A Semidefinite Programming Approach to Optimal Unambiguous Discrimination of Quantum States

Yonina C. Eldar, *Member, IEEE*

***Abstract*—In this paper, we consider the problem of unambiguous discrimination between a set of linearly independent pure quantum states. We show that the design of the optimal measurement that minimizes the probability of an inconclusive result can be formulated as a semidefinite programming problem. Based on this formulation, we develop a set of necessary and sufficient conditions for an optimal quantum measurement. We show that the optimal measurement can be computed very efficiently in polynomial time by exploiting the many well-known algorithms for solving semidefinite programs, which are guaranteed to converge to the global optimum.**

Using the general conditions for optimality, we derive necessary and sufficient conditions so that the measurement that results in an equal probability of an inconclusive result for each one of the quantum states is optimal. We refer to this measurement as the *equal-probability measurement (EPM)*. We then show that for any state set, the prior probabilities of the states can be chosen such that the EPM is optimal.

Finally, we consider state sets with strong symmetry properties and equal prior probabilities for which the EPM is optimal. We first consider geometrically uniform (GU) state sets that are defined over a group of unitary matrices and are generated by a single generating vector. We then consider compound GU state sets which are generated by a group of unitary matrices using *multiple* generating vectors, where the generating vectors satisfy a certain (weighted) norm constraint.

***Index Terms*—Compound geometrically uniform (CGU) quantum states, equal-probability measurement (EPM), geometrically uniform (GU) quantum states, quantum detection, semidefinite programming, unambiguous discrimination.**

I. Introduction

In recent years, research into the foundations of quantum physics has led to the emerging field of quantum information theory [1]. 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. To convey information using quantum states, we may prepare a quantum system in a pure quantum state, drawn from a collection of known states $\{|\phi_i\rangle,\ 1 \leq i \leq m\}$. To detect the information, the system is subjected to a quantum measurement. If the given states $|\phi_i\rangle$ are not orthogonal, then no measurement can distinguish perfectly between them [2]. A fundamental problem, therefore, is to design measurements optimized to distinguish between pure nonorthogonal quantum states.

We may formulate this problem within the framework of quantum detection, and seek the measurement that minimizes the probability of a detection error, or more generally, the Bayes cost [3]–[6]. More recently, a different approach to the problem has emerged, which in some cases may be more useful. This approach, referred to as unambiguous discrimination of quantum states, combines error-free discrimination with a certain fraction of inconclusive results. The basic idea, pioneered by Ivanovic [7], is to design a measurement that with a certain probability returns an inconclusive result, but such that if the measurement returns an answer, then the answer is correct with probability 1. Given an ensemble consisting of m states, the measurement therefore 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.

Ivanovic [7] developed a measurement which discriminates unambiguously between a pair of nonorthogonal pure states. The measurement gives the smallest possible probability of obtaining an inconclusive result for unambiguous discrimination, when distinguishing between two linearly independent nonorthogonal states with equal prior probabilities. This measurement was then further investigated by Dieks [8] and Peres [9], and was later extended by Jaeger and Shimony [10] to the case in which the two states have unequal prior probabilities.

Although the two-state problem is well developed, the problem of unambiguous discrimination between multiple quantum states has received considerably less attention. In [11], Peres and Terno consider unambiguous discrimination between three quantum states. Chefles [12] showed that a necessary and sufficient condition for the existence of unambiguous measurements for distinguishing between m quantum states is that the states are linearly independent. He also proposed a simple suboptimal measurement for unambiguous discrimination for which the probability of an inconclusive result is the same regardless of the state of the system. Equivalently, the measurement yields an equal probability of correctly detecting each one of the ensemble states. We refer to such a measurement as an equal-probability measurement (EPM). Chefles and Barnett [13] developed the optimal measurement for the special case in which the state vectors form a cyclic set, i.e., the vectors are generated by a cyclic group of unitary matrices using a single generating vector, and showed that it coincides with the EPM. In their paper, they raise the question of whether or not this is the only case for which the EPM is optimal.

Manuscript received June 18, 2002; revised September 18, 2002. This work was supported in part by BAE Systems Cooperative Agreement RP6891 under Army Research Laboratory Grant DAAD19-01-2-0008, by the Army Research Laboratory Collaborative Technology Alliance through BAE Systems Subcontract RK78554, and by Texas Instruments through the TI Leadership University Consortium.

The author was with the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA. She is now with the Technion–Israel Institute of Technology, Haifa 32000, Israel (e-mail: yonina@ee.technion.ac.il).

Communicated by P. W. Shor, Associate Editor for Quantum Information Theory.

Digital Object Identifier 10.1109/TIT.2002.807291

In this paper, we develop a general framework for unambiguous state discrimination which can be applied to any number of states with arbitrary prior probabilities. For our measurement, we consider general positive operator-valued measures [3], [14], consisting of $m+1$ measurement operators. We derive a set of necessary and sufficient conditions for an optimal measurement that minimizes the probability of an inconclusive result, by exploiting principles of duality theory in vector space optimization. In analogy to the quantum detection problem, deriving a closed-form analytical expression for the optimal measurement directly from these conditions is a difficult problem. However, our formulation has several advantages. First, it readily lends itself to efficient computational methods. Specifically, we show that the optimal measurement can be found by solving a standard semidefinite program (SDP) [15], which is a convex optimization problem. By exploiting the many well-known algorithms for solving SDPs [16], [17], the optimal measurement can be computed very efficiently in polynomial time. Since an SDP is convex, it does not suffer from local optimums, so that SDP-based algorithms are guaranteed to converge to the *global* optimum. Second, although the necessary and sufficient conditions are hard to solve directly, they can be used to verify a solution. Finally, the necessary and sufficient conditions lead to further insight into the optimal measurement. In particular, using these conditions we derive necessary and sufficient conditions on the state vectors, so that the EPM minimizes the probability of an inconclusive result. In contrast with the general optimality conditions, these conditions can be easily verified given the state ensemble and the prior probabilities. Using these conditions we show that for *any* set of state vectors the prior probabilities can be chosen such that the EPM is optimal.

Based on the necessary and sufficient conditions, we develop the optimal measurement for state sets with broad symmetry properties. In particular, we consider geometrically uniform (GU) state sets [18]–[20] defined over a group of unitary matrices. For such state sets, we show that the optimal measurement is the EPM, and we obtain a convenient characterization of the EPM that exploits the state symmetries. We then consider *compound GU (CGU)* state sets [21], [20] in which the state vectors are generated by a group of unitary matrices using *multiple* generating vectors. We obtain a convenient characterization of the EPM in this case, and show that when the generating vectors satisfy a certain constraint, the EPM is optimal.

The paper is organized as follows. After a statement of the problem in Section II, in Section III, we derive the necessary and sufficient conditions for the optimal measurement that minimizes the probability of an inconclusive result, by formulating the problem as an SDP. In Section IV, we consider the EPM and derive necessary and sufficient conditions on the state set and the prior probabilities so that the EPM is optimal. Efficient iterative algorithms for minimizing the probability of an inconclusive result which are guaranteed to converge to the global optimum are considered in Section V. In Sections VI and VII, we derive the optimal measurement for state sets with certain symmetry properties, and show that the optimal measurement coincides with the EPM.

II. Unambiguous Discrimination of Quantum States

Assume that a quantum system is prepared in a pure quantum state drawn from a collection of given states $\{|\phi_i\rangle,\ 1 \le i \le m\}$ in an r-dimensional complex Hilbert space $\mathcal{H}$, with $r \ge m$. The states span a subspace $\mathcal{U}$ of $\mathcal{H}$. To detect the state of the system, a measurement is constructed comprising $m+1$ measurement operators $\{\Pi_i,\ 0 \le i \le m\}$ that satisfy

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

The measurement operators are constructed so that either the state is correctly detected, or the measurement returns an inconclusive result. Thus, each of the operators $\Pi_i,\ 1 \le i \le m$ corresponds to detection of the corresponding states $|\phi_i\rangle,\ 1 \le i \le m$, and Π_0 corresponds to an inconclusive result.

Given that the state of the system is $|\phi_i\rangle$, the probability of obtaining outcome k is $\langle\phi_i|\Pi_k|\phi_i\rangle$. Therefore, to ensure that each state is either correctly detected or an inconclusive result is obtained, we must have

$$\langle\phi_i|\Pi_k|\phi_i\rangle = p_i\delta_{ik}, \qquad 1 \le i, k \le m \tag{2}$$

for some $0 \le p_i \le 1$. Since from (1), $\Pi_0 = I_r - \sum_{i=1}^{m} \Pi_i$, (2) implies that $\langle\phi_i|\Pi_0|\phi_i\rangle = 1 - p_i$, so that given that the state of the system is $|\phi_i\rangle$, the state is correctly detected with probability p_i, and an inconclusive result is returned with probability $1-p_i$.

