m} p_i \rho_i \Pi_i \right) \Pi_k
= \Pi_j \left( p_j \rho_j - p_k \rho_k \right) \Pi_k
= 0,
\] (33)

where we used the fact that \( \{\Pi_i\}_{i=1}^{m} \) is a projective measurement. Thus \( \{\Pi_i\}_{i=1}^{m} \) satisfy the condition (4). Also, note that \( \sum_{i=1}^{m} p_i \rho_i \Pi_i \geq 0 \), so by Theorem 8 \( \{\Pi_j\}_{j=1}^{m} \) is the optimal POVM for the MED of \( P \).

Hence Theorem 8 and Corollary 1 tell us that the optimality conditions for the MED of ensembles of LI states are actually simpler than for the case of more general ensembles of states. This also generalizes the results in (21), (22).

**VI. FIXED POINTS OF \( \mathcal{R} \)**

Let \( P \in \mathcal{E}(r_1, \cdots, r_m) \) be a fixed point of \( \mathcal{R} \), i.e., \( \mathcal{R}(P) = P \). Then by (12) we have
\[
\text{OP}(P) = \text{PGM}(P).
\]

In other words, if \( P \) is a fixed point of \( \mathcal{R} \), then its PGM is the optimal POVM. In the following theorem, we give necessary and sufficient conditions for \( P \) to be a fixed point of \( \mathcal{R} \).

**Theorem 9.** Let \( P = \{p_i, \rho_i\}_{i=1}^{m} \) be an element in \( \mathcal{E}(r_1, \cdots, r_m) \). Then \( \mathcal{R}(P) = P \) if and only if \( \sum_{i=1}^{m} \Pi_i \rho_i^{1/2} \Pi_i = c \text{Id} \), for some constant \( c > 0 \), where \( \{\Pi_i\}_{i=1}^{m} = \text{PGM}(P) \) and \( \rho = \sum_{i=1}^{m} p_i \rho_i \).

**Proof.** Suppose that \( \sum_{i=1}^{m} \Pi_i \rho_i^{1/2} \Pi_i = c \text{Id} \), for some constant \( c > 0 \), where \( \{\Pi_i\}_{i=1}^{m} = \text{PGM}(P) \) and \( \rho = \sum_{i=1}^{m} p_i \rho_i \). Then for each \( i = 1, 2, \cdots, m \),
\[
\Pi_i \rho_i^{1/2} \Pi_i = c \text{Id}.
\] (34)

Let \( \mathcal{R}^{-1}(P) = P' = \{p'_i, \rho'_i\}_{i=1}^{m} \). By (29), \( \text{OP}(P') = \text{PGM}(P) = \{\Pi_i\}_{i=1}^{m} \) and \( \{\Pi_i\}_{i=1}^{m}, \rho_i^{1/2} \) is the optimal dual pair for MED of \( P' \), where \( m > 0 \) is some constant.

Now we follow the same sequence of steps as in proof of Theorem 7 to show that \( \mathcal{R}^{-1}(P) = P \) by using the relation (31).

Let us fix an orthonormal basis which diagonalizes \( \Pi_i \) and we use this basis to obtain matrix representations. Consider the following matrix representation of \( m \rho_i^{1/2} \)
\[
m \rho_i^{1/2} \longleftrightarrow \begin{pmatrix} A_i & B_i \\ B_i^† & D_i \end{pmatrix},
\] (35)

where \( A_i \) represents \( m \Pi_i \rho_i^{1/2} \Pi_i \). Then by (34), \( A_i = m \text{Id}_{r_i} \). By the optimality conditions (4) we have \( m \rho_i^{1/2} \Pi_i = p_i \rho'_i \Pi_i \) and thus the matrix representation of \( p_i \rho'_i \) is given by
\[
p_i \rho'_i \longleftrightarrow \begin{pmatrix} m \text{Id}_{r_i} & B_i \\ B_i^† & \frac{1}{m} B_i^† B_i \end{pmatrix},
\] (36)

where \( \frac{1}{m} B_i^† B_i \) is obtained from equation (31) with \( A_i^{-1} = \frac{1}{m} \text{Id}_{r_i} \). Note that from equation (11), \( p_i \rho_i = \frac{p_i^2 \rho'_i \Pi_i}{m^2} \), which has the following matrix representation
\[
p_i \rho_i \longleftrightarrow \begin{pmatrix} m^2 \text{Id}_{r_i} & B_i \\ B_i^† & \frac{1}{m} B_i^† B_i \end{pmatrix}
= \frac{c}{m} \begin{pmatrix} m \text{Id}_{r_i} & B_i \\ B_i^† & \frac{1}{m} B_i^† B_i \end{pmatrix}.
\] (37)

