# Optimal unambiguous discrimination of quantum states

M. A. Jafarizadeh,[1,3,*] M. Rezaei,[2,†] N. Karimi,[1,‡] and A. R. Amiri[1,§]
[1]*The Department of Theoretical Physics and Astrophysics, University of Tabriz, Tabriz, 51664, Iran*
[2]*The Institute for Studies in Theoretical Physics and Mathematics, Tehran 19395-1795, Iran*
[3]*The Research Institute for Fundamental Sciences, Tabriz 51664, Iran*
(Received 17 August 2007; published 16 April 2008)

Unambiguous discrimination between nonorthogonal but linearly independent quantum states is a challenging problem in quantum information processing. In this study, using the connection between Lewenstein-Sanpera decomposition (LSD) and optimal unambiguous state discrimination (OPUSD), an analytic relation for the feasible region of $N$ linearly independent quantum states is presented in terms of inner product of reciprocal states. Then, for the real inner product of states, an exact analytic solution for the OPUSD problem involving an arbitrary number of pure linearly independent quantum states is presented using the Karush-Kuhn-Tucker convex optimization method. In another approach, an analytic relation for the feasible region for an arbitrary number of pure linearly independent quantum states is presented in terms of the inner product of states. To this end, the relevant semidefinite programming task is reduced to a linear programming (LP) one with a feasible region of polygon type which can be solved via simplex method. Moreover, using the obtained feasible region, an exact analytic solution to an OPUSD problem involving two and three pure linearly independent quantum states is provided.

DOI: 10.1103/PhysRevA.77.042314 PACS number(s): 03.67.Hk, 03.65.Ta

## I. INTRODUCTION

In quantum information and quantum computing, the carrier of information is some quantum system and information is encoded in its state. A quantum state describes what we know about a quantum system. Given a single copy of a quantum system which can be prepared in several known quantum states, our aim is to determine in which state the system is. This can be well understood in a communication context where only a single copy of the system is given and only a single-shot measurement is performed. This is in contrast with usual experiments in physics where many copies of a system are measured to get the probability distribution of the system. In quantum state discrimination, no statistics is built since only a single-shot measurement is performed on a single copy of the system. Actually there are fundamental limitations to the precision with which the state of the system can be determined with a single measurement. A fundamental aspect of quantum information theory is that nonorthogonal quantum states cannot be perfectly distinguished. Therefore, a central problem in quantum mechanics is to design measurements optimized to distinguish between a collection of nonorthogonal quantum states. The topic of quantum state discrimination was firmly established in the 1970s by the pioneering work of Helstrom [1], who considered a minimum error discrimination of two known quantum states. In this case, the state identification is probabilistic. Another possible discrimination strategy is the so-called unambiguous state discrimination (USD) where the states are successfully identified with nonunit probability, but without error. USD was originally formulated and analyzed by Ivanovic, Dieks, and Peres [2–4] in 1987. The solution for unambiguous discrimination of two known pure states appearing with arbitrary prior probabilities was obtained by Jaeger and Shimony [5]. Although the two-state problem is well developed, the problem of unambiguous discrimination between multiple quantum states has received considerably less attention. The problem of discrimination among three nonorthogonal states was first considered by Peres and Terno [4]. They developed a geometric approach and applied it numerically on several examples. A different method was considered by Duan and Guo [6] and Sun *et al.* [7]. Chefles [8] showed that a necessary and sufficient condition for the existence of unambiguous measurements for distinguishing between $N$ 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 correct detection of each one of the ensemble states.

Over the past years, semidefinite programming (SDP) has been recognized as a valuable numerical tool for control system analysis and design. In SDP, one minimizes a linear function subject to the constraint that an affine combination of symmetric matrices is positive semidefinite. SDP has been studied (under various names) as far back as the 1940s. Subsequent research in semidefinite programming during the 1990s was driven by applications in combinatorial optimization [9], communications and signal processing [10–12], and other areas of engineering [13]. Although semidefinite programming is designed to be applied in numerical methods, it can be used for analytic computations, too. In the context of quantum computation and quantum information, Barnum, Saks, and Szegedy have reformulated quantum query complexity in terms of a semidefinite program [14], while Jafarizadeh, Mirzaee, and Rezaee have revealed the close con-

*Jafarizadeh@tabrizu.ac.ir
†Karamaty@tabrizu.ac.ir
‡karimi@tabrizu.ac.ir
§amiri@tabrizu.ac.ir

 

nection between Lewenstein-Sanpera decomposition, robustness of entanglement, finite quantum tomography, and semidefinite programming algorithm [15–18].

The problem of finding the optimal measurement to distinguish between a set of quantum states was first formulated as a semidefinite program in 1974 by Holevo, who gave optimality conditions equivalent to the complementary slackness conditions [1]. Recently, Eldar, Megretski, and Verghese showed that the optimal measurements can be found efficiently by solving the dual followed by the use of linear programming [19]. Also in Ref. [20], SDP has been used to show that the standard algorithm implements the optimal set of measurements. All of the abovementioned applications indicate that the method of SDP is very useful. The reason why the area has shown relatively slow progress until recently within the rapidly evolving field of quantum information is that it poses quite formidable mathematical challenges. Except for a handful of very special cases, no general exact solution has been available involving more than two arbitrary states and mostly numerical algorithms are proposed for finding optimal measurements for quantum-state discrimination, where the theory of the semidefinite programming provides a simple check of the optimality of the numerically obtained results.

In this study, we obtain the feasible region in terms of the inner product of the reciprocal states and show that LSD is equivalent to optimal unambiguous state discrimination, and thus one can use LSD to solve the problem of optimal unambiguous state discrimination. Also, for the real inner product of reciprocal states, an exact analytic solution for OPUSD problem involving an arbitrary number of pure linearly independent quantum states is presented using KKT convex optimization method. This method is completely illustrated for three linearly independent states. Moreover, we obtain the feasible region in terms of the inner product of the states which enables us to solve the problem without using reciprocal states. Using the feasible region, an analytic solution for two and three states, and geometrical interpretation for minimum inconclusive result for unambiguous discrimination of two pure states are given. For more than three states, the analytic calculations are very complicated to write down, and therefore we will consider specific cases such as geometrical uniform states and Welch bound equality (WBE) sequences. To solve the problem in general with the use of the following prescription of Refs. [[21–25]], we have reduced it to the LP one, where the computation can be done at a very fast pace and with high precision. Since the analytical relation for feasible region is given here and it is a convex region, one can obtain optimal positive operator-valued measure (POVM) for unambiguous discrimination of any number of linearly independent quantum states numerically.

The organization of the paper is as follows. First, the definition of the unambiguous quantum state discrimination is given. Then, Lewenstein-Sanpera decomposition is studied as an optimal unambiguous discrimination of quantum states, and for the real inner product of reciprocal states, an exact analytic solution for OPUSD problem involving an arbitrary number of pure linearly independent quantum states is presented by using KKT convex optimization method in Sec. III. Next, the unambiguous discrimination of quantum states by introducing feasible region and using linear programming is discussed. Finally, discrimination of nonorthogonal quantum states using approximated linear programming is analyzed. The paper ends with a brief conclusion and three appendixes.

## II. UNAMBIGUOUS QUANTUM STATE DISCRIMINATION

In quantum theory, measurements are represented by positive operator valued measures (POVMs). A measurement is described by a collection $\{M_k\}$ of measurement operators. These operators are acting on the state space of the system being measured. The index $k$ refers to the measurement outcomes that may occur in the experiment. In quantum information theory the measurement operators $\{M_k\}$ are often called Kraus operators [26]. If we define the operator

$$\Pi_k = M_k^\dagger M_k, \tag{1}$$

the probability of obtaining the outcome $k$ for a given state $\rho_i$ is given by $p(k|i)=\mathrm{Tr}(\Pi_k\rho_i)$. Thus, the set of operators $\Pi_k$ is sufficient to determine the measurement statistics.

*Definition of the POVM.* A set of operators $\{\Pi_k\}$ is named a positive operator valued measure if and only if the following two conditions are met: (1) each operator $\Pi_k$ is positive positive $\Leftrightarrow \langle\psi|\Pi_k|\psi\rangle \geq 0, \quad \forall |\psi\rangle$ and (2) the completeness relation is satisfied, i.e.,

$$\sum_k \Pi_k = 1. \tag{2}$$

The elements of $\{\Pi_k\}$ are called effects or POVM elements. On its own, a given POVM $\{\Pi_k\}$ is enough to give complete knowledge about the probabilities of all possible outcomes; measurement statistics is the only item of interest. Consider a set of known states $\rho_i$, $i=1,\ldots,N$ with their prior probabilities $\eta_i$. We are looking for a measurement that either identifies a state uniquely (conclusive result) or fails to identify it (inconclusive result). The goal is to minimize the probability of inconclusive results. The measurements involved are typically generalized measurements. A measurement described by a POVM $\{\Pi_k\}_{k=1}^N$ is called an unambiguous state discrimination measurement (USDM) on the set of states $\{\rho_i\}_{i=1}^N$ if and only if the following conditions are satisfied. (1) The POVM contains the elements $\{\Pi_k\}_{k=0}^N$ where $N$ is the number of different signals in the set of states. The element $\Pi_0$ is related to an inconclusive result, while the other elements correspond to an identification of one of the states $\rho_i$, $i=1,\ldots,N$. (2) No states are wrongly identified, that is, $\mathrm{Tr}(\rho_i\Pi_k)=0 \quad \forall i\neq k \quad i,k=1,\ldots,N$.