It was shown in [12] that (2) can be satisfied if and only if the vectors $|\phi_i\rangle$ are linearly independent, or equivalently, $\dim\mathcal{U} = m$. We, therefore, make this assumption throughout the paper. In this case, we may choose

$$\Pi_i = p_i|\tilde{\phi}_i\rangle\langle\tilde{\phi}_i| \triangleq p_i Q_i, \qquad 1 \le i \le m \tag{3}$$

where

$$Q_i = |\tilde{\phi}_i\rangle\langle\tilde{\phi}_i|, \qquad 1 \le i \le m \tag{4}$$

and the vectors $|\tilde{\phi}_i\rangle \in \mathcal{U}$ are the *reciprocal states* associated with the states $|\phi_i\rangle$, i.e., they are the unique vectors in $\mathcal{U}$ such that

$$\langle\tilde{\phi}_i|\phi_k\rangle = \delta_{ik}, \qquad 1 \le i, k \le m. \tag{5}$$

With Φ and $\tilde{\Phi}$ denoting the matrices of columns $|\phi_i\rangle$ and $|\tilde{\phi}_i\rangle$, respectively,

$$\tilde{\Phi} = \Phi(\Phi^*\Phi)^{-1}. \tag{6}$$

Since the vectors $|\phi_i\rangle$ are linearly independent, $\Phi^*\Phi$ is always invertible. Alternatively

$$\tilde{\Phi} = (\Phi\Phi^*)^{\dagger}\Phi \tag{7}$$

so that

$$|\tilde{\phi}_i\rangle = (\Phi\Phi^*)^{\dagger}|\phi_i\rangle \tag{8}$$

where $(\cdot)^{\dagger}$ denotes the *Moore–Penrose pseudoinverse* [22]; the inverse is taken on the subspace spanned by the columns of the matrix.

We can immediately verify that the measurement operators given by (3) satisfy (2). If $r = m$ so that the dimension of $\mathcal{H}$ is equal to the dimension of the space $\mathcal{U}$ spanned by the vectors $|\phi_i\rangle$, then these operators are the unique operators satisfying (2). If, on the other hand, $r > m$, then the measurement operators are not strictly unique. Indeed, any measurement operators of the form

$$\Pi_i = p_i Q_i + |\mu_i\rangle\langle\mu_i|, \qquad 1 \le i \le m \tag{9}$$

where $|\mu_i\rangle \in \mathcal{U}^\perp$, also satisfy (2). Since $|\phi_i\rangle \in \mathcal{U}$, $\langle\phi_i|\mu_k\rangle = 0$ for every i, k so that the measurement operators given by (3) and (9) lead to the same detection probabilities $\langle\phi_i|\Pi_k|\phi_i\rangle = p_i\delta_{ik}$. We may, therefore, assume without loss of generality that the operators Π_i are restricted to $\mathcal{U}$, so that they have the form given by (3).

If the state $|\phi_i\rangle$ is prepared with prior probability η_i, then the total probability of correctly detecting the state is

$$P_D = \sum_{i=1}^{m} \eta_i \langle\phi_i|\Pi_i|\phi_i\rangle = \sum_{i=1}^{m} \eta_i p_i. \tag{10}$$

Our problem, therefore, is to choose the measurement operators $\Pi_i = p_i Q_i$, or equivalently, the probabilities $p_i \ge 0$, to maximize P_D, subject to the constraint (1). We can express this constraint directly in terms of the probabilities p_i as

$$\sum_{i=1}^{m} \Pi_i = \sum_{i=1}^{m} p_i Q_i \le I_r. \tag{11}$$

Note that (11) implies that $p_i \le 1$.

III. Semidefinite Programming (SDP) Formulation

We now show that our maximization problem (10) and (11) can be formulated as a standard SDP [15], [16], which is a convex optimization problem. There are several advantages to this formulation. First, the SDP formulation readily lends itself to efficient computational methods. Specifically, by exploiting the many well-known algorithms for solving SDPs [15], e.g., interior point methods[1] [16], [17], the optimal measurement can be computed very efficiently in polynomial time. Furthermore, SDP-based algorithms are guaranteed to converge to the global optimum. Second, by exploiting principles of duality theory in vector space optimization, the SDP formulation can be used to derive a set of necessary and sufficient conditions for the probabilities p_i to maximize P_D of (10) subject to the constraint (11).

We note that recently SDP-based methods have been employed in a variety of different problems in quantum detection and quantum information [6], [23]–[27].

After a description of the general SDP problem in Section III-A, in Section III-B we show that our maximization problem can be formulated as an SDP. Based on this formulation, we derive a set of necessary and sufficient conditions on the measurement operators, or equivalently, the probabilities p_i, to minimize the probability of an inconclusive result. Although in general obtaining a closed-form analytical solution directly from these conditions is a difficult problem, the conditions can be used to verify whether or not a set of measurement operators is optimal. Furthermore, these conditions lead to further insight into the optimal measurement operators. In particular, in Section IV, we use these conditions to develop necessary and sufficient conditions on the state vectors and the prior probabilities so that the EPM is optimal.

A. Semidefinite Programming

A standard SDP is the problem of minimizing

$$P(x) = \langle c|x\rangle \tag{12}$$

subject to

$$F(x) \ge 0 \tag{13}$$

where

$$F(x) = F_0 + \sum_{i=1}^{m} x_i F_i. \tag{14}$$

Here $|x\rangle \in \mathcal{R}^m$ is the vector to be optimized, x_i denotes the ith component of $|x\rangle$, $|c\rangle$ is a given vector in $\mathcal{R}^m$, and F_i are given matrices in the space $\mathcal{B}_n$ of $n \times n$ Hermitian matrices.[2]

The problem of (12) and (13) is referred to as the *primal problem*. A vector $|x\rangle$ is said to be *primal feasible* if $F(x) \ge 0$, and is *strictly primal feasible* if $F(x) > 0$. If there exists a strictly feasible point, then the primal problem is said to be strictly feasible. We denote the optimal value of $P(x)$ by $\hat{P}$.

An SDP is a convex optimization problem and can be solved very efficiently. Furthermore, iterative algorithms for solving SDPs are guaranteed to converge to the global minimum. The SDP formulation can also be used to derive necessary and sufficient conditions for optimality by exploiting principles of duality theory. The essential idea is to formulate a *dual problem* of the form $\max_Z D(Z)$ for some linear functional D whose maximal value $\hat{D}$ serves as a certificate for $\hat{P}$. That is, for all feasible values of $Z \in \mathcal{B}_n$, i.e., values of $Z \in \mathcal{B}_n$ that satisfy a certain set of constraints, and for all feasible values of $|x\rangle$, $D(Z) \le P(x)$, so that the dual problem provides a lower bound on the optimal value of the original (primal) problem. If in addition, we can establish that $\hat{P} = \hat{D}$, then this equality can be used to develop conditions for optimality on $|x\rangle$.

The dual problem associated with the SDP of (12) and (13) [15] is the problem of maximizing

$$D(Z) = -\mathrm{Tr}(F_0 Z) \tag{15}$$

subject to

$$\mathrm{Tr}(F_i Z) = c_i, \qquad 1 \le i \le m \tag{16}$$

$$Z \ge 0 \tag{17}$$

[1]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 an SDP 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.

[2]Although typically in the literature the matrices F_i are restricted to be real and symmetric, the SDP formulation can be easily extended to include Hermitian matrices F_i; see, e.g., [28]. In addition, many of the standard software packages for efficiently solving SDPs, for example the Self-Dual-Minimization (SeDuMi) package [29], [30], allow for Hermitian matrices.

where $Z \in \mathcal{B}_n$. A matrix $Z \in \mathcal{B}_n$ is said to be *dual feasible* if it satisfies (16) and (17) and is *strictly dual feasible* if it satisfies (16) and $Z > 0$. If there exists a strictly feasible point, then the dual problem is said to be strictly feasible.

For any feasible $|x\rangle$ and Z we have that

$$\begin{aligned} P(x) - D(Z) &= \langle c|x\rangle + \mathrm{Tr}(F_0 Z) \\ &= \sum_{i=1}^{m} x_i \mathrm{Tr}(F_i Z) + \mathrm{Tr}(F_0 Z) \\ &= \mathrm{Tr}(F(x)Z) \geq 0 \end{aligned} \tag{18}$$

so that as required, $D(Z) \leq P(x)$. Furthermore, it can be shown [15] that if both the primal problem and the dual problem are strictly feasible, then $\hat{P} = \hat{D}$ and $|x\rangle$ is an optimal primal point if and only if $|x\rangle$ is primal feasible, and there exists a dual feasible $Z \in \mathcal{B}_n$ such that

$$ZF(x) = 0. \tag{19}$$

Equation (19) together with (16), (17), and (13) constitute a set of necessary and sufficient conditions for $|x\rangle$ to be an optimal solution to the problem of (12) and (13), when both the primal and the dual are strictly feasible.