Comparing equations (36) and (37) we get that \( p_i \rho_i = (c/m)p_i \rho'_i \). Summing over \( i \) and taking trace gives us that \( c = m \). Thus \( p_i \rho_i = p_i \rho'_i \), for all \( 1 \leq i \leq m \). Thus we get that \( P' = P \), or that \( \mathcal{R}(P) = P \).

Conversely, for some \( P = \{p_i, \rho_i\}_{i=1}^{m} \in \mathcal{E}(r_1, \cdots, r_m) \), let \( \mathcal{R}(P) = P \). Then \( \text{OP}(P) = \text{PGM}(P) = \{\Pi_i\}_{i=1}^{m} \). Let \( \{\Pi_i\}_{i=1}^{m}, Z \) be the optimal dual pair for MED of \( P \). Then by (13) and (35), we have, for some constant \( c > 0 \),
\[
Z = cp_i^{1/2} = \sum_{i=1}^{m} p_i \rho_i \Pi_i.
\] (38)

Since \( \Pi_i = \rho_i^{-1/2} p_i \rho_i \rho_i^{-1/2} \) with \( \rho = \sum_{i=1}^{m} p_i \rho_i \), we get \( p_i \rho_i = \rho_i^{1/2} \Pi_i \rho_i^{1/2} \). Thus we have \( p_i \rho_i \Pi_i = \rho_i^{1/2} \Pi_i \rho_i^{1/2} \Pi_i \).
Then by (68),
\[ Z = \sum_{i=1}^{m} p_i \rho_i \Pi_i = \rho^{1/2} \sum_{i=1}^{m} \Pi_i \rho^{1/2} \Pi_i \]
and hence \[ \sum_{i=1}^{m} \Pi_i \rho^{1/2} \Pi_i = c I. \]

Theorem 5 tells us that the PGM is the optimal POVM when the probability of successfully detecting the \( i \)-th state is proportional to \( \text{Rank} \rho_i \), i.e., \( p_i = \text{Tr} \rho_i \rho_i \) for all \( 1 \leq i \leq m \). In [20] it was shown that when the states \( \rho_i \) are LI and pure, and thus if one can invert \( \rho_i \), then the PGM is the optimal POVM when the probability of successfully identifying the \( i \)-th state is independent of \( i \), i.e., \( p_i = \frac{1}{m} \). Hence Theorem 5 reduces to the result in [20] for the case of linearly independent pure state ensembles.

VII. DISCUSSION AND CONCLUSION

In this work we generalize the results for the MED problem of LI pure state ensembles to mixed state ensembles. Firstly, we show that there exists a map \( \mathcal{R} \) on the set of LI ensembles, such that the pretty good measurement of the image of this map is the optimal POVM for the MED of the pre-image. Next, we show that \( \mathcal{R} \) is bijective, and we explicitly construct \( \mathcal{R}^{-1} \). This generalizes results obtained in [2]. The fixed points of \( \mathcal{R} \) are seen to be ensembles whose pretty good measurements are optimal for MED. In Theorem 8 we obtain necessary and sufficient conditions for an ensemble to be a fixed point of \( \mathcal{R} \). It is seen that for such cases, the probability of successfully detecting the \( i \)-th state is proportional to the rank of that state for all \( 1 \leq i \leq m \). This generalizes the result for LI pure state ensembles in [20], where it was shown that the probability of successfully detecting the \( i \)-th state is independent of \( i \). Also, in Theorem 5 and Corollary 1 we show that the optimality conditions for the MED of LI states is in fact simpler than the optimality conditions for general ensembles of states. This generalizes a result in obtained in [21, 22].