Each USD measurement gives rise to a failure probability, that is, the rate of inconclusive result. This can be calculated as

$$Q = \sum_i \eta_i \mathrm{Tr}(\rho_i\Pi_0). \tag{3}$$

The success probability can be calculated as

$$P = 1 - Q = \sum_i \eta_i \mathrm{Tr}(\rho_i \Pi_i). \tag{4}$$

A measurement described by a POVM $\{\Pi_k^{\mathrm{opt}}\}$ is called an optimal unambiguous state discrimination measurement (OUSDM) on a set of states $\{\rho_i\}$ with the corresponding prior probabilities $\{\eta_i\}$ if and only if the following conditions are satisfied. (1) The POVM $\{\Pi_k^{\mathrm{opt}}\}$ is a USD measurement on $\{\rho_i\}$. (2) The probability of inconclusive result is minimal, that is, $Q(\{\Pi_k^{\mathrm{opt}}\}) = \min Q(\{\Pi_k\})$, where the minimum is taken over all USDM.

Unambiguous state discrimination is an error-free discrimination. This implies a strong constraint on the measurement. Suppose that a quantum system is prepared in a pure quantum state drawn from a collection of given states $\{|\psi_i\rangle, 1 \le i \le N\}$ in $d$-dimensional complex Hilbert space $\mathcal{H}$ with $d \ge N$. These states span a subspace $\mathcal{U}$ of $\mathcal{H}$. In order to detect the state of the system, a measurement is constructed comprising $N+1$ measurement operators $\{\Pi_i, 0 \le i \le N\}$. Given that the state of the system is $|\psi_i\rangle$, the probability of obtaining outcome $k$ is $\langle\psi_i|\Pi_k|\psi_i\rangle$. Therefore, in order to ensure that each state is either correctly detected or an inconclusive result is obtained, we must have

$$\langle\psi_i|\Pi_k|\psi_i\rangle = p_i \delta_{ik}, \quad 1 \le i,k \le N, \tag{5}$$

for some $0 \le p_i \le 1$. Since $\Pi_0 = I_d - \Sigma_{i=1}^N \Pi_i$, we have $\langle\psi_i|\Pi_0|\psi_i\rangle = 1 - p_i$. So a system with given state $|\psi_i\rangle$, the state of the system is correctly detected with probability $p_i$ and an inconclusive result is obtained with probability $1-p_i$. It was shown in Ref. [8] that Eq. (5) is satisfied if and only if the vectors $|\psi_i\rangle$ are linearly independent, or equivalently, $\dim \mathcal{U} = N$. Therefore, we will take this assumption throughout the paper. In this case, we may choose [27]

$$\Pi_i = p_i |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|, \quad 1 \le i \le N, \tag{6}$$

where the vectors $|\tilde{\psi}_i\rangle \in \mathcal{U}$ are the reciprocal states associated with the states $|\psi_i\rangle$, i.e., there are unique vectors in $\mathcal{U}$ such that

$$\langle\tilde{\psi}_i|\psi_k\rangle = \delta_{ik}, \quad 1 \le i,k \le N. \tag{7}$$

With $\Phi$ and $\tilde{\Phi}$ we denote the matrices such that their columns are $|\psi_i\rangle$ and $|\tilde{\psi}_i\rangle$, respectively. Then, one can show that $\tilde{\Phi}$ is

$$\tilde{\Phi} = \Phi(\Phi^*\Phi)^{-1}. \tag{8}$$

Since the vectors $|\psi_i\rangle$, $i = 1, \ldots, N$ are linearly independent, $\Phi^*\Phi$ is always invertible. Alternatively,

$$\tilde{\Phi} = (\Phi\Phi^*)^{\ddagger}\Phi, \tag{9}$$

so that

$$|\tilde{\psi}_i\rangle = (\Phi\Phi^*)^{\ddagger}|\psi_i\rangle, \tag{10}$$

where $(\cdots)^{\ddagger}$ denotes the Moore-Penrose pseudoinverse [28,29]. The inverse is taken on the subspace spanned by the columns of the matrix. If the state $|\psi_i\rangle$ is prepared with prior probability $\eta_i$, then the total probability of correctly detecting the state is

$$P = \sum_{i=1}^N \eta_i \langle\psi_i|\Pi_i|\psi_i\rangle = \sum_{i=1}^N \eta_i p_i \tag{11}$$

and the probability of the inconclusive result is given by

$$Q = 1 - P = \sum_{i=1}^N \eta_i \langle\psi_i|\Pi_0|\psi_i\rangle = 1 - \sum_{i=1}^N \eta_i p_i. \tag{12}$$

In general, an optimal measurement for a given strategy depends on the quantum states and the prior probabilities of their appearance. In the unambiguous discrimination for a given strategy and a given ensemble of states, the goal is to find a measurement which minimizes the inconclusive result. In fact, it is known that USD (of both pure and mixed states) is a convex optimization problem. Mathematically, this means that the quantity which is to be optimized as well as the constraints on the unknowns, are convex functions. Practically, this implies that the optimal solution can be computed in an extremely efficient way. This is therefore a very useful tool. Nevertheless, our aim is to understand the structure of USD in order to relate it with neat and relevant quantities and to find feasible region for numerical and analytic solutions. Thus, by using SDP we determine feasible region via reciprocal states and reduce the theory of the SDP to a linear programming one with a feasible region of polygon type which can be solved via simplex method precisely or approximately.

## III. LEWENSTEIN-SANPERA DECOMPOSITION (LSD) AS AN OPTIMAL UNAMBIGUOUS DISCRIMINATION

The idea of Refs. [30,31] is based on the method of subtracting projections on product vectors from a given state, that is, for a given density matrix $\rho$ and any set $V = \{|\tilde{\psi}_i\rangle\}$ of the states belonging to the range of $\rho$, one can subtract a density matrix $\rho' = \Sigma_i p_i \Pi_i$ (not necessarily normalized) with $p_i \ge 0$ such that $\delta\rho = \rho - \rho' \ge 0$, in the sense that $\mathrm{Tr}(\rho') \le 1$. In the following we recall some important definitions and theorems about LSD.

*Definition 1.* A non-negative parameter $p$ is called maximal with respect to a (not necessarily normalized) density matrix $\rho$ and the projection operator $\Pi = |\tilde{\psi}\rangle\langle\tilde{\psi}|$ if and only if $\rho - p\Pi \ge 0$, and for every $\epsilon \ge 0$, the matrix $\rho - (p+\epsilon)\Pi$ is not positive definite. The maximal $p$ thus determines the maximal contribution of $\Pi$ that can be subtracted from $\rho$ maintaining the non-negativity of the difference.

*Lemma 1.* $p$ is maximal with respect to $\rho$ and $\Pi = |\tilde{\psi}\rangle\langle\tilde{\psi}|$ if and only if: (1) $|\tilde{\psi}\rangle \notin \mathcal{R}(\rho)$, then $p = 0$, (2) $|\tilde{\psi}\rangle \in \mathcal{R}(\rho)$, then

$$0 < p = \frac{1}{\langle\tilde{\psi}|\rho^{-1}|\tilde{\psi}\rangle}, \tag{13}$$

where $\mathcal{R}(\rho)$, denotes the range of $\rho$, i.e., the set of all $|\psi\rangle \in \mathcal{H}$ for which $\exists |\phi\rangle \in \mathcal{H}$ such that $|\psi\rangle = \rho|\phi\rangle$.

*Definition 2.* We say that a pair of non-negative $(p_1, p_2)$ is maximal with respect to $\rho$ and a pair of projection operators $\Pi_1 = |\tilde{\psi}_1\rangle\langle\tilde{\psi}_1|$, $\Pi_2 = |\tilde{\psi}_2\rangle\langle\tilde{\psi}_2|$ if and only if $\rho - p_1\Pi_1 - p_2\Pi_2 \ge 0$, $p_1$ is maximal with respect to $\rho - p_2\Pi_2$ and to the pro-

jector $\Pi_1$, $p_2$ is maximal with respect to $\rho - p_1\Pi_1$ and to the projector $\Pi_2$, and the sum $p_1+p_2$ is maximal.