If $\hat{Z}$ maximizes $D(Z)$ so that $D(\hat{Z}) = \hat{D}$, then $|x\rangle$ is optimal if and only if $F(x) \geq 0$ and $\hat{Z}F(x) = 0$.

B. SDP Formulation of Unambiguous Discrimination

We now show that the unambiguous discrimination problem of (10) and (11) can be formulated as an SDP. Denote by $|p\rangle$ the vector of components p_i and by $|c\rangle$ the vector of components $-\eta_i$. Then our problem is to minimize

$$P(p) = \langle c|p\rangle \tag{20}$$

subject to

$$\sum_{i=1}^{m} p_i Q_i \leq I_r; \qquad p_i \geq 0,\ 1 \leq i \leq m. \tag{21}$$

To formulate this problem as an SDP, let F_i, $0 \leq i \leq m$ be the block-diagonal matrices defined by

$$F_0 = \begin{bmatrix} I_r & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix},\ F_1 = \begin{bmatrix} -Q_1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix}, \ldots,$$

$$F_m = \begin{bmatrix} -Q_m & & & \\ & 0 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}. \tag{22}$$

Then

$$F(p) = F_0 + \sum_{i=1}^{m} p_i F_i = \begin{bmatrix} I_r - \sum_{i=1}^{m} p_i Q_i & & & \\ & p_1 & & \\ & & \ddots & \\ & & & p_m \end{bmatrix} \tag{23}$$

so that the constraint $F(p) \geq 0$ is equivalent to $\sum_{i=1}^{m} p_i Q_i \leq I_r$ and $p_i \geq 0$, $1 \leq i \leq m$. Thus, the problem of (10) and (11) reduces to the SDP

$$\min_{p \in \mathcal{R}^m} \langle c|p\rangle \ \text{ subject to } \ F(p) \geq 0 \tag{24}$$

where $|c\rangle$ is the vector of components $-\eta_i$ with η_i being the prior probability of $|\phi_i\rangle$, and $F(p)$ is given by (23).

To derive a set of necessary and sufficient conditions for optimality on $|p\rangle$, we use the dual problem formulation of a general SDP (15)–(17) to formulate the dual problem associated with (24), which reduces to

$$\max_{X \in \mathcal{B}_r} -\mathrm{Tr}(X) \tag{25}$$

subject to

$$\mathrm{Tr}(Q_i X) - z_i = \eta_i, \qquad 1 \leq i \leq m \tag{26}$$

$$X \geq 0 \tag{27}$$

$$z_i \geq 0, \qquad 1 \leq i \leq m. \tag{28}$$

We can immediately verify that both the primal and the dual problem are strictly feasible. Therefore, it follows that $|p\rangle$ is optimal if and only if the components p_i of $|p\rangle$ satisfy (21), there exists a matrix X and scalars z_i, $1 \leq i \leq m$ that satisfy (26)–(28), and

$$X\left(I_r - \sum_{i=1}^{m} p_i Q_i\right) = 0 \tag{29}$$

$$z_i p_i = 0, \qquad 1 \leq i \leq m. \tag{30}$$

Note that (29) implies that for the optimal choice of p_i, the largest eigenvalue of $\sum_{i=1}^{m} p_i Q_i$ must be equal to 1. This condition has already been derived in [12].

If $\hat{X}$ and $\hat{z}_i$ maximize (25) subject to (26)–(28), then the optimal values of p_i can be found by solving (29) and (30) with $X = \hat{X}$, $z_i = \hat{z}_i$.

We summarize our results in the following theorem.

Theorem 1: Let $\{|\phi_i\rangle,\ 1 \leq i \leq m\}$ denote a set of state vectors with prior probabilities $\{\eta_i,\ 1 \leq i \leq m\}$ in an r-dimensional Hilbert space $\mathcal{H}$ that span an m-dimensional subspace $\mathcal{U}$ of $\mathcal{H}$, let $\{|\tilde{\phi}_i\rangle,\ 1 \leq i \leq m\}$ denote the reciprocal states in $\mathcal{U}$ defined by $\langle\tilde{\phi}_i|\phi_k\rangle = \delta_{ik}$, and let $Q_i = |\tilde{\phi}_i\rangle\langle\tilde{\phi}_i|$. Let Λ denote the set of all ordered sets of constants $\{p_i,\ 1 \leq i \leq m\}$ that satisfy $p_i \geq 0$ and $\sum_{i=1}^{m} p_i Q_i \leq I_r$, and let Γ denote the set of $r \times r$ Hermitian matrices X satisfying $X \geq 0$ and scalars $z_i \geq 0$, $1 \leq i \leq m$ such that $\mathrm{Tr}(Q_i X) - z_i = \eta_i$. Consider the problem $\min_{p_i \in \Lambda} P(p)$ where $P(p) = -\sum_{i=1}^{m} \eta_i p_i$ and the dual problem $\max_{X,\, z_i \in \Gamma} D(X)$ where $D(X) = -\mathrm{Tr}(X)$. Then

1) for any $p_i \in \Lambda$ and X, $z_i \in \Gamma$, $P(p) \geq D(X)$;
2) there is an optimal $|p\rangle$, denoted $|\hat{p}\rangle$, such that $\hat{P} = P(\hat{p}) \leq P(p)$ for any $|p\rangle \in \Lambda$;
3) there is an optimal X and optimal z_i, denoted $\hat{X}$ and $\hat{z}_i$, such that $\hat{D} = D(\hat{X}) \geq D(X)$ for any X, $z_i \in \Gamma$;
4) $\hat{P} = \hat{D}$;

5) a set of necessary and sufficient conditions on $|p\rangle$ to minimize $P(p)$ is that $p_i \in \Lambda$ and there exists X, $z_i \in \Gamma$ such that $X(I_r - \sum_{i=1}^m p_i Q_i) = 0$ and $z_i p_i = 0$, $1 \le i \le m$;
6) given $\hat{X}$ and $\hat{z}_i$ a set of necessary and sufficient conditions on $|p\rangle$ to minimize $P(p)$ is that $p_i \in \Lambda$, $\hat{X}(I_r - \sum_{i=1}^m p_i Q_i) = 0$ and $\hat{z}_i p_i = 0$, $1 \le i \le m$.

As we indicated at the outset, the necessary and sufficient conditions given by Theorem 1 are in general hard to solve directly, although they can be used to verify a solution. In addition, these conditions can be used to gain insight into the optimal measurement operators. In the next section, we will use Theorem 1 to develop necessary and sufficient conditions on a set of state vectors so that the EPM is optimal. Contrary to the conditions given by Theorem 1, these conditions can be easily verified.

IV. Equal-Probability Measurement (EPM)

A. EPM

A simple measurement that has been employed for unambiguous state discrimination is the measurement in which $p_i = p$, $1 \le i \le m$. This measurement results in equal probability of correctly detecting each of the states. We, therefore, refer to this measurement as the EPM.

To determine the value of p, let Φ have a singular value decomposition (SVD) [22], [19] of the form $\Phi = U\Sigma V^*$ where U is an $r \times r$ unitary matrix, Σ is a diagonal $r \times m$ matrix with diagonal elements $\sigma_i > 0$ arranged in descending order so that $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_m$, and V is an $m \times m$ unitary matrix. Then from (6) it follows that

$$\tilde{\Phi} = U(\Sigma^\dagger)^* V^* \tag{31}$$

where $\Sigma^\dagger$ is a diagonal $m \times r$ matrix with diagonal elements $1/\sigma_i$. Thus,

$$\sum_{i=1}^m Q_i = \sum_{i=1}^m |\tilde{\phi}_i\rangle\langle\tilde{\phi}_i| = \tilde{\Phi}\tilde{\Phi}^* = U(\Sigma^\dagger)^*\Sigma^\dagger U^* \tag{32}$$

and the largest eigenvalue of $\sum_{i=1}^m Q_i$ is equal to $1/\sigma_m^2$. To satisfy the condition (29), the largest eigenvalue of $p\sum_i Q_i$ must be equal to 1, so that

$$p = \sigma_m^2. \tag{33}$$

Therefore, our problem reduces to finding necessary and sufficient conditions on the vectors $|\phi_i\rangle$ such that $\Pi_i = \sigma_m^2 Q_i$ minimizes the probability of an inconclusive result.

In the next section, we develop conditions under which the EPM is optimal for unambiguous discrimination. In our development, we consider separately the case in which σ_m has multiplicity 1 and the case in which σ_m has multiplicity greater than 1. We derive a set of necessary and sufficient conditions for optimality of the EPM in the first case, and sufficient conditions for optimality in the second case. Two broad classes of state sets that satisfy these conditions are discussed in Sections VI and VII.