While the geometric structure of the MED problem has been employed to study it, particularly for the case of qubit systems [14, 15, 17], the structure which Belavkin introduced in [14] has received scant attention. In [2], Mochon rediscovered the structure for the case of pure state ensembles, and proved the existence of the map \( \mathcal{R} \) for LI pure state ensembles. This map was later employed in [9] to obtain the optimal POVM. Equations (14) tells us that to solve the MED problem it suffices to know the map \( \mathcal{R} \). However the construction of \( \mathcal{R} \) requires the optimal POVM. In fact it is a difficult problem to get an exact form of \( \mathcal{R} \). On the other hand, we have constructed \( \mathcal{R}^{-1} \) and thus if one can invert \( \mathcal{R}^{-1} \), then one solves the MED problem. This was done for the case of LI pure state ensembles in [2], where the authors used the implicit function theorem to do so. We would like to see if this can be generalized to the case of LI mixed state ensembles as well. Work for this is under progress.

[1] V. P. Belavkin (1975) Optimal multiple quantum statistical hypothesis testing. Stochastics, 1:1-4, 315-345, DOI: 10.1080/17442507508833114
[2] C. Mochon, ‘Family of generalized “pretty good” measurements and the minimal-error-pure-state discrimination problems for which they are optimal’, Phys. Rev. A 73 (2006), 032328.
[3] V.P. Belavkin and V. Maslov, ‘Design of Optimal Dynamic Analyzers: Mathematical Aspects of Wave Pattern Recognition’, Mathematical Aspects Of Computer Engineering Advances in Science and Technology in the USSR Mir Publishers, (1988).
[4] C.W. Helstrom, Quantum Detection and Estimation Theory (Academic Press, New York, 1976)
[5] Y.C. Eldar, A. Magretska and G.C. Verghese, ‘Designing optimal quantum detectors via semidefinite programming’, IEEE Trans. Inform. Theory 49 (2003), 1007-1012.
[6] [2] A.S. Holevo, ‘Statistical decision in quantum theory’, J. Multivar. Anal. 3 (1973), 337-394.
[7] H.P. Yuen, R.S. Kennedy, and M. Lax, IEEE Trans. Inform. Theory IT-21, 125 (1975)
[8] J. Bae, ‘Structure of minimum-error quantum state discrimination’, New J. Phys. 15 (2013), 073037.
[9] T. Singal and S. Ghosh, ‘Minimum error discrimination for an ensemble of linearly independent pure states’, J. Phys. A: Math. Theor. 49 (2016), 165304.
[10] Y.C. Eldar, ‘von Neumann measurement is optimal for detecting linearly independent mixed quantum states’ Phys. Rev. A 68 (2003) 052303.
[11] Boyd, S. and Vandenberghe, L. (2004), ”Convex Optimization”, Cambridge University Press (Appendix A.5.5, page 651)
[12] Ban, M., Kurokawa, K., Momose, R. et al. Int J Theor Phys (1997) 36: 1269. https://doi.org/10.1007/BF02435921
[13] Sasaki M, Barnett S M, Jozsa R, Osaki M and Hirota O Phys. Rev. A 59 3325, (1999)
[14] D. Ha and Y. Kwon, Complete analysis for three-qubit mixedstate discrimination, Phys. Rev. A 87, 062302 (2013).
[15] D. Ha and Y. Kwon, Discriminating N-qudit states using geometric structure Phys. Rev. A 90, 022330 (2014)
[16] M. Jeek, J. chek, and J. Fiurek, Phys. Rev. A 65, 060301(R) (2002)
[17] J. Bae and Won-Young Hwang Phys. Rev. A 87, 012334 (2013)
[18] Weir G, Barnett S, and Croke S, Phys. Rev. A 96, 022312 (2017)
[19] Jon Tyson, https://arxiv.org/abs/0902.0395
[20] M. Sasaki, K. Kato, M. Izutsu, and O. Hirota, Phys. Rev. A 58, 146 (1998)
[21] C. W. Helstrom, IEEE Trans. Inf. Theory IT-28, 359 (1982)

[22] Nicola Dalla Pozza and Gianfranco Pierobon, Phys. Rev. A 91, 042334 (2015)