*Lemma 2.* A pair $(p_1,p_2)$ is maximal with respect to $\rho$ and a pair of projectors $(\Pi_1,\Pi_2)$ if and only if (a) $|\tilde{\psi}_1\rangle,|\tilde{\psi}_2\rangle$ do not belong to $\mathcal{R}(\rho)$ then $p_1=p_2=0$, (b) $|\tilde{\psi}_1\rangle$ does not belong to $\mathcal{R}(\rho)$, while $|\tilde{\psi}_2\rangle\in\mathcal{R}(\rho)$ then $p_1=0$, $p_2=\langle\tilde{\psi}_2|\frac{1}{\rho}|\tilde{\psi}_2\rangle^{-1}$, (c) $|\tilde{\psi}_1\rangle,|\tilde{\psi}_2\rangle\in\mathcal{R}(\rho)$ and $\langle\tilde{\psi}_1|\frac{1}{\rho}|\tilde{\psi}_2\rangle=0$ then $p_i=\langle\tilde{\psi}_i|\frac{1}{\rho}|\tilde{\psi}_i\rangle$, $i=1,2$, and (d) finally, if $|\tilde{\psi}_1\rangle,|\tilde{\psi}_2\rangle\in\mathcal{R}(\rho)$ and $\langle\tilde{\psi}_1|\frac{1}{\rho}|\tilde{\psi}_2\rangle\neq 0$ then

$$p_1=\frac{1}{D}\left(\langle\tilde{\psi}_2|\frac{1}{\rho}|\tilde{\psi}_2\rangle-\left|\langle\tilde{\psi}_1|\frac{1}{\rho}|\tilde{\psi}_2\rangle\right|\right), \tag{14}$$

$$p_2=\frac{1}{D}\left(\langle\tilde{\psi}_1|\frac{1}{\rho}|\tilde{\psi}_1\rangle-\left|\langle\tilde{\psi}_1|\frac{1}{\rho}|\tilde{\psi}_2\rangle\right|\right), \tag{15}$$

where $D=\langle\tilde{\psi}_1|\frac{1}{\rho}|\tilde{\psi}_1\rangle\langle\tilde{\psi}_2|\frac{1}{\rho}|\tilde{\psi}_2\rangle-|\langle\tilde{\psi}_1|\frac{1}{\rho}|\tilde{\psi}_2\rangle|^2$.

*Lemma 3.* Let a Hermitian density matrix $\rho$ has a decomposition of the form $\rho=\rho'+(1-p)\delta\rho$, where $\rho'$ is a part of density operator $\rho$ which has the structure $\rho'=\Sigma_{i=1}^{N}p_i\Pi_i$, with $\Pi_i$ being projection operator onto state $|\tilde{\psi}_i\rangle$ and $\Sigma_{i=1}^{N}p_i=p$. Then the set of $\{p_i\}$, which are maximal with respect to the density matrix $\rho$ and the set of the projection operators $\{\Pi_i\}$ form a manifold which generically has dimension $N-1$ and is determined by the following equation:

$$1-\sum_i D_i p_i+\sum_{i<j}D_{ij}p_ip_j-\sum_{i<j<k}D_{ijk}p_ip_jp_k + \cdots+(-1)^N\sum_{i_1<i_2\ldots<i_N}p_{i_1}p_{i_2}\cdots p_{i_N}D_{i_1i_2\ldots i_N}=0, \tag{16}$$

where the set of $\{D_{i_1,i_2\ldots i_N}\}$ are the subdeterminants (minors) of matrix $D$ defined by

$$D=\begin{pmatrix}\tilde{a}_{11} & \tilde{a}_{12} & \cdots & \tilde{a}_{1N}\\ \tilde{a}_{21} & \tilde{a}_{22} & \cdots & \tilde{a}_{2N}\\ \vdots & \vdots & \ddots & \vdots\\ \tilde{a}_{N1} & \tilde{a}_{N2} & \cdots & \tilde{a}_{NN}\end{pmatrix}, \tag{17}$$

with $\tilde{a}_{ij}:=\langle\tilde{\psi}_i|\frac{1}{\rho}|\tilde{\psi}_j\rangle$. Equation (16) determines feasible region via reciprocal states, that is, it gives the domain of acceptable values of $p_i$. One way to derive Eq. (16) is using semidefinite programming [15,17].

In the rest of this section, we show that optimal unambiguous discrimination for $N$ linearly independent states can be reduced to the LSD method. Suppose a quantum system is prepared in a state secretly drawn from a known set $|\psi_1\rangle,\ldots,|\psi_N\rangle$, where each $|\psi_i\rangle$ is a pure state in the Hilbert space $\mathcal{H}$. In order to discriminate $|\psi_1\rangle,\ldots,|\psi_N\rangle$ unambiguously, one can construct a most general POVM consisting of $N+1$ elements $\Pi_0,\Pi_1,\ldots,\Pi_N$ such that

$$\Pi_i\geq 0,\quad i=0,1,\ldots,N,\quad \text{and}\quad \sum_{i=0}^{N}\Pi_i=I, \tag{18}$$

where $I$ denotes the identity matrix in $\mathcal{H}$. Each element $\Pi_i$, $i=1,\ldots,N$ of POVM corresponds to an identification of the corresponding state $|\psi_i\rangle$, while $\Pi_0$ corresponds to the inconclusive answer. For the sake of simplicity, we often specify only $\Pi_1,\ldots,\Pi_N$ for a given POVM since the left element $\Pi_0$ is uniquely determined by

$$\Pi_0=I-\sum_{i=1}^{N}\Pi_i=I-\sum_{i=1}^{N}p_i|\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|. \tag{19}$$

The goal of LSD is maximizing $p_i$'s such that $\Sigma_{i=1}^{N}p_i$ is maximized. With assuming that density matrix in LSD method is equal to identity we obtain LSD as follows:

$$\rho-\sum_{i=1}^{N}\Pi_i=I-\sum_{i=1}^{N}p_i|\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|. \tag{20}$$

Through the comparison of this relation with Eq. (19), it is clear that we can maximize success probability by using LSD method or minimize the inconclusive probability. Then, LSD can be considered the same as OUSDM and can be used to obtain the elements of the optimal POVM.

### A. Optimal unambiguous discrimination of two states using LSD method

Suppose that two pure states $|\psi_1\rangle$ and $|\psi_2\rangle$ with arbitrary prior probabilities $\eta_1$ and $\eta_2$ are given. In order to obtain optimal POVM set for these two states by using LSD, we use lemma 2 of LSD for two states $|\tilde{\psi}_1\rangle$ and $|\tilde{\psi}_2\rangle$. Let the corresponding density matrix in Hilbert space $\mathcal{H}$ be the identity operator, i.e., $\rho=I$ and $\Pi_1'=\eta_1|\tilde{\psi}_1\rangle\langle\tilde{\psi}_1|$ and $\Pi_2'=\eta_2|\tilde{\psi}_2\rangle\langle\tilde{\psi}_2|$. Then, a pair $(p_1,p_2)$ is maximal with respect to $\rho$ and the pair of operators $\Pi_1'$ and $\Pi_2'$ if and only if $\rho-p_1\eta_1|\tilde{\psi}_1\rangle\langle\tilde{\psi}_1|-p_2\eta_2|\tilde{\psi}_2\rangle\langle\tilde{\psi}_2|\geq 0$. Therefore, from lemma 2 of LSD, the maximal pair $(p_1,p_2)$ is given by

$$p_1=\frac{\tilde{a}_{22}-\sqrt{\frac{\eta_2}{\eta_1}}|\tilde{a}_{12}|}{\tilde{a}_{11}\tilde{a}_{22}-|\tilde{a}_{12}|^2},\quad p_2=\frac{\tilde{a}_{11}-\sqrt{\frac{\eta_1}{\eta_2}}|\tilde{a}_{12}|}{\tilde{a}_{11}\tilde{a}_{22}-|\tilde{a}_{12}|^2}, \tag{21}$$

where $\tilde{a}_{ij}=\langle\tilde{\psi}_i|\tilde{\psi}_j\rangle$. If the condition

$$\frac{|\tilde{a}_{12}|}{\tilde{a}_{11}}\leq\sqrt{\frac{\eta_2}{\eta_1}}\leq\frac{\tilde{a}_{22}}{|\tilde{a}_{12}|} \tag{22}$$

is held, then Eq. (21) is the optimal solution.

If $\eta_1<(\frac{|\tilde{a}_{12}|}{\tilde{a}_{22}})^2\eta_2$, then the optimal solution reads

$$p_1=0,\quad p_2=\frac{1}{\tilde{a}_{22}}. \tag{23}$$

If $\eta_2<(\frac{|\tilde{a}_{12}|}{\tilde{a}_{11}})^2\eta_1$, then the optimal solution is given by

$$p_2=0,\quad p_1=\frac{1}{\tilde{a}_{11}}. \tag{24}$$

For more than two linearly independent states, as the LSD method leads to a set of coupled equations which are in general difficult to solve, we embark on KKT (see Appendix B) method which makes the problem extremely easy to solve. Since this condition is necessary and sufficient, the answer will be exactly optimal for unambiguous discrimination.