B. Conditions for Optimality

1) Necessary and Sufficient Conditions: Let s denote the multiplicity of σ_m so that $\sigma_m = \sigma_{m-1} = \cdots = \sigma_{m+s-1}$. We first consider the case in which $s = 1$. In this case, to satisfy (29) and (27) we must have that

$$X = b|u_m\rangle\langle u_m| \tag{34}$$

where $|u_k\rangle$ are the columns of U and $b \ge 0$. In addition, since $p_i = p > 0$, it follows from (30) that $z_i = 0$, $1 \le i \le m$ so that from (26)

$$\mathrm{Tr}(Q_i X) = b|\langle\tilde{\phi}_i|u_m\rangle|^2 = \eta_i, \qquad 1 \le i \le m. \tag{35}$$

Now, from (31) we have that

$$|\tilde{\phi}_i\rangle = U(\Sigma^\dagger)^*|v_i\rangle \tag{36}$$

where $|v_i\rangle$ denotes the ith column of V^*. Substituting into (35)

$$\frac{b}{\sigma_m^2}|v_i(m)|^2 = \eta_i, \qquad 1 \le i \le m \tag{37}$$

where $v_i(k)$ denotes the kth component of $|v_i\rangle$. Since

$$\sum_{i=1}^m |v_i(m)|^2 = \sum_{i=1}^m \eta_i = 1 \tag{38}$$

b must be equal to σ_m^2.

We conclude that when the multiplicity of σ_m is equal to 1, the EPM is optimal if and only if $|v_i(m)|^2 = \eta_i$, $1 \le i \le m$, i.e., if and only if each of the elements in the last row of V^* is equal to the prior probability of the corresponding state.

2) Sufficient Conditions: We now consider the case in which $s > 1$. To derive a set of sufficient conditions for the EPM to be optimal we construct a matrix X that satisfies the conditions of Theorem 1.

To satisfy (29) and (27) we let

$$X = \sum_{k=1}^s b_k |u_{m-k+1}\rangle\langle u_{m-k+1}| \tag{39}$$

with $b_k \ge 0$. Since $p_i = p > 0$, it follows from (30) that $z_i = 0$, $1 \le i \le m$ so that from (26), X must satisfy

$$\mathrm{Tr}(Q_i X) = \sum_{k=1}^s b_k |\langle\tilde{\phi}_i|u_{m-k+1}\rangle|^2 = \eta_i, \qquad 1 \le i \le m. \tag{40}$$

Substituting $|\tilde{\phi}_i\rangle = U(\Sigma^\dagger)^*|v_i\rangle$ into (40), we have that the constants b_k must satisfy

$$\frac{1}{\sigma_m^2}\sum_{k=1}^s b_k |v_i(m-k+1)|^2 = \eta_i, \qquad 1 \le i \le m \tag{41}$$

where $v_i(k)$ denotes the kth component of $|v_i\rangle$.

We conclude that the EPM is optimal if there exists constants $b_i \ge 0$, $1 \le i \le s$ such that

$$\begin{bmatrix} |v_1(m)|^2 & \cdots & |v_1(m-s+1)|^2 \\ |v_2(m)|^2 & \cdots & |v_2(m-s+1)|^2 \\ \vdots & & \vdots \\ |v_m(m)|^2 & \cdots & |v_m(m-s+1)|^2 \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_s \end{bmatrix} = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_m \end{bmatrix}. \tag{42}$$

The problem of determining whether there exists a $|b\rangle$ with components $b_i \geq 0$ such that (42) is satisfied is equivalent to verifying whether a standard linear program is feasible. Specifically, in a linear program the objective is to minimize a linear functional of the vector $|b\rangle$ of the form $\langle d|b\rangle$ for some vector $|d\rangle$, subject to the constraints $A|b\rangle = |y\rangle$ and[3] $|b\rangle \geq 0$ for some given matrix A and vector $|y\rangle$. A linear program is feasible if there exists a vector $|b\rangle$ that satisfies the constraints [31]. Thus, we can use standard linear programming techniques to determine whether a $|b\rangle$ exists that satisfies (42), or equivalently, whether given a set of state vectors with given prior probabilities, the EPM is optimal.

Note, that given a set of state vectors, we can always choose the prior probabilities η_i so that the EPM is optimal. This follows from the fact that the matrix in (42) depends only on the state vectors. Thus, any set of coefficients $b_i \geq 0$ will give a set of $\eta_i \geq 0$ that satisfy (42). The coefficients η_i will correspond to probabilities if $\sum_i \eta_i = 1$. Since $\sum_{i=1}^m |v_i(k)|^2 = 1$ for all k, $\sum_{i=1}^m \eta_i = \sum_{i=1}^s b_i$, and any set of coefficients $b_i \geq 0$ such that $\sum_i b_i = 1$ will result in a set of probabilities η_i for which the EPM is optimal.

In [13], the authors raise the question of whether or not cyclic state sets with equal prior probabilities are the only state sets for which the EPM is optimal. Here we have shown that the EPM can be optimal for *any* state set, as long as we choose the prior probabilities correctly. In Sections VI and VII, we consider state sets with equal prior probabilities for which the EPM is optimal, generalizing the result in [13].

We summarize our results regarding the EPM in the following theorem.

Theorem 2: Let $\{|\phi_i\rangle,\ 1 \leq i \leq m\}$ denote a set of state vectors with prior probabilities $\{\eta_i,\ 1 \leq i \leq m\}$ in a Hilbert space $\mathcal{H}$ that span an m-dimensional subspace $\mathcal{U}$ of $\mathcal{H}$, let $\{|\tilde{\phi}_i\rangle,\ 1 \leq i \leq m\}$ denote the reciprocal vectors in $\mathcal{U}$ defined by $\langle\tilde{\phi}_i|\phi_k\rangle = \delta_{ik}$, and let $Q_i = |\tilde{\phi}_i\rangle\langle\tilde{\phi}_i|$. Let $\Phi = U\Sigma V^*$ denote the matrix of columns $|\phi_i\rangle$, let $|v_i\rangle$ denote the columns of V^* and $v_i(k)$ the kth component of $|v_i\rangle$, let $\sigma_1 \geq \cdots \geq \sigma_m > 0$ denote the singular values of Φ, and let s be the multiplicity of σ_m. Let $\Pi_i = \sigma_m^2 Q_i$ denote the EPM operators. Then we have the following.

1) If $s = 1$, then the EPM minimizes the probability of an inconclusive result if and only if $|v_i(m)|^2 = \eta_i$ for $1 \leq i \leq m$.
2) If $s > 1$, then the EPM minimizes the probability of an inconclusive result if there exists constants $b_i \geq 0$, $1 \leq i \leq s$ such that (42) is satisfied.
3) Given a set of state vectors, we can always choose the prior probabilities η_i so that the EPM is optimal. Specifically, η_i is given by (42) where b_i are arbitrary coefficients satisfying $b_i \geq 0$, and $\sum_{i=1}^m b_i = 1$.

Theorem 2 provides necessary and sufficient conditions in the case $s = 1$ and sufficient conditions in the case $s > 1$ for the EPM to be optimal, which depend on the SVD of Φ and the prior probabilities η_i. It may also be useful to have a criterion which depends explicitly on the given states $|\phi_i\rangle$ and the prior probabilities. Theorem 3 provides a set of sufficient conditions on the states $|\phi_i\rangle$ and the prior probabilities η_i so that the EPM is optimal. The proof of the Theorem is given in the Appendix. In Sections VI and VII, we discuss some general classes of state sets that satisfy these conditions.

Theorem 3: Let $\{|\phi_i\rangle,\ 1 \leq i \leq m\}$ denote a set of state vectors with prior probabilities $\{\eta_i,\ 1 \leq i \leq m\}$ in a Hilbert space $\mathcal{H}$ that span an m-dimensional subspace $\mathcal{U}$ of $\mathcal{H}$. Let Φ denote the matrix of columns $|\phi_i\rangle$, and let q denote the number of distinct singular values of Φ. Then the equal-probability measurement minimizes the probability of an inconclusive result if $\langle\phi_i|(\Phi\Phi^*)^{t/2-1}|\phi_i\rangle = \eta_i a_t$ for $1 \leq i \leq m$ and $1 \leq t \leq q$, for some constants a_t.

V. Computational Aspects

In the general case, there is no closed-form analytical solution to the maximization problem (20) subject to (21). However, since this problem is a convex optimization problem, there are very efficient methods for solving (20). In particular, the optimal vector $|p\rangle$ can be computed on Matlab using the linear matrix inequality (LMI) Toolbox. Convenient interfaces for using the LMI toolbox are the Matlab packages IQCβ [32] and SeDuMi [29], [30]. These algorithms are guaranteed to converge to the global optimum in polynomial time within any desired accuracy.