### B. Analytical calculation of optimal POVM for unambiguous discrimination of quantum states using KKT method

In this section, for the real inner product of reciprocal states, an exact analytic solution for OPUSD problem involving an arbitrary number of pure linearly independent quantum states is presented by using KKT convex optimization method. The KKT conditions for unambiguous discrimination of $N$ linearly independent states are given by (see Appendix C)

$$
\begin{aligned}
&I-\sum_{i=1}^{N} p_i|\tilde{\psi}_i\rangle\langle\tilde{\psi}_i| \geq 0, \quad p_i \geq 0,\\
&\left(I-\sum_{i=1}^{N} p_i|\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|\right)\\
&X=X\left(I-\sum_{i=1}^{N} p_i|\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|\right)=0,\\
&z_i p_i=0,\\
&\mathrm{Tr}(X|\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|)=z_i+\eta_i, \quad \eta_i \geq 0\\
&i=1,\ldots,N
\end{aligned}
\tag{25}
$$

If $\tilde{a}_{ij}$ is real, then using KKT conditions one can show that

$$
\begin{pmatrix}
1-p_1\tilde{a}_{11} & -p_2\tilde{a}_{12} & \cdots & -p_N\tilde{a}_{1N}\\
-p_1\tilde{a}_{12} & 1-p_2\tilde{a}_{22} & \cdots & -p_N\tilde{a}_{2N}\\
\vdots & \vdots & \ddots & \vdots\\
-p_1\tilde{a}_{1N} & -p_2\tilde{a}_{2N} & \cdots & 1-p_N\tilde{a}_{NN}
\end{pmatrix}
\times
\begin{pmatrix}
x_{11} & x_{12} & \cdots & x_{1N}\\
x_{12} & x_{22} & \cdots & x_{2N}\\
\vdots & \vdots & \ddots & \vdots\\
x_{1N} & x_{2N} & \cdots & x_{NN}
\end{pmatrix}=0. \tag{26}
$$

If $p_i \neq 0$ ($\underline{i=1},\ldots,N$), using Eq. (25) with $x_{ij}=\langle\tilde{\psi}_i|X|\tilde{\psi}_j\rangle = (-1)^{\epsilon_i+\epsilon_j}\sqrt{\eta_i\eta_j}$, where $\epsilon_i=0$ or 1 and $\epsilon_i+\epsilon_j=0$ or 1 module 1, one can obtain the optimal solutions. Then, $X$ is a rank 1 matrix. Consequently, the Eq. (26) can be written as

$$
\begin{pmatrix}
1-p_1\tilde{a}_{11} & -p_2\tilde{a}_{12} & \cdots & -p_N\tilde{a}_{1N}\\
-p_1\tilde{a}_{12} & 1-p_2\tilde{a}_{22} & \cdots & -p_N\tilde{a}_{2N}\\
\vdots & \vdots & \ddots & \vdots\\
-p_1\tilde{a}_{1N} & -p_2\tilde{a}_{2N} & \cdots & 1-p_N\tilde{a}_{NN}
\end{pmatrix}
\begin{pmatrix}
x_{11}\\ x_{12}\\ \vdots\\ x_{1N}
\end{pmatrix}=0. \tag{27}
$$

Or, equivalently,

$$
\begin{aligned}
&\begin{pmatrix}
\tilde{a}_{11} & \tilde{a}_{12} & \cdots & \tilde{a}_{1N}\\
\tilde{a}_{21} & \tilde{a}_{22} & \cdots & \tilde{a}_{2N}\\
\vdots & \vdots & \ddots & \vdots\\
\tilde{a}_{N1} & \tilde{a}_{N2} & \cdots & \tilde{a}_{NN}
\end{pmatrix}
\begin{pmatrix}
x_{11} & 0 & \cdots & 0\\
0 & x_{12} & \cdots & 0\\
\vdots & \vdots & \ddots & \vdots\\
0 & 0 & \cdots & x_{1N}
\end{pmatrix}
\begin{pmatrix}
p_1\\ p_2\\ \vdots\\ p_N
\end{pmatrix}\\
&=\begin{pmatrix}
x_{11}\\ x_{12}\\ \vdots\\ x_{1N}
\end{pmatrix}.
\end{aligned}
\tag{28}
$$

If the first condition in Eq. (25) is satisfied, using above equation one can show that the optimal solution for $p_i$ is given by

$$
\begin{aligned}
p_i=&\frac{1}{(-1)^{\epsilon_i}\sqrt{\eta_i}\,\det(D)}\\
&\times\det\begin{pmatrix}
\tilde{a}_{11} & \cdots & \tilde{a}_{1,i-1} & (-1)^{\epsilon_1}\sqrt{\eta_1} & \tilde{a}_{1,i+1} & \cdots & \tilde{a}_{1N}\\
\tilde{a}_{21} & \cdots & \tilde{a}_{2,i-1} & (-1)^{\epsilon_2}\sqrt{\eta_2} & \tilde{a}_{2,i+1} & \cdots & \tilde{a}_{2N}\\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots\\
\tilde{a}_{N1} & \cdots & \tilde{a}_{N,i-1} & (-1)^{\epsilon_N}\sqrt{\eta_N} & \tilde{a}_{N,i+1} & \cdots & \tilde{a}_{NN}
\end{pmatrix},
\end{aligned}
\tag{29}
$$

using the fact that the matrix whose components are $\tilde{a}_{ij}$ is inverse matrix of the one whose components are $a_{ij}$, we can rewrite the probability $p_i$ in a simpler form

$$
p_i=(-1)^{i+1}\sum_{j=1}^{N}(-1)^{\epsilon_i+\epsilon_j}\sqrt{\frac{\eta_j}{\eta_i}}a_{ij}, \tag{30}
$$

where $(-1)^{\epsilon_i+\epsilon_j}a_{ij}=\pm|a_{ij}|$ and the signs of $\pm|a_{ij}|$ are determined such that $p_i$ satisfy the condition $0\leq p_i\leq 1$. Then the success probability with respect to $\tilde{a}_{ij}$ components is given by

$$
\begin{aligned}
P=&-\frac{1}{\det(D)}\\
&\times\det\begin{pmatrix}
0 & (-1)^{\epsilon_1}\sqrt{\eta_1} & \cdots & (-1)^{\epsilon_N}\sqrt{\eta_N}\\
(-1)^{\epsilon_1}\sqrt{\eta_1} & & & \\
\vdots & & D & \\
(-1)^{\epsilon_N}\sqrt{\eta_N} & & &
\end{pmatrix},
\end{aligned}
\tag{31}
$$

and with respect to $a_{ij}$ components is as follows

$$
P=\sum_{i=1}^{N}\eta_i p_i, \tag{32}
$$

where values of $\epsilon_i$ and $p_i$ are determined by the $0\leq p_i\leq 1$ condition.

If the first condition of KKT is not satisfied, for a specified $i$, $p_i$ does not lie in the feasible region ($p_i\leq 0$ or $p_i\geq 1$). In this case, one can omit the $j$th row and $j$th column in the square $N\times N$ matrices and $j$th row in the row matrices in Eqs. (27) and (17), and find the optimal solutions with $p_i=0$. After this reduction, if any other $p_i$ does not lie in the

feasible region, the same procedure will be repeated.

In general, for various prior probabilities, different solutions are obtained. One of them can discriminate all of the states provided that the KKT conditions are satisfied and the inner product of the reciprocal state is real. Other solutions can only discriminate less than $N$ quantum states.

*Example 1.* In this example we consider reciprocal independent states $\{\tilde{\psi}_i\}_{i=1}^N$ such that

$$\langle\tilde{\psi}_i|\tilde{\psi}_j\rangle = a \quad \text{for} \quad i \neq j, \quad \langle\tilde{\psi}_i|\tilde{\psi}_j\rangle = 1 \quad \text{for} \quad i = j. \tag{33}$$

Thus, for optimal distinguishing of independent vectors that are prepared with equal probabilities, we minimize the inconclusive result $\Pi_0$ given by

$$\Pi_0 = I - \sum_{i=1}^{N} p_i|\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|. \tag{34}$$

Now using Eq. (29) we show that all of $p_i$'s are equal, i.e., all of the elements of POVM possess equal probabilities, i.e.,

$$p_1 = p_2 = \cdots = p_N = \frac{1}{1 + a(N-1)}. \tag{35}$$

Similar to geometrically uniform (GU) states, one can prove that the optimal distinguishability corresponds to equal measurement probabilities $p_i=p$ for all $i$ such that $p$ is equal to the inverse of maximum eigenvalue of the corresponding frame operator. In order to prove this fact, we define frame operator as

$$S = \sum_{i=1}^{N} |\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|, \tag{36}$$

such that

$$S_{kl} = \sum_{i=1}^{N} \langle k|\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|l\rangle. \tag{37}$$

Then, $S=AA^\dagger$ with $A_{ki}=\langle k|\tilde{\psi}_i\rangle$. On the other hand, the Gramm matrix is defined as

$$\tilde{G} = A^\dagger A. \tag{38}$$

Again, one can easily show that $S$ and $G$ have equal eigenvalues, thus we evaluate eigenvalues of Gramm matrix. The Gramm matrix $\tilde{G}$ can be written as

$$\tilde{G} = I + a(C-1), \tag{39}$$

such that its eigenvalues are given by

$$1 + a(N-1), \quad 1 - a, \tag{40}$$

where $a<1$. Therefore, maximum eigenvalue of frame operator is equal to $1+a(N-1)$ and thus $p$ is given by $[1+a(N-1)]^{-1}$.

## IV. UNAMBIGUOUS DISCRIMINATION OF QUANTUM STATES USING LINEAR PROGRAMMING

The method presented in this section seems to be a powerful method which enables us to discriminate $N$ linearly independent pure quantum states by some well-known numerical methods. We naturally want the probabilities $p_i$ to be as large as possible in order to increase the detection probabilities, but their values are bounded by the demand of positivity of $\Pi_0$. Let

$$|\psi\rangle := \sum_{i=1}^{N} a_i|\psi_i\rangle + |\psi^\perp\rangle$$

with $\langle\psi_i|\psi^\perp\rangle=0$ and where the normalization condition is defined by

$$\langle\psi|\psi\rangle = \sum_{i,j=1}^{N} a_i^* a_j\langle\psi_i|\psi_j\rangle + \langle\psi^\perp|\psi^\perp\rangle = 1. \tag{41}$$

The relation (41) can be written as

$$\sum_{i,j=1}^{N} a_i^* a_j G_{ij} \leq 1, \tag{42}$$

where $G_{ij}=\langle\psi_i|\psi_j\rangle$ are matrix elements of the Gramm matrix. Now the following vector representation is defined:

$$X := \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix}. \tag{43}$$

Then Eq. (42) can be rewritten as the following constraint:

$$X^\dagger G X \leq 1. \tag{44}$$

Positivity of $\mathrm{Tr}(\Pi_0)$ gives

$$\sum_i |a_i|^2 p_i \leq 1.$$

This last condition is a decisive one that actually determines the domain of acceptable values of $p_i$. This result leads us to the optimization problem defined as

$$\text{maximize} \ \sum_i \eta_i p_i \leq 1, \quad \text{such that} \ \begin{cases} \sum_i |a_i|^2 p_i \leq 1 \\ X^\dagger G X \leq 1. \end{cases} \tag{45}$$

If we write $p_i=\lambda'\xi_i$ with $0\leq\xi_i\leq 1$, then we will have

$$\sum_i |a_i|^2 p_i = \lambda' \sum_i |a_i|^2 \xi_i \leq 1 \Rightarrow \lambda' \leq \frac{1}{\sum_i |a_i|^2 \xi_i}. \tag{46}$$

Then, we must compute the maximum value of $\Sigma_i|a_i|^2\xi_i$ such that $\lambda'$ possesses its lowest possible value. That is,

$$\text{maximize} \ \sum_i |a_i|^2 \xi_i, \quad \text{such that} \ X^\dagger G X \leq 1. \tag{47}$$

By defining a diagonal matrix as follows:

$$D:=\begin{pmatrix} \xi_1 & 0 & \cdots & 0 \\ 0 & \xi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \xi_N \end{pmatrix}, \tag{48}$$

we can write $\Sigma_i |a_i|^2 \xi_i = X^\dagger D X$. This leads us to the following optimization problem:

$$\text{maximize } X^\dagger D X, \quad \text{such that } X^\dagger G X \leq 1. \tag{49}$$

Now, let $Y=\sqrt{D}X$. Then, Eq. (49) can be rewritten as

$$\text{maximize } Y^\dagger Y, \quad \text{such that } Y^\dagger D^{-1/2} G D^{-1/2} Y \leq 1. \tag{50}$$

Suppose $D^{-1/2}GD^{-1/2}=\hat{G}$. Then, we have

$$\text{maximize } Y^\dagger Y, \quad \text{such that } Y^\dagger \hat{G} Y \leq 1. \tag{51}$$

The determinant of $\tilde{G}-\lambda I=0$ identifies the feasible region provided that $\lambda$ coincides with $\lambda'$. Now, in order to show the ability of the method used in this study, we calculate the optimal failure probabilities corresponding to unambiguous discrimination of two and three linearly independent states with arbitrary prior probabilities. In the simple case of two pure states $|\psi_1\rangle$ and $|\psi_2\rangle$ with arbitrary prior probabilities $\eta_1$ and $\eta_2$, $\hat{G}$ is given by

$$\hat{G} = \begin{pmatrix} \frac{1}{\sqrt{\xi_1}} & 0 \\ 0 & \frac{1}{\sqrt{\xi_2}} \end{pmatrix} \begin{pmatrix} 1 & a_{12} \\ a_{12}^* & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{\xi_1}} & 0 \\ 0 & \frac{1}{\sqrt{\xi_2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\xi_1} & \sqrt{\frac{1}{\xi_1 \xi_2}} a_{12} \\ \sqrt{\frac{1}{\xi_1 \xi_2}} a_{12}^* & \frac{1}{\xi_2} \end{pmatrix}, \tag{52}$$

where $a_{ij}=\langle \psi_i | \psi_j \rangle$.

The characteristic equation is given by

$$\lambda^2 - \left(\frac{1}{\xi_1}+\frac{1}{\xi_2}\right)\lambda + \frac{1}{\xi_1 \xi_2}(1-|a_{12}|^2) = 0. \tag{53}$$

Thus,

$$\xi_1 \xi_2 \lambda^2 - \lambda(\xi_1+\xi_2) + (1-|a_{12}|^2) = 0. \tag{54}$$

In the feasible region we have $p_1=\lambda\xi_1$ and $p_2=\lambda\xi_2$, then Eq. (54) is equivalent to

$$p_1 p_2 - (p_1+p_2) + (1-|a_{12}|^2) = 0. \tag{55}$$

This equation determines the feasible region (see Fig. 1). To calculate the minimum probability of inconclusive result, we put the gradient of line $\eta_1 p_1 + \eta_2 p_2$ equal to the gradient of Eq. (27), then we will have

$$p_2 - 1 = \Lambda \eta_1 \text{ and } p_1 - 1 = \Lambda \eta_2 \tag{56}$$

By substituting Eq. (28) into Eq. (27), we obtain

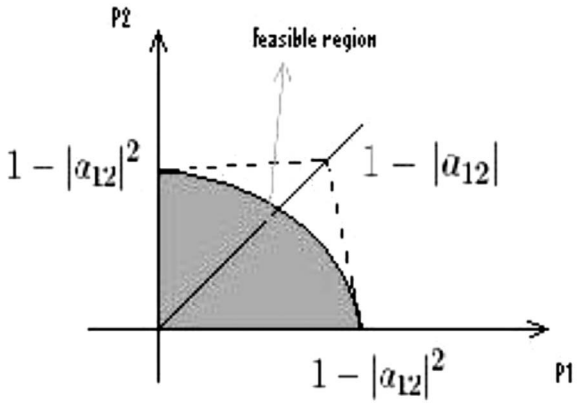


FIG. 1. Feasible region for unambiguous discrimination of two linearly independent states is shown by the shaded region and the approximately feasible region is showed by a polygon.

$$(1+\Lambda\eta_1)(1+\Lambda\eta_2) - (2+\Lambda) + (1-|a_{12}^2|) = 0, \tag{57}$$

which implies that $\Lambda = \pm \frac{|a_{12}|}{\sqrt{\eta_1 \eta_2}}$. Substituting $\Lambda$ into Eq. (28) gives the following solutions:

$$p_1 = 1 \pm \sqrt{\frac{\eta_2}{\eta_1}}|a_{12}| \text{ and } p_2 = 1 \pm \sqrt{\frac{\eta_1}{\eta_2}}|a_{12}|. \tag{58}$$

Since $p_i \leq 1$, thus we conclude that

$$p_1 = 1 - \sqrt{\frac{\eta_2}{\eta_1}}|a_{12}| \text{ and } p_2 = 1 - \sqrt{\frac{\eta_1}{\eta_2}}|a_{12}|. \tag{59}$$

By choosing $(-1)^{\epsilon_i+\epsilon_j}a_{ij}=-|a_{ij}|$ in Eq. (30) we can easily find relations (21).

From the positivity of $p_1$ and $p_2$, we have

$$|a_{12}| \leq \sqrt{\frac{\eta_2}{\eta_1}} \leq \frac{1}{|a_{12}|}.$$

In this case, the minimum probability of inconclusive result is

$$Q = 1 - (\eta_1 p_1 + \eta_2 p_2) = 2\sqrt{\eta_1 \eta_2}|a_{12}|. \tag{60}$$

If $\sqrt{\frac{\eta_2}{\eta_1}} < |a_{12}|$, then $p_2=0$ and if $\frac{1}{|a_{12}|} < \sqrt{\frac{\eta_2}{\eta_1}}$, then $p_1=0$. Then, by using Eq. (55), we obtain

$$\begin{aligned} Q &= \eta_2 + \eta_1 |a_{12}|^2 \text{ if } \sqrt{\frac{\eta_2}{\eta_1}} < |a_{12}|, \\ Q &= 2\sqrt{\eta_1 \eta_2}|a_{12}| \text{ if } |a_{12}| \leq \sqrt{\frac{\eta_2}{\eta_1}} \leq \frac{1}{|a_{12}|}, \\ Q &= \eta_1 + \eta_2 |a_{12}|^2 \text{ if } \frac{1}{|a_{12}|} < \sqrt{\frac{\eta_2}{\eta_1}}. \end{aligned} \tag{61}$$

At this stage, geometrical interpretation of optimal unambiguous discrimination of two pure states on Bloch sphere is discussed. One can show that the minimum inconclusive result for unambiguous discrimination of two pure states is equivalent to the distance between sphere center and the line connecting $\rho_1$ to $\rho_2$ (see Fig. 2). Density matrix for a pure qubit state is defined in the Bloch form as