The number of operations required for each iteration of a general SDP where $|x\rangle \in \mathcal{R}^m$ and $F_i \in \mathcal{B}_n$ is $O(m^2n^2)$. However, the computational load can be reduced substantially by exploiting structure in the matrices F_i. In our problem, these matrices are block diagonal, so that each iteration requires on the order of $O(m^4)$ operations [15].

To illustrate the computational steps involved in computing the optimal measurement, we now consider a specific example.

Consider the case in which the ensemble consists of three state vectors with equal probability $1/3$, where

$$|\phi_1\rangle = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix}, \quad |\phi_2\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\\0\end{bmatrix}, \quad |\phi_3\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}0\\1\\1\end{bmatrix}. \tag{43}$$

To find the optimal measurement operators, we first find the reciprocal states $|\tilde{\phi}_i\rangle$. With Φ denoting the matrix of columns $|\phi_i\rangle$, we have

$$\tilde{\Phi} = \Phi(\Phi^*\Phi)^{-1} = \begin{bmatrix} 1.73 & 0 & -1.41 \\ -1.73 & 1.41 & 1.41 \\ 1.73 & -1.41 & 0 \end{bmatrix} \tag{44}$$

and the vectors $|\tilde{\phi}_i\rangle$ are the columns of $\tilde{\Phi}$. Next, we form the matrices $Q_i = |\tilde{\phi}_i\rangle\langle\tilde{\phi}_i|$ which results in

$$Q_1 = 3\begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & -2 \\ 0 & -2 & 2 \end{bmatrix}$$
$$Q_3 = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{45}$$

[3]The inequality is to be understood as a component-wise inequality.

We can now find the optimal vector $|p\rangle$ using the IQCβ package on Matlab. To this end, we first define the matrices F_i according to (22), and define

$$|c\rangle = -\frac{1}{3}\begin{bmatrix}1\\1\\1\end{bmatrix}. \tag{46}$$

We then generate the code shown at the bottom of the page, assuming that the matrices F_i and the vector $|c\rangle$ have already been defined in Matlab. The optimal vector $|p\rangle$ is given by

$$|p\rangle = \begin{bmatrix}0\\0.17\\0.17\end{bmatrix} \tag{47}$$

and the optimal measurement operators $\Pi_i = p_i Q_i$ are

$$\Pi_1 = \begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix}, \quad \Pi_2 = \begin{bmatrix}0&0&0\\0&0.34&-0.34\\0&-0.34&0.34\end{bmatrix}$$
$$\Pi_3 = \begin{bmatrix}0.34&-0.34&0\\-0.34&0.34&0\\0&0&0\end{bmatrix}. \tag{48}$$

We can now use the necessary and sufficient conditions derived in Section III-B and summarized in Theorem 1 to verify that $|p\rangle$ given by (47) is the optimal probability vector. To this end, we first form the matrix $T = I_r - \sum_{i=1}^{3}\Pi_i$. Using the eigendecomposition of T, we conclude that the null space of T has dimension 1 and is spanned by the vector

$$|u\rangle = \begin{bmatrix}-0.81\\0.41\\0.41\end{bmatrix}. \tag{49}$$

Therefore, to satisfy (29) and (27), X must be equal to $X = a|u\rangle\langle u|$ for some $a \geq 0$. Since $p_1 = 0$ and p_2, $p_3 > 0$, (30) and (28) imply that $z_2 = z_3 = 0$ and $z_1 \geq 0$. Therefore, from (26), we must have that

$$\mathrm{Tr}(Q_2X) = \mathrm{Tr}(Q_3X) = \frac{1}{3} \tag{50}$$

and

$$\mathrm{Tr}(Q_1X) \geq \frac{1}{3}. \tag{51}$$

To satisfy (50), we choose

$$a = \frac{1}{3\langle u|Q_2|u\rangle} = 0.11. \tag{52}$$

With this choice of a, $\mathrm{Tr}(Q_3X) = 1/3$ and $\mathrm{Tr}(Q_1X) = 0.89 > 1/3$, so that the necessary and sufficient conditions are satisfied.

Now, suppose that instead of equal prior probabilities we assume that the prior probabilities are $\eta_1 = 0.6$, $\eta_2 = 0.2$, $\eta_3 = 0.2$. These priors were chosen to be equal to the elements of the last row of V^*, where $\Phi = U\Sigma V^*$. Since the smallest square singular value of Φ, $\sigma_3^2 = 0.07$, has multiplicity 1, (42) is satisfied and the EPM, consisting of the measurement operators $\Pi_i = pQ_i$ with $p = 0.07$, minimizes the probability of an inconclusive result. As before, we can immediately verify that this is indeed the correct solution using the necessary and sufficient conditions of Theorem 1. For this choice of Π_i, $T = I_r - p\sum_{i=1}^{3} Q_i$, and the null space of T is spanned by the vector

$$|u\rangle = \begin{bmatrix}0.68\\-0.52\\-0.52\end{bmatrix}. \tag{53}$$

Therefore, X must be equal to $X = a|u\rangle\langle u|$ for some $a \geq 0$. Since $p_i = p > 0$ for all i, $z_i = 0$, $1 \leq i \leq 3$ so that we must have

$$\mathrm{Tr}(Q_1X) = 0.6, \quad \mathrm{Tr}(Q_2X) = 0.2, \quad \mathrm{Tr}(Q_3X) = 0.2. \tag{54}$$

If we choose $a = 0.6/\langle u|Q_1|u\rangle = 0.07$, then (54) is satisfied, and the EPM is optimal.

In the remainder of the paper, we use the sufficient conditions of Theorem 3 to derive the optimal unambiguous measurement for state sets with certain symmetry properties. The symmetry properties we consider are quite general, and include many cases of practical interest. Specifically, in Section VI we consider GU state sets, and in Section VII we consider compound GU state sets. It is interesting to note that for these classes of state sets, the optimal measurement that minimizes the probability of a detection error is also known explicitly [19], [20].

VI. Geometrically Uniform (GU) State Sets

In this section, we consider the case in which the state vectors $|\phi_i\rangle$ are defined over a group of unitary matrices and are generated by a single generating vector. Such a state set is called *geometrically uniform (GU)* [18]. We first obtain a convenient characterization of the EPM in this case and then show that the EPM is optimal. This result generalizes a similar result of Chefles and Barnett [13].

```
>> abst_init_lmi                    % Initializing the LMI toolbox
>> p = rectangular(3,1);            % Defining a vector |p⟩ of length 3
>> F = F0;                          % Defining the matrix F(p); here Fi = Fi
>> for i = 1:3,
>>    eval(['W = F' num2str(i)]);
>>    F = F + p(i)*W;
>> end
>> F > 0;                           % Imposing the constraint
>> lmi_mincx_tbx(c'*p);             % Minimizing ⟨c|p⟩ subject to the constraint
>> P = value(p)                     % Getting the optimal value of p.
```

A. GU State Sets

Let $\mathcal{G}$ be a finite group of m unitary matrices U_i on $\mathcal{H}$. That is, $\mathcal{G}$ contains the identity matrix I_r; 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 $\mathcal{G}$ [33].

A state set generated by $\mathcal{G}$ using a single generating vector $|\phi\rangle$ is a set $\mathcal{S} = \{|\phi_i\rangle = U_i|\phi\rangle, U_i \in \mathcal{G}\}$. The group $\mathcal{G}$ will be called the *generating group* of $\mathcal{S}$. For concreteness, we assume that $U_1 = I_r$ so that $|\phi_1\rangle = |\phi\rangle$. 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 $\mathcal{S}$.

Alternatively, a state set is GU if given any two states $|\phi_i\rangle$ and $|\phi_j\rangle$ in the set, there is an isometry (a norm-preserving linear transformation) that transforms $|\phi_i\rangle$ into $|\phi_j\rangle$ while leaving the set invariant [18]. Intuitively, a state set is GU if it "looks the same" geometrically from any of the states in the set. Some examples of GU state sets are considered in [18], [19].

We note that in [19], a GU state set was defined over an *abelian* group of unitary matrices. Here we are not requiring the group $\mathcal{G}$ to be abelian.

A cyclic state set is a special case of a GU state set in which the generating group $\mathcal{G}$ has elements $U_i = Z^{i-1}$, $1 \le i \le m$, where Z is a unitary matrix with $Z^m = I_r$. A cyclic group generates a cyclic state set $\mathcal{S} = \{|\phi_i\rangle = Z^{i-1}|\phi\rangle, 1 \le i \le m\}$, where $|\phi\rangle$ is arbitrary.

Any binary state set $\mathcal{S} = \{|\phi_1\rangle, |\phi_2\rangle\}$ is a GU cyclic state set, because it can be generated by the binary group $\mathcal{G} = \{I_r, R\}$, where R is the reflection about the hyperplane halfway between the two states. Since R represents a reflection, R is unitary and $R^2 = I_r$.

B. The EPM for GU States

To derive the EPM for a GU state set with generating group $\mathcal{G}$, we need to determine the reciprocal states $|\tilde{\phi}_i\rangle$. It was shown in [21], [20] that for a GU state set with generating group $\mathcal{G}$, $\Phi\Phi^*$ commutes with each of the matrices $U_i \in \mathcal{G}$. For completeness, we repeat the argument here. Expressing $\Phi\Phi^*$ as

$$\Phi\Phi^* = \sum_{i=1}^{m} |\phi_i\rangle\langle\phi_i| = \sum_{i=1}^{m} U_i|\phi\rangle\langle\phi|U_i^* \tag{55}$$

we have that for all j

$$\begin{aligned}\Phi\Phi^*U_j &= \sum_{i=1}^{m} U_i|\phi\rangle\langle\phi|U_i^*U_j \\ &= U_j \sum_{i=1}^{m} U_j^*U_i|\phi\rangle\langle\phi|U_i^*U_j \\ &= U_j \sum_{i=1}^{m} U_i|\phi\rangle\langle\phi|U_i = U_j\Phi\Phi^*\end{aligned} \tag{56}$$

since $\{U_j^*U_i, 1 \le i \le m\}$ is just a permutation of $\mathcal{G}$.

If $\Phi\Phi^*$ commutes with U_j, then $T = (\Phi\Phi^*)^\dagger$ also commutes with U_j for all j. Thus, from (8) the reciprocal states are

$$|\tilde{\phi}_i\rangle = T|\phi_i\rangle = TU_i|\phi\rangle = U_iT|\phi\rangle = U_i|\tilde{\phi}\rangle \tag{57}$$

where

$$|\tilde{\phi}\rangle = T|\phi\rangle = (\Phi\Phi^*)^\dagger|\phi\rangle. \tag{58}$$

It follows that the reciprocal states are also GU with generating group $\mathcal{G}$ and generating vector $|\tilde{\phi}\rangle$ given by (58). Therefore, to compute the reciprocal states for a GU state set all we need is to compute the generating vector $|\tilde{\phi}\rangle$. The remaining vectors are obtained by applying the group $\mathcal{G}$ to $|\tilde{\phi}\rangle$. The EPM is then given by the measurement operators

$$Q_i = pU_i|\tilde{\phi}\rangle\langle\tilde{\phi}|U_i \tag{59}$$

where p is equal to the smallest eigenvalue of $\Phi\Phi^*$.

C. Optimality of the EPM

We now show that the EPM is optimal for GU state sets with equal prior probabilities $\eta_i = 1/m$. Since $\Phi\Phi^*$ commutes with U_j for all j, $(\Phi\Phi^*)^a$ also commutes with U_j for any a. Therefore, for all t

$$\begin{aligned}\langle\phi_i|(\Phi\Phi^*)^{t/2-1}|\phi_i\rangle &= \langle\phi|U_i^*(\Phi\Phi^*)^{t/2-1}U_i|\phi\rangle \\ &= \langle\phi|(\Phi\Phi^*)^{t/2-1}U_i^*U_i|\phi\rangle \\ &= \langle\phi|(\Phi\Phi^*)^{t/2-1}|\phi\rangle.\end{aligned} \tag{60}$$

Since $\langle\phi_i|(\Phi\Phi^*)^{t/2-1}|\phi_i\rangle$ does not depend on i, from Theorem 3 we conclude that the EPM is optimal.

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

Theorem 4 (GU State Sets): Let $\mathcal{S} = \{|\phi_i\rangle = U_i|\phi\rangle, U_i \in \mathcal{G}\}$ be a GU state set generated by a finite group $\mathcal{G}$ of unitary matrices, where $|\phi\rangle$ is an arbitrary state, and let Φ be the matrix of columns $|\phi_i\rangle$. Then the measurement that minimizes the probability of an inconclusive result is equal to the equal-probability measurement, and consists of the measurement operators

$$\Pi_i = p|\tilde{\phi}_i\rangle\langle\tilde{\phi}_i|$$

where $\{|\tilde{\phi}_i\rangle = U_i|\tilde{\phi}\rangle, U_i \in \mathcal{G}\}$

$$|\tilde{\phi}\rangle = (\Phi\Phi^*)^\dagger|\phi\rangle,$$

and p is the smallest eigenvalue of $\Phi\Phi^*$.

D. Example of a GU State Set

We now consider an example of a GU state set.

Consider the group $\mathcal{G}$ of $m = 4$ unitary matrices U_i, where

$$U_1 = I_4, \quad U_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

$$U_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \quad U_4 = U_2U_3. \tag{61}$$

Let the state set be $\mathcal{S} = \{|\phi_i\rangle = U_i|\phi\rangle, 1 \le i \le 4\}$, where $|\phi\rangle = 1/(3\sqrt{2})[2 \;\; 2 \;\; 1 \;\; 3]^*$, so that

$$\Phi = \frac{1}{3\sqrt{2}} \begin{bmatrix} 2 & 2 & 2 & 2 \\ 2 & -2 & 2 & -2 \\ 1 & 1 & -1 & -1 \\ 3 & -3 & -3 & 3 \end{bmatrix}. \tag{62}$$

From Theorem 4, the measurement that minimizes the probability of an inconclusive result is the EPM. Furthermore, the

reciprocal states $|\tilde{\phi}_i\rangle$ are also GU with generating group $\mathcal{G}$ and generator

$$|\tilde{\phi}\rangle = (\Phi\Phi^*)^\dagger|\phi\rangle = \frac{1}{4\sqrt{2}}\begin{bmatrix}3\\3\\6\\2\end{bmatrix} \tag{63}$$

so that $\{|\tilde{\phi}_i\rangle = U_i|\tilde{\phi}\rangle,\ 1 \le i \le 4\}$. Since

$$\Phi\Phi^* = \tfrac{2}{9}\begin{bmatrix}4&0&0&0\\0&4&0&0\\0&0&1&0\\0&0&0&9\end{bmatrix} \tag{64}$$

$p = 2/9$ and the EPM measurement operators are

$$\Pi_i = (2/9)Q_i = (2/9)U_i|\tilde{\phi}\rangle\langle\tilde{\phi}|U_i^*.$$

We can now use the necessary and sufficient conditions of Theorem 1 to verify that $\Pi_i = (2/9)|\tilde{\phi}_i\rangle\langle\tilde{\phi}_i|$ are indeed the optimal measurement operators. To this end, we first form the matrix $T = I_r - \sum_{i=1}^4 \Pi_i$. Using the eigendecomposition of T, we conclude that the null space of T has dimension 1 and is spanned by the vector

$$|u\rangle = \begin{bmatrix}0\\0\\1\\0\end{bmatrix}. \tag{65}$$

Therefore, to satisfy (29) and (27), X must be equal to $X = a|u\rangle\langle u|$ for some $a \ge 0$. Since $p_i = 2/9 > 0$, $1 \le i \le 4$, (30) and (28) imply that $z_i = 0$, $1 \le i \le 4$. Therefore, from (26) we must have that

$$\mathrm{Tr}(Q_1X) = \mathrm{Tr}(Q_2X) = \mathrm{Tr}(Q_3X) = \mathrm{Tr}(Q_4X) = \frac{1}{4}. \tag{66}$$

To satisfy (66) we choose

$$a = \frac{1}{4\langle u|Q_1|u\rangle} = \frac{2}{9}. \tag{67}$$

With this choice of a, $\mathrm{Tr}(Q_2X) = \mathrm{Tr}(Q_3X) = \mathrm{Tr}(Q_4X) = 1/4$, so that as we expect the necessary and sufficient conditions are satisfied.

VII. Compound Geometrically Uniform (CGU) State Sets

In Section VI, we showed that the optimal measurement for a GU state set is the EPM associated with this set. We also showed that the reciprocal states are themselves GU and can, therefore, be computed using a single generator. In this section, we consider state sets which consist of subsets that are GU, and are therefore referred to as *compound geometrically uniform (CGU)* [21]. As we show, the reciprocal states are also CGU so that they can be computed using a *set* of generators. Under a certain condition on the generating vectors, we also show that the EPM associated with a CGU state set is optimal.