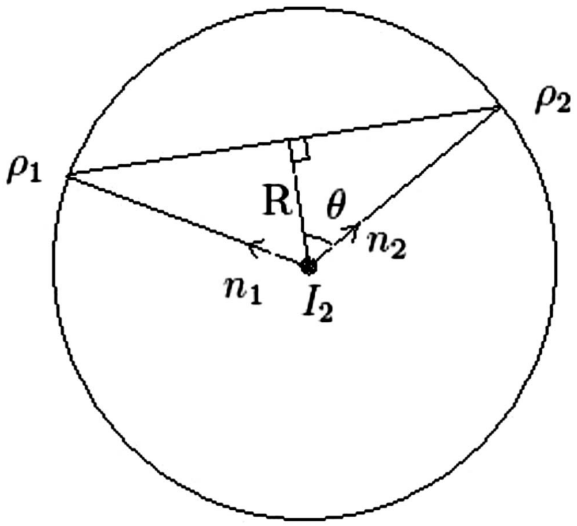


FIG. 2. Unambiguous discrimination of two pure states in Bloch sphere.

$$\rho_1 = \frac{1}{2}(I_2 + \vec{n}_1 \cdot \vec{\sigma}) = |n_1\rangle\langle n_1|, \quad |n_1\rangle = \begin{pmatrix} \cos\frac{\theta_1}{2} \\ e^{i\phi_1}\sin\frac{\theta_1}{2} \end{pmatrix},$$

$$\rho_2 = \frac{1}{2}(I_2 + \vec{n}_2 \cdot \vec{\sigma}) = |n_2\rangle\langle n_2|, \quad |n_2\rangle = \begin{pmatrix} \cos\frac{\theta_2}{2} \\ e^{i\phi_2}\sin\frac{\theta_2}{2} \end{pmatrix}, \tag{62}$$

where $\theta_i$ and $\phi_i$ for $i=1,2$ are spherical coordinates.

For unambiguous discrimination we will have

$$\mathrm{Tr}(\Pi_1\rho_2) = 0, \quad \mathrm{Tr}(\Pi_2\rho_1) = 0, \tag{63}$$

where $\Pi_1$ and $\Pi_2$ are the POVM elements in the pure Bloch form

$$\Pi_1 = \frac{p_1}{2}(I_2 + \vec{n}_2' \cdot \vec{\sigma}) = p_1|n_2'\rangle\langle n_2'|, \quad |n_2'\rangle = \begin{pmatrix} -\sin\frac{\theta_2}{2} \\ e^{i\phi_2}\cos\frac{\theta_2}{2} \end{pmatrix},$$

$$\Pi_2 = \frac{p_2}{2}(I_2 + \vec{n}_1' \cdot \vec{\sigma}) = p_2|n_1'\rangle\langle n_1'|, \quad |n_1'\rangle = \begin{pmatrix} -\sin\frac{\theta_1}{2} \\ e^{i\phi_1}\cos\frac{\theta_1}{2} \end{pmatrix}. \tag{64}$$

For $\eta_1=\eta_2=1/2$, using Eq. (61) optimum $p$'s are given by

$$p_1 = p_2 = 1 - |a_{12}| = 1 - \cos\theta, \tag{65}$$

where $2\theta=n_1n_2$ and then minimum inconclusive result is $Q=\cos\theta$.

Regarding Fig. 2, it can be seen that $R=\cos\theta=Q$. That is, the minimum inconclusive result for unambiguous discrimination of two pure states is equivalent to the distance between sphere center and the line connecting $\rho_1$ to $\rho_2$.

Now, we give analytical solution for three linearly independent normalized state vectors $|\psi_1\rangle$, $|\psi_2\rangle$, and $|\psi_3\rangle$ in the three-dimensional complex vector space with arbitrary prior probabilities $\eta_1$, $\eta_2$, and $\eta_3$. In this case, $\hat{G}$ is given by

$$\hat{G} = \begin{pmatrix} \frac{1}{\sqrt{\xi_1}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{\xi_2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{\xi_3}} \end{pmatrix} \begin{pmatrix} 1 & a_{12} & a_{13} \\ a_{12}^* & 1 & a_{23} \\ a_{13}^* & a_{23}^* & 1 \end{pmatrix} \times \begin{pmatrix} \frac{1}{\sqrt{\xi_1}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{\xi_2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{\xi_3}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\xi_1} & \frac{a_{12}}{\sqrt{\xi_1\xi_2}} & \frac{a_{13}}{\sqrt{\xi_1\xi_3}} \\ \frac{a_{12}^*}{\sqrt{\xi_1\xi_2}} & \frac{1}{\xi_2} & \frac{a_{23}}{\sqrt{\xi_2\xi_3}} \\ \frac{a_{13}^*}{\sqrt{\xi_1\xi_3}} & \frac{a_{23}^*}{\sqrt{\xi_2\xi_3}} & \frac{1}{\xi_3} \end{pmatrix}. \tag{66}$$

The characteristic equation is given by

$$\lambda^3 - \left(\frac{1}{\xi_1} + \frac{1}{\xi_2} + \frac{1}{\xi_3}\right)\lambda^2 + \left(\frac{1}{\xi_1\xi_2} + \frac{1}{\xi_1\xi_3} + \frac{1}{\xi_2\xi_3} - \frac{|a_{12}|^2}{\xi_1\xi_2} - \frac{|a_{23}|^2}{\xi_2\xi_3} - \frac{|a_{13}|^2}{\xi_1\xi_3}\right)\lambda + \frac{1}{\xi_1\xi_2\xi_3}(|a_{12}|^2 + |a_{23}|^2 + |a_{13}|^2) - \frac{2\,\mathrm{Re}(a_{12}a_{23}^*a_{13}^*)}{\xi_1\xi_2\xi_3} - \frac{1}{\xi_1\xi_2\xi_3} = 0, \tag{67}$$

where $\mathrm{Re}(a_{12}a_{23}^*a_{13}^*)$, means real part of $a_{12}a_{23}^*a_{13}^*$. Since for the feasible region $p_i=\lambda\xi_i$, we have

$$p_1p_2p_3 - (p_2p_3 + p_1p_3 + p_1p_2) + (1 - |a_{23}|^2)p_1 + (1 - |a_{13}|^2)p_2 + (1 - |a_{12}|^2)p_3 + |a_{12}|^2 + |a_{23}|^2 + |a_{13}|^2 - 2\,\mathrm{Re}(a_{12}a_{23}^*a_{13}^*) - 1 = 0. \tag{68}$$

If the gradient of the plane $\eta_1p_1+\eta_2p_2+\eta_3p_3$ is equal to that of Eq. (68), we will have

$$p_2p_3 - (p_2 + p_3) + 1 - |a_{23}|^2 = \Lambda\eta_1, \quad p_1p_3 - (p_1 + p_3) + 1 - |a_{13}|^2 = \Lambda\eta_2, \quad p_1p_2 - (p_1 + p_2) + 1 - |a_{12}|^2 = \Lambda\eta_3. \tag{69}$$

By solving Eqs. (68) and (69), one can show that if $a_{ij}$ is real, only $\Lambda=0$ and $\Lambda=-\frac{a_{13}a_{23}}{\sqrt{\eta_1\eta_2}}+\frac{a_{12}(a_{13}\sqrt{\eta_1}-a_{23}\sqrt{\eta_2})}{\sqrt{\eta_1\eta_2\eta_3}}$ give the acceptable values for $p_i$. If the point of contact lies in the first octant and $0\le p_i\le 1$, it gives the optimal solution. If not, then an optimal contact point occurs on one of the coordinate planes or even at one of the vertices.

### A. Equal-probability measurement

A simple measurement that has been employed for unambiguous state discrimination is the measurement with $p_i=p$, for all $i=1,2,\ldots,N$. This measurement results in equal probability of correctly detecting each of the states and is called the equal-probability measurement (EPM). Using the feasible region, we are able to calculate the prior probabilities, so that EPM is optimal. Using Eqs. (68) and (69), the prior probabilities in the optimal EPM measurement for unambitious discrimination of three states are given by

$$\eta_1 = \frac{(a_{12}^2-1)^2}{3-2a_{12}^2-a_{13}^2-a_{12}^2a_{13}^2-2a_{23}+2a_{23}a_{12}^2+a_{12}^4},$$
$$\eta_2 = \frac{1-a_{12}^2-a_{13}^2+a_{12}^2a_{13}^2}{3-2a_{12}^2-a_{13}^2-a_{12}^2a_{13}^2-2a_{23}+2a_{23}a_{12}^2+a_{12}^4},$$
$$\eta_3 = 1-\eta_1-\eta_2. \tag{70}$$

In this case the value of $p$ is calculated from Eq. (68). If we consider more than three states, the functionality of the $\eta_i$ in terms of $\{a_{ij}\}$ are too complicated to be written down, so it is not included here. However, if we consider geometrically uniform states, the problem will be easy. Let $S=\{|\psi_i\rangle = U_i|\psi\rangle, U_i\in\mathcal{G}\}$ be a set of GU states generated by a finite group $\mathcal{G}$ of unitary matrices, where $|\psi\rangle$ is an arbitrary state. Now, let $\Phi$ be the matrix with columns $|\psi_i\rangle$. Then, the measurement which minimizes the probability of an inconclusive result could be reduced to an equal-probability measurement [27] and consists of the measurement operators

$$\Pi_i = p_i|\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|, \tag{71}$$

where $p_i$ is the inverse of the maximum eigenvalue of frame operator [27] and $|\tilde{\psi}_i\rangle = U_i|\tilde{\psi}\rangle$, $U_i\in\mathcal{G}$ with

$$|\tilde{\psi}\rangle = (\Phi\Phi^*)^{\ddagger}|\psi\rangle. \tag{72}$$

In this case, using the feasible region it is easy to show that for optimal EPM measurement all of prior probabilities are equal. In general, one can prove that for optimal EPM measurement all of the $p_i$'s are equal to the inverse of the maximum eigenvalue of the frame operator.