A CGU state set is defined as a set of vectors

$$\mathcal{S} = \{|\phi_{ik}\rangle,\ 1 \le i \le l,\ 1 \le k \le r\}$$

such that $|\phi_{ik}\rangle = U_i|\phi_k\rangle$, where the matrices $\{U_i,\ 1 \le i \le l\}$ are unitary and form a group $\mathcal{G}$, and the vectors $\{|\phi_k\rangle,\ 1 \le k \le r\}$ are the generating vectors. For consistency with the symmetry of $\mathcal{S}$, we will assume equiprobable prior probabilities on $\mathcal{S}$.

A CGU state set is in general not GU. However, for every k, the vectors $\{|\phi_{ik}\rangle,\ 1 \le i \le l\}$ are a GU state set with generating group $\mathcal{G}$. Examples of CGU state sets are considered in [21], [20].

A. The EPM for CGU State Sets

We now derive the EPM for a CGU state set with equal prior probabilities. Let Φ denote the matrix of columns $|\phi_{ik}\rangle$, where the first l columns correspond to $k = 1$, and so forth. Then, for a CGU state set with generating group $\mathcal{G}$, it was shown in [21], [20] that $\Phi\Phi^*$ commutes with each of the matrices $U_i \in \mathcal{G}$. If $\Phi\Phi^*$ commutes with U_i, then $T = (\Phi\Phi^*)^\dagger$ also commutes with U_i for all i. Thus, the reciprocal states are

$$|\tilde{\phi}_{ik}\rangle = T|\phi_{ik}\rangle = TU_i|\phi_k\rangle = U_iT|\phi_k\rangle = U_i|\tilde{\phi}_k\rangle \tag{68}$$

where

$$|\tilde{\phi}_k\rangle = T|\phi_k\rangle = (\Phi\Phi^*)^\dagger|\phi_k\rangle. \tag{69}$$

Therefore, the reciprocal states are also CGU with generating group $\mathcal{G}$ and generating vectors $|\tilde{\phi}_k\rangle$ given by (69). To compute these vectors all we need is to compute the generating vectors $|\tilde{\phi}_k\rangle$. The remaining vectors are then obtained by applying the group $\mathcal{G}$ to each of the generating vectors.

B. CGU State Sets With GU Generators

A special class of CGU state sets is *CGU state sets with GU generators* [21] in which the generating vectors $\{|\phi_k\rangle,\ 1 \le k \le r\}$ are themselves GU. Specifically, $\{|\phi_k\rangle = V_k|\phi\rangle\}$ for some generator $|\phi\rangle$, where the matrices $\{V_k,\ 1 \le k \le r\}$ are unitary, and form a group $\mathcal{Q}$. Examples of CGU state sets with GU generators are considered in [20].

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 indexes i and k. In this case, we say that $\mathcal{G}$ and $\mathcal{Q}$ commute up to a phase factor (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 [21]). Then for all $i,\ k$, $\Phi\Phi^*$ commutes with U_iV_k [21], [20]. The reciprocal states $|\tilde{\phi}_{ik}\rangle$ of the vectors $|\phi_{ik}\rangle$ are, therefore, given by

$$|\tilde{\phi}_{ik}\rangle = T|\phi_{ik}\rangle = TU_iV_k|\phi\rangle = U_iV_kT|\phi\rangle = U_iV_k|\overline{\phi}\rangle \tag{70}$$

where $|\overline{\phi}\rangle = T|\phi\rangle$. Thus, even though the state set is not in general GU, the reciprocal states can be computed using a single generating vector.

Alternatively, we can express $|\tilde{\phi}_{ik}\rangle$ as $|\tilde{\phi}_{ik}\rangle = U_i|\tilde{\phi}_k\rangle$ where the generators $|\tilde{\phi}_k\rangle$ are given by

$$|\tilde{\phi}_k\rangle = V_k|\overline{\phi}\rangle. \tag{71}$$

From (71), it follows that the generators $|\tilde{\phi}_k\rangle$ are GU with generating group $\mathcal{Q} = \{V_k,\ 1 \le k \le r\}$ and generator $|\overline{\phi}\rangle$.

We conclude that for a CGU state set with commuting GU generators and generating group $\mathcal{Q}$, the reciprocal states are also CGU with commuting GU generators and generating group $\mathcal{Q}$.

C. The Optimal Measurement for CGU State Sets Satisfying a Weighted Norm Constraint

We now show that if the generating vectors $|\phi_k\rangle$ satisfy

$$\langle\phi_k|(\Phi^*\Phi)^{t/2-1}|\phi_k\rangle = a_t, \qquad 1 \le k \le r,\ 1 \le t \le q \tag{72}$$

$$H=\begin{bmatrix} \sum_{i=1}^{s_1}|v_1(i)|^2 & \sum_{i=1}^{s_1}|v_2(i)|^2 & \cdots & \sum_{i=1}^{s_1}|v_2(i)|^2 \\ \sum_{i=1}^{s_2}|v_1(s_1+i)|^2 & \sum_{i=1}^{s_2}|v_2(s_1+i)|^2 & \cdots & \sum_{i=1}^{s_2}|v_2(s_1+i)|^2 \\ \vdots & \vdots & & \vdots \\ \sum_{i=1}^{s_q}\beta_i|v_1(m-s_q+i)|^2 & \sum_{i=1}^{s_q}\beta_i|v_2(m-s_q+i)|^2 & \cdots & \sum_{i=1}^{s_q}\beta_i|v_2(m-s_q+i)|^2 \end{bmatrix} \tag{78}$$

where q is the number of distinct singular values of Φ, then the EPM is optimal.

From Theorem 3, it follows that it is sufficient to show that (72) implies

$$\langle\phi_{ik}|(\Phi^*\Phi)^{t/2-1}|\phi_{ik}\rangle = a_t, \qquad 1\le i\le l,\ 1\le k\le r,\ 1\le t\le q. \tag{73}$$

Now

$$\begin{aligned}(\Phi^*\Phi)^{t/2-1}|\phi_{ik}\rangle &= (\Phi^*\Phi)^{t/2-1}U_i|\phi_k\rangle \\ &= U_i(\Phi^*\Phi)^{t/2-1}|\phi_k\rangle \end{aligned}\tag{74}$$

so that

$$\begin{aligned}\langle\phi_{ik}|(\Phi^*\Phi)^{t/2-1}|\phi_{ik}\rangle &= \langle\phi_k|U_i^*U_i(\Phi^*\Phi)^{t/2-1}|\phi_k\rangle \\ &= \langle\phi_k|(\Phi^*\Phi)^{t/2-1}|\phi_k\rangle = a_t \end{aligned}\tag{75}$$

establishing (73).

For CGU state sets with GU generators $\{|\phi_k\rangle = V_k|\phi\rangle\}$ where $V_k\in\mathcal{Q}$ and $\mathcal{G}$ and $\mathcal{Q}$ commute up to a phase factor, the EPM is optimal. This follows from the fact that in this case (72) is always satisfied. To see this, we first note that V_k commutes with $\Phi\Phi^*$ for each k [21]. Therefore, for all k

$$\begin{aligned}\langle\phi_k|(\Phi^*\Phi)^{t/2-1}|\phi_k\rangle &= \langle\phi|V_k^*(\Phi^*\Phi)^{t/2-1}V_k|\phi\rangle \\ &= \langle\phi|V_k^*V_k(\Phi^*\Phi)^{t/2-1}|\phi\rangle \\ &= \langle\phi|(\Phi^*\Phi)^{t/2-1}|\phi\rangle. \end{aligned}\tag{76}$$

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

Theorem 5 (CGU State Sets): Let

$$\mathcal{S}=\{|\phi_{ik}\rangle = U_i|\phi_k\rangle,\ 1\le i\le l,\ 1\le k\le r\}$$

be a CGU state set generated by a finite group $\mathcal{G}=\{U_i,\ 1\le i\le l\}$ of unitary matrices and generating vectors $\{|\phi_k\rangle,\ 1\le k\le r\}$, and let Φ be the matrix of columns $|\phi_{ik}\rangle$. Then the EPM consists of the measurement operators

$$\Pi_i = p|\tilde\phi_{ik}\rangle\langle\tilde\phi_{ik}|$$

where $\{|\tilde\phi_{ik}\rangle = U_i|\tilde\phi_k\rangle,\ 1\le i\le l,\ 1\le k\le r\}$

$$|\tilde\phi_k\rangle = (\Phi\Phi^*)^\dagger|\phi_k\rangle,$$

and p is equal to the smallest eigenvalue of $\Phi\Phi^*$.

The EPM has the following properties.

1) If $\langle\phi_k|(\Phi\Phi^*)^{t/2-1}|\phi_k\rangle = a_t$ for $1\le k\le r$, $1\le t\le q$, where q is the number of distinct eigenvalues of $\Phi\Phi^*$, then the EPM minimizes the probability of an inconclusive result.
2) If the generating vectors $\{|\phi_k\rangle = V_k|\phi\rangle,\ 1\le k\le r\}$ are GU with $U_iV_k = V_kU_ie^{j\theta(i,k)}$ for all i, k, then
   a) $|\tilde\phi_{ik}\rangle = U_iV_k|\overline{\phi}\rangle$ where $|\overline{\phi}\rangle = (\Phi\Phi^*)^\dagger|\phi\rangle$ so that the reciprocal states are CGU with GU generators;
   b) the EPM is optimal;
   c) if in addition $\theta(i,k)=0$ for all i, k, then the vectors $\{\phi_{ik},\ 1\le i\le l,\ 1\le k\le r\}$ form a GU state set.

Appendix
Proof of Theorem 3

In this appendix we prove Theorem 3.

Let λ_i, $1\le i\le q$ denote the singular values of Φ without multiplicity so that $\lambda_1=\sigma_1$ and $\lambda_q=\sigma_m$, and let s_i denote the multiplicity of λ_i. Define

$$A=\begin{bmatrix}\lambda_1 & \lambda_2 & \cdots & \lambda_q \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_q^2 \\ \vdots & \vdots & & \vdots \\ \lambda_1^q & \lambda_2^q & \cdots & \lambda_q^q\end{bmatrix} \tag{77}$$

and (78) as shown at the top of the page, for some $\beta_i\ge 0$. Finally, let N be the matrix with ith column equal to $\eta_i|a\rangle$ where $|a\rangle$ is an arbitrary vector.

Now, suppose that $AH=N$. Then $A|h_i\rangle=\eta_i|a\rangle$, where $|h_i\rangle$ denotes the ith column of H. Since A is invertible, this implies that

$$\frac{1}{\eta_i}h_i(k)=\frac{1}{\eta_j}h_j(k), \qquad 1\le i,j\le m,\ 1\le k\le q. \tag{79}$$

For $k=q$, (79) reduces to (42). We, therefore, conclude that a sufficient condition for the EPM to be optimal is that $AH=N$ for some $\beta_i\ge 0$. Taking $\beta_i=1$ for each i, we can express AH as

$$AH=\begin{bmatrix}\sigma_1 & \sigma_2 & \cdots & \sigma_m \\ \sigma_1^2 & \sigma_2^2 & \cdots & \sigma_m^2 \\ \vdots & \vdots & & \vdots \\ \sigma_1^q & \sigma_2^q & \cdots & \sigma_m^q\end{bmatrix} \cdot \begin{bmatrix}|v_1(1)|^2 & |v_2(1)|^2 & \cdots & |v_2(m)|^2 \\ |v_1(2)|^2 & |v_2(2)|^2 & \cdots & |v_2(2)|^2 \\ \vdots & \vdots & & \vdots \\ |v_1(m)|^2 & |v_2(m)|^2 & \cdots & |v_2(m)|^2\end{bmatrix} \triangleq Y. \tag{80}$$

Then, we have that

$$Y_{tl} = \sum_{i=1}^{m} \sigma_i^t |v_l(i)|^2 = \langle \phi_l | (\Phi\Phi^*)^{t/2-1} | \phi_l \rangle. \qquad (81)$$

Therefore, $AH = N$ reduces to the condition that

$$\langle \phi_l | (\Phi\Phi^*)^{t/2-1} | \phi_l \rangle = \eta_l a_t, \qquad 1 \le l \le m, 1 \le t \le q \qquad (82)$$

for some constants a_t.

ACKNOWLEDGMENT

The author wishes to thank Prof. A. Megretski and Prof. G. C. Verghese for many helpful discussions on semidefinite programming.

REFERENCES

[1] C. H. Bennett and P. W. Shor, "Quantum information theory," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2724–2742, Oct. 1998.
[2] A. Peres, *Quantum Theory: Concepts and Methods*. Boston, MA: Kluwer, 1995.
[3] C. W. Helstrom, *Quantum Detection and Estimation Theory*. New York: Academic, 1976.
[4] A. S. Holevo, "Statistical decisions in quantum theory," *J. Multivar. Anal.*, vol. 3, pp. 337–394, Dec. 1973.
[5] H. P. Yuen, R. S. Kennedy, and M. Lax, "Optimum testing of multiple hypotheses in quantum detection theory," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 125–134, Mar. 1975.
[6] Y. C. Eldar, A. Megretski, and G. C. Verghese, "Designing optimal quantum detectors via semidefinite programming," *IEEE Trans. Inform. Theory*. [Online]. Available: http://www.arXiv.org/abs/quant-ph/0205178, to be published.
[7] I. D. Ivanovic, "How to differentiate between nonorthogonal states," *Phys. Lett. A*, vol. 123, pp. 257–259, Aug. 1987.
[8] D. Dieks, "Overlap and distinguishability of quantum states," *Phys. Lett. A*, vol. 126, pp. 303–307, 1988.
[9] A. Peres, "How to differentiate between nonorthogonal states," *Phys. Lett. A*, vol. 128, p. 19, Mar. 1988.
[10] G. Jaeger and A. Shimony, "Optimal distinction between two nonorthogonal quantum states," *Phys. Lett. A*, vol. 197, pp. 83–87, 1995.
[11] A. Peres and D. R. Terno, "Optimal distinction between nonorthogonal quantum states," *J. Phys. A*, vol. 31, pp. 7105–7111, 1998.
[12] A. Chefles, "Unambiguous discrimination between linearly independent quantum states," *Phys. Lett. A*, vol. 239, pp. 339–347, Apr. 1998.
[13] A. Chefles and S. M. Barnett, "Optimum unambiguous discrimination between linearly independent symmetric states," *Phys. Lett. A*, vol. 250, pp. 223–229, 1998.
[14] A. Peres, "Neumark's theorem and quantum inseparability," *Found. Phys.*, vol. 20, no. 12, pp. 1441–1453, 1990.
[15] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.*, vol. 38, no. 1, pp. 40–95, Mar. 1996.
[16] Y. Nesterov and A. Nemirovski, *Interior-Point Polynomial Algorithms in Convex Programming*. Philadelphia, PA: SIAM, 1994.
[17] F. Alizadeh, "Combinatorial optimization with interior point methods and semi-definite matrices," Ph.D. dissertation, Univ. Minnesota, Minneapolis, Oct. 1991.
[18] G. D. Forney, Jr., "Geometrically uniform codes," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1241–1260, Sept. 1991.
[19] Y. C. Eldar and G. D. Forney, Jr., "On quantum detection and the square-root measurement," *IEEE Trans. Inform. Theory*, vol. 47, pp. 858–872, Mar. 2001.
[20] Y. C. Eldar, A. Megretski, and G. C. Verghese. Optimal detection of symmetric mixed quantum states. [Online] http://www.arXiv.org/abs/quant-ph/0211111.
[21] Y. C. Eldar and H. Bölcskei, "Geometrically uniform frames," *IEEE Trans. Inform. Theory*. [Online]. Available: http://arXiv.org/abs/math.FA/0108096, to be published.
[22] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1996.
[23] M. Ježek, J. Řeháček, and J. Fiurášek. Finding optimal strategies for minimum-error quantum-state discrimination. [Online]. Available: http://www.arXiv.org/abs/quant-ph/0201109.
[24] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri. Distinguishing separable and entangled states. [Online]. Available: http://www.arXiv.org/abs/quant-ph/0112007.
[25] E. M. Rains, "A semidefinite program for distillable entanglement," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2921–2933, Nov. 2001.
[26] K. Audenaert and B. De Moor. Optimizing completely positive maps using semidefinite programming. [Online]. Available: http://www.arXiv.org/abs/quant-ph/0109155.
[27] X. Sun, S. Zhang, Y. Feng, and M. Ying, "Mathematical nature of and a family of lower bounds for the success probability of unambiguous discrimination," *Phys. Rev. A*, vol. 65, Sept. 2002.
[28] M. X. Goemans and D. P. Williamson, "Approximation algorithms for MAX-3-CUT and other problems via complex semidefinite programming," in *Proc. ACM Symp. Theory of Computing*, 2001, pp. 443–452.
[29] J. F. Strum, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimiz. Methods and Software*, vol. 11–12, pp. 625–653, 1999.
[30] D. Peaucelle, D. Henrion, and Y. Labit. Users guide for SeDuMi interface 1.03. [Online]. Available: http://www.laas.fr/peaucell/SeDuMiInt.html.
[31] D. Bertsimas and J. Tsitsiklis, *Introduction to Linear Optimization*. Belmont, MA: Athena Scientific, 1997.
[32] A. Megretski, C.-Y. Kao, U. Jönsson, and A. Rantzer. A guide to IQCβ: Software for robustness analysis. [Online]. Available: http://web.mit.edu/cykao/www.
[33] M. A. Armstrong, *Groups and Symmetry*. New York: Springer-Verlag, 1988.