*Example 2.* We now consider an example of a set of GU states. Consider the group $\mathcal{G}$ of four unitary matrices $U_i$, where

$$U_1 = I_4, \quad U_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$
$$U_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad U_4 = U_2U_3. \tag{73}$$

Now, let the set of GU states is given by $S=\{|\psi_i\rangle = U_i|\psi\rangle, 1\le i\le 4\}$ with $|\psi\rangle = \frac{1}{3\sqrt{2}}[2,2,1,3]^T$. Then, we obtain

$$\Phi = \frac{1}{3\sqrt{2}}\begin{pmatrix} 2 & 2 & 2 & 2 \\ 2 & -2 & 2 & -2 \\ 1 & 1 & -1 & -1 \\ 3 & -3 & -3 & 3 \end{pmatrix}. \tag{74}$$

It should be noticed that the reciprocal states $|\tilde{\psi}_i\rangle = U_i|\tilde{\psi}\rangle$ for $i=1,\ldots,4$ with

$$|\tilde{\psi}\rangle = (\Phi\Phi^*)^{\ddagger}|\psi\rangle = \frac{1}{4\sqrt{2}}\begin{pmatrix} 3 \\ 3 \\ 6 \\ 2 \end{pmatrix} \tag{75}$$

are also GU states with generating group $\mathcal{G}$. Therefore, we can provide the elements of POVM as $\Pi_i=p_i|\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|$. Using feasible region it easy to show that the optimal contact point is given by

$$p_1=p_2=p_3=p_4=\frac{2}{9}, \tag{76}$$

then the equal probability measurement operators are given by

$$\Pi_i = \frac{2}{9}|\tilde{\psi}_i\rangle\langle\tilde{\psi}_i|, \quad i=1,\ldots,4, \tag{77}$$

where these results are in agreement with those of Ref. [27].

### B. Discrimination of quantum states using approximated linear programming

In general, finding an exact analytic solution for the OPUSD problem involving an arbitrary number of pure linearly independent quantum states is hard. Then, the approximated methods are useful for unambiguous discrimination of $N$ linearly independent quantum states. Since we have presented an analytic relation for the feasible region of $N$ linearly independent quantum states, and this region is convex, then one can easily obtain the optimal POVM by some well-known numerical methods such as constrained linear or nonlinear least-squares, interior points, and simplex and quadratic programming methods [32]. The simplex method is the easiest way to solve it. The simplex algorithm is a common one used to solve an optimization problem with a polytope feasible region, such as a linear programming problem. It is an improvement over the algorithm to test all feasible solutions of the convex feasible region and then to choose the optimal feasible solution. It does this by moving from one vertex to an adjacent vertex, such that the objective function is improved. This algorithm still guarantees that the optimal point will be discovered. In addition, only in the worst case scenario, all vertices will be tested. Here, considering the scope of this paper, a complete treatment of the simplex algorithm is unnecessary; for a more complete treatment, please refer to any LP text such as Refs. [32,33]. Using the feasible region, the extremal points are identified and through connecting them, a hyperplane is made. Next, the line crossing through the origin and center of hyperplane will

TABLE I. Calculation of the maximum success probability using simplex method.

| Number of extremal points | $p_1$ | $p_2$ | $p_3$ | $\mathrm{Max}(P_{\mathrm{suc}})$ |
|---|---|---|---|---|
| 4 | 0.04569739556 | 0.102819140 | 0.102819140 | 0.0837785585 |
| 7 | 0.01609505944 | 0.144855535 | 0.144855535 | 0.1019353765 |
| 10 | 0.00544829025 | 0.15936249 | 0.15936249 | 0.1080577568 |
| 13 | 0.00182476888 | 0.16422920 | 0.16422920 | 0.1100943896 |
| 16 | 0.00060919706 | 0.1658539 | 0.1658539 | 0.1107723324 |
| 19 | 0.00020316947 | 0.1663958 | 0.1663958 | 0.1109982565 |
| 22 | 0.00006773447 | 0.166576 | 0.166576 | 0.1110732448 |
| 25 | 0.00002257967 | 0.166638 | 0.166638 | 0.1110995266 |
| 28 | 0.00000752678 | 0.16666 | 0.16666 | 0.1111091756 |

contact with the feasible region, thus forming a point. This point constructs a polygon with the extremal points. With the use of simplex method, we obtain the optimal solution to the problem. To raise precision and accuracy, we need to increase the points. We make more new hyper-planes to arrive at this precision. These hyperplanes are formed through connecting extremal points and the center of the previous hyperplane. Moreover, the center of these hyperplanes and previous points construct new polygon with more points. Then, the precision and accuracy of OPUSD increase. For obtaining optimal solution, one can reiterate this procedure.

As an example, we consider three state vectors with equal probabilities 1/3, which were introduced in Ref. [7]

$$|\psi_1\rangle = \frac{1}{\sqrt{3}}[1,1,1]^T, \quad |\psi_2\rangle = \frac{1}{\sqrt{2}}[1,1,0]^T,$$

$$|\psi_3\rangle = \frac{1}{\sqrt{2}}[0,1,1]^T. \tag{78}$$

Using feasible region equation for these states and the above method for defining the convex region as a polygon one can solve maximization problem of the success probability by numerical methods such as simplex. We need to increase the points number on the feasible region to raise precision and accuracy. In Table I, we try to calculate the maximum success probability using simplex method. One can easily see that the success probability is improved and closed to the analytical result

$$p_1 = 0, \quad p_2 = p_3 = \frac{1}{6} \text{ and } \mathrm{Max}(P_{\mathrm{suc}}) = \frac{1}{9}, \tag{79}$$

by increasing the identified points on the feasible region.

## V. CONCLUSION

Here in this work by using the close connection between the Lewenstein-Sanpera decomposition and semidefinite programming, we could obtain an analytic relation for the feasible region of $N$ linearly independent quantum states in terms of the inner product of dual states. Also, for the real inner product of states, we have been able to obtain an exact analytic solution for OPUSD problem involving an arbitrary number of pure linearly independent quantum states by using KKT convex optimization method. Moreover, by reducing the theory of the semidefinite programming to a linear programming one with a feasible region of polygon type, we have been able to obtain optimal measurements to unambiguous discrimination of an arbitrary number of pure linearly independent quantum states numerically. With regard to the fact that we could demonstrate the feasible region for $N$ states throughout this study, it is possible to solve the problem, by using some numerical methods such as constrained linear or nonlinear least-squares, interior points, and simplex and quadratic programming methods. This paper represented the procedure through providing one example. Solving further problems in broad terms numerically and completely will hopefully be carried out in the future studies by the authors. Furthermore, unambiguous discrimination of mixed states is under investigation.

## APPENDIX A

Semidefinite programming. A SDP problem requires minimizing a linear function subject to a linear matrix inequality (LMI) constraint

$$\text{minimize } p = c^T x, \quad \text{such that } F(x) \geq 0, \tag{A1}$$

where $c^T$ is a given vector, $x=(x_1,\ldots,x_n)$, and $F(x)=F_0+\Sigma_i x_i F_i$, for some fixed Hermitian matrices $F_i$. The inequality sign in $F(x)\geq 0$ means that $F(x)$ is positive semidefinite.

This problem is called the primal problem. Vectors $x$ whose components are the variables of the problem and satisfy the constraint $F(x)\geq 0$ are called primal feasible points, and if they satisfy $F(x)>0$ they are called strictly feasible points. The minimal objective value $c^T x$ is by convention denoted by $P^\star$ and is called the primal optimal value. Due to the convexity of set of feasible points, SDP has a nice duality structure with the associated dual program being

$$\text{maximize } -\mathrm{Tr}[F_0 Z], \quad Z \geq 0, \quad \mathrm{Tr}[F_i Z] = c_i. \tag{A2}$$

Here the variable is the real symmetric (or Hermitian) matrix $Z$, and the data $c$, $F_i$ are the same as in the primal problem. Correspondingly, matrix $Z$ satisfying the constraints are called dual feasible (or strictly dual feasible if $Z>0$). The

maximal objective value of $-\mathrm{Tr}(F_0 Z)$, i.e., the dual optimal value, is denoted by $d_\star$. The objective value of a primal (dual) feasible point is an upper (lower) bound on $P^\star(d^\star)$. The main reason why one is interested in the dual problem is that one can prove that $d^\star \leq P^\star$, and under relatively mild assumptions, we can have $P^\star = d^\star$. If the equality holds, one can prove the following optimality condition on $x$.

A primal feasible $x$ and a dual feasible $Z$ are optimal which is denoted by $\hat{x}$ and $\hat{Z}$ if and only if

$$F(\hat{x})\hat{Z} = \hat{Z}F(\hat{x}) = 0. \tag{A3}$$

This latter condition is called the complementary slackness condition. In one way or another, numerical methods for solving SDP problems always exploit the inequality $d \leq d^\star \leq P^\star \leq P$, where $d$ and $P$ are the objective values for any dual feasible point and primal feasible point, respectively. The difference

$$p^\star - d^\star = c^T x + \mathrm{Tr}[F_0 Z] = \mathrm{Tr}[F_x Z] \geq 0 \tag{A4}$$

is called the duality gap. If the equality holds $d^\star = P^\star$, i.e., the optimal duality gap is zero, then we say that strong duality holds.

## APPENDIX B

### 1. Karush-Kuhn-Tucker (KKT) theorem

Assuming that functions $g_i$, $h_i$ are differentiable and that strong duality holds, there exists vectors $\zeta \in R^k$ and $y \in R^m$, such that the gradient of dual Lagrangian $L(x^*, \zeta^*, y^*) = f(x^*) + \Sigma_i \zeta_i^* h_i(x^*) + \Sigma_i y_i^* g_i(x^*)$ over $x$ vanishes at $x^*$:

$$\begin{aligned} &h_i(x^*) = 0 \ \text{(primal feasible)}, \\ &g_i(x^*) \leq 0 \ \text{(primal feasible)}, \\ &y_i^* \geq 0 \ \text{(dual feasible)}, \quad y_i^* g_i(x^*) = 0, \\ &\nabla f(x^*) + \sum_i \zeta_i^* \nabla h_i(x^*) + \sum_i y_i^* \nabla g_i(x^*) = 0. \end{aligned} \tag{B1}$$

Then $x^*$ and $(\zeta_i^*, y_i^*)$ are primal and dual optimal with zero duality gap. In summary, for any convex optimization problem with differentiable objective and constraint functions, the points which satisfy the KKT conditions are primal and dual optimal, and have zero duality gap. Necessary KKT conditions satisfied by any primal and dual optimal pair and for convex problems, KKT conditions are also sufficient. If a convex optimization problem with differentiable objective and constraint functions satisfies Slater's condition, then the KKT conditions provide necessary and sufficient conditions for optimality: Slater's condition implies that the optimal duality gap is zero and the dual optimum is attained, so $x$ is optimal if and only if there are $(\zeta_i^*, y_i^*)$ such that they, together with $x$, satisfy the KKT conditions.

### 2. Slater's condition

Suppose $x^*$ solves

$$\text{minimize } f(x) g_i(x) \geq b_i, \quad i = 1, \ldots, m \tag{B2}$$

and the feasible set is nonempty. Then there is a nonnegative vector $\zeta$ such that for all $x$

$$L(x, \zeta) = f(x) + \zeta^T [b - g(x)] \leq f(x^*) = L(x^*, \zeta). \tag{B3}$$

In addition, if $f(\cdots), g_i(\cdots), i = 1, \ldots, m$ are continuously differentiable, then

$$\frac{\partial f(x^*)}{\partial (x_j)} - \zeta \frac{\partial g(x^*)}{\partial (x)} = 0. \tag{B4}$$

In the spatial case the vector $x$ is a solution of the linear program

$$\text{minimize } c^T x, \quad \text{such that } Ax = bx \geq 0, \tag{B5}$$

if and only if there exist vectors $\zeta \in R^k$, and $y \in R^m$ for which the following conditions hold for $(x, \zeta, y) = (x^*, \zeta^*, y^*)$:

$$\begin{aligned} &A^T \zeta + y = c, \quad Ax = b, \quad x_i \geq 0; \quad y_i \geq 0; \\ &x_i y_i = 0, \ i = 1, \ldots, m. \end{aligned} \tag{B6}$$

A solution $(x^*, \zeta^*, y^*)$ is called strictly complementary, if $x^* + y^* > 0$, i.e., if there exists no index $i \in 1, \ldots, m$ such that $x_i^* = y_i^* = 0$.

## APPENDIX C: SDP FORMULATION OF UNAMBIGUOUS DISCRIMINATION

Eldar, Megretski, and Verghese in Ref. [19] have showed that the unambiguous discrimination problem can be reduced to SDP method and the KKT conditions can be defined as

$$\begin{aligned} &F(p) = \sum_{i=1}^N p_i F_i + F_0 \geq 0, \quad \mathrm{Tr}(\Pi_i X) = z_i + \eta_i, \\ &z_i \geq 0, 1 \leq i \leq N, \quad AF(p) = 0, \quad X(I_d - \sum_{i=1}^N p_i \Pi_i) = 0, \\ &z_i p_i = 0, \quad 1 \leq i \leq N, \quad \exists\, p_i : F_0 + \sum_{i=1}^N p_i F_i \geq 0, \\ &\exists\, X, z : X \geq 0, z \geq 0 \end{aligned} \tag{C1}$$

such that

$$|p\rangle := \begin{pmatrix} p_1 \\ \vdots \\ p_N \end{pmatrix}, \quad |c\rangle := \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_N \end{pmatrix}, \tag{C2}$$

$$F_0 := \begin{pmatrix} I_d & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

$$F_1 := \begin{pmatrix} -\Pi_1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix},$$

$$F_2 := \begin{pmatrix} -\Pi_2 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & \cdots & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}, \dots,$$

$$F_N := \begin{pmatrix} -\Pi_N & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}, \tag{C3}$$

$$A =: \begin{pmatrix} X_d & Y \\ Y^T & \begin{pmatrix} z_1 & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & \dots & z_N \end{pmatrix} \end{pmatrix}, \tag{C4}$$

where $X_d$ is the $d \times d$ matrix and $Y$ is the $N \times d$ matrix.


[1] A. S. Holevo, Probl. Peredachi Inf. **10**, 51 (1974); A. S. Holevo, Probl. Inf. Transm. **10**, 51 (1974).
[2] I. D. Ivanovic, Phys. Lett. A **123**, 257 (1987).
[3] D. Dieks, Phys. Lett. A **126**, 303 (1988).
[4] A. Peres and D. R. Terno, J. Phys. A **31**, 7105 (1998).
[5] G. Jaeger and A. Shimony, Phys. Lett. A **197**, 83 (1995).
[6] L. M. Duan and G. C. Guo, Phys. Rev. Lett. **80**, 4999 (1998).
[7] Y. Sun, M. Hillery, and J. A. Bergou, Phys. Rev. A **64**, 022311 (2001).
[8] A. Chefles, Phys. Lett. A **239**, 339 (1998).
[9] M. X. Goemans and D. P. Williamson, J. Assoc. Comput. Mach. **42**, 1115 (1995).
[10] Z. Q. Luo, Math. Program. **97**, 587 (2003).
[11] T. N. Davidson, Z.-Q. Luo, and K. M. Wong, IEEE Trans. Signal Process. **48**, 1433 (2000).
[12] Wing-Kin Ma, T. N. Davidson, K. M. Wong, Z.-Q. Luo, and P. C. Ching, IEEE Trans. Signal Process. **50**, 912 (2002).
[13] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, Studies in Applied Mathematics, SIAM (Philadelphia, PA 1994), Vol. 15.
[14] H. Barnum, M. Saks, and M. Szegedy, in *Proceedings of the 18th IEEE Annual Conference on Computational Complexity* (IEEE Computer Society, New York, 2003), pp. 179–193.
[15] M. A. Jafarizadeh, M. Mirzaee, M. Rezaee, Int. J. Quantum Inf. **2**, 541 (2004).
[16] M. A. Jafarizadeh, M. Mirzaee, M. Rezaee. Int. J. Quantum Inf. **3**, 511 (2005).
[17] M. A. Jafarizadeh, M. Mirzaee, and M. Rezaee, Quantum Inf. Process. **4**, 199 (2005).
[18] M. Mirzaee, M. Rezaee, M. A. Jafarizadeh, Int. J. Theor. Phys. **46**, 1471(2007).
[19] Y. C. Eldar, A. Megretski, and G. C. Verghese, IEEE Trans. Inf. Theory **49**, 1007 (2003).
[20] L. Ip, http://www.qcaustralia.org
[21] M. A. Jafarizadeh, M. Rezaee, and S. K. A. Seyed Yagoobi, Phys. Rev. A **72**, 062106 (2005).
[22] M. A. Jafarizadeh, M. Rezaee, and S. Ahadpour, Phys. Rev. A **74**, 042335 (2006).
[23] M. A. Jafarizadeh, G. Najarbashi, and H. Habibian, Phys. Rev. A **75**, 052326 (2007).
[24] M. A. Jafarizadeh and R. Sufiani, e-print arXiv:quant-ph/0703106.
[25] M. A. Jafarizadeh, G. Najarbashi, Y. Akbari and H. Habibian, e-print arXiv:quant-ph/0704.2414.
[26] K. Kraus, *States, Effects, and Operations*, *Lecture Notes in Physics* No. 190 (Springer, Berlin, 1983).
[27] Y. C. Eldar, IEEE Trans. Inf. Theory **49**, 446 (2003).
[28] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. (Johns Hopkins University Press, Baltimore, MD, 1996).
[29] L. R. Welch, IEEE Trans. Inf. Theory **20**, 397 (1974).
[30] M. Lewenstein and A. Sanpera, Phys. Rev. Lett. **80**, 2261 (1998).
[31] S. Karnas and M. Lewenstein, J. Phys. A **34**, 6919 (2001).
[32] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, 2004).
[33] E. K. P. Chong and S. H. Żak, *An Introduction to Optimization* (John Wiley, New York, 2001).