Mathematical nature of and a family of lower bounds for the success probability of unambiguous discrimination

Xiaoming Sun,* Shengyu Zhang, Yuan Feng, and Mingsheng Ying†
State Key Laboratory of Intelligent Technology and Systems, Department of Computer Science and Technology, Tsinghua University, Beijing, China 100084
(Received 8 November 2001; published 10 April 2002)

Unambiguous discrimination is a strategy to the discrimination problem that identifies the state with certainty, leaving a possibility of undecidability. This paper points out that the optimal success probability of unambiguous discrimination is mathematically the well-known semidefinite programming problem. A family of lower bounds of the optimal success probability is also given.

DOI: 10.1103/PhysRevA.65.044306 PACS number(s): 03.67.−a, 03.65.Ta

Unambiguous discrimination is one of the most important strategies to the problem of quantum state discrimination [1]. It describes the situation that a quantum system is secretly prepared in one of the finite known states, and we hope to identify the state which the system is actually in. Sometimes we prefer the requirement that once a result is reported, it must be true. This kind of discrimination is conventionally called unambiguous discrimination, and we sometimes call the success probability the efficiency of the discrimination. Unambiguous discrimination was first considered by Ivanovic [2], and then by Dieks [3] and Peres [4], all of which focused on the two-state case. The optimal efficiency they got is $1-|\langle p|q\rangle|$, later known as IDP-limit, where p and q are the two states to be distinguished. Jaeger and Shimony in [5] extended the result by augmenting unequal *a priori* probabilities to the two states. Peres and Terno [7] gave a discussion of the three-state case, but unfortunately, they did not give an explicit expression as the result. A more interesting and important extension is the general n-state case. In [8] Chefles showed that only linearly independent states can be unambiguously discriminated. Chefles and Barnett also considered a special case, known as equally probable symmetrical states, in [9]. For the general n-state discrimination, Duan and Guo gave a beautiful equivalent condition to the efficiency of the discrimination [10]. In a rather recent paper [11], Zhang, Feng, and Sun gave a neat upper bound for the optimal success probability of the n-state case. But what is the problem in a mathematical sense? This paper points out that the problem of success probability of unambiguous discrimination is the semidefinite programming (SDP) problem—a well-known mathematical problem. We first give a different proof from the one in [10], and then present a brief introduction of the SDP problem. Finally, we show a family of lower bounds for the optimal efficiency [6].

In what follows, we assume that a quantum system is secretly prepared in one of the states $|\varphi_1\rangle,|\varphi_2\rangle,\ldots,|\varphi_n\rangle$, which are linearly independent vectors in m-dimensional Hilbert space where $m\geqslant n$. We shall discriminate the states by a general measurement. A general measurement is a set of linear operators $\{M_1,M_2,\ldots,M_k\}$ such that $\Sigma_{i=1}^k M_i^\dagger M_i=I$. Without loss of generosity, we can further consider a measurement in the form of $\{M_0,M_1,\ldots,M_n\}$ such that $\Sigma_{i=0}^n M_i^\dagger M_i=I$ and $\langle\varphi_i|M_j^\dagger M_j|\varphi_i\rangle=0$, if i, $j=1,\ldots,n$, $i\neq j$. Intuitively, if outcome $i(i\neq 0)$ occurs, one may claim with certainty that the system is originally in the state $|\varphi_i\rangle$; if outcome 0 occurs, the identification fails to give a report. We begin with Lemma 1, which intuitively reduces the original problem to the discrimination within the subspace spanned by $|\varphi_1\rangle,|\varphi_2\rangle,\ldots,|\varphi_n\rangle$. We denote the m-dimensional Hilbert space and its subspace *span* $\{|\varphi_1\rangle,|\varphi_2\rangle,\ldots,|\varphi_n\rangle\}$ by V^m and V^n, respectively.

Lemma 1. For any measurement $\{M_i:V^m\rightarrow V^m|i=0,1,\ldots,n\}$ such that $I_m-\Sigma_{i=1}^n M_i^\dagger M_i$ is positive and $\langle\varphi_i|M_j^\dagger M_j|\varphi_i\rangle=0$ for all $i\neq j$, there exists a measurement $\{\tilde{M}_i:V^n\rightarrow V^n|i=0,1,\ldots,n\}$ such that $I_n-\Sigma_{i=1}^n\tilde{M}_i^\dagger\tilde{M}_i$ is positive and $\langle\varphi_i|M_j^\dagger M_j|\varphi_i\rangle=\langle\varphi_i|\tilde{M}_j^\dagger\tilde{M}_j|\varphi_i\rangle$ for all i and j.

Proof. Let $\tilde{V}^n$ denote *span* $\{M_1|\varphi_1\rangle,M_2|\varphi_2\rangle,\ldots,M_n|\varphi_n\rangle\}$, then $\tilde{V}^n$ has the dimensionality less than or equal to n. So any linear operator $f:\tilde{V}^n\rightarrow V^n$ mapping a orthonormal basis of $\tilde{V}^n$ to a set of orthonormal vectors in V^n preserves the norm of all vectors in $\tilde{V}^n$. Now let $\tilde{M}_i=f\circ M_i|_{\tilde{V}^n}$. We have

$$\langle\varphi_i|\tilde{M}_j^\dagger\tilde{M}_j|\varphi_i\rangle=|\tilde{M}_j|\varphi_i\rangle|^2=|f\circ M_j|\varphi_i\rangle|^2$$
$$=|M_j|\varphi_1\rangle|^2=\langle\varphi_i|M_j^\dagger M_j|\varphi_i\rangle$$

by noting that when $i\neq j$, $|M_j|\varphi_i\rangle|=0\in\widetilde{V^n}$, thus $f\circ M_j|\varphi_i\rangle=0$. And what is more, for all $|\varphi\rangle\in V^n$, we have

$$\langle\varphi|\varphi\rangle-\sum_{i=1}^n\langle\varphi|\tilde{M}_i^\dagger\tilde{M}_i|\varphi\rangle=\langle\varphi|\varphi\rangle-\sum_{i=1}^n|\tilde{M}_i\varphi\rangle|^2$$
$$=\langle\varphi|\varphi\rangle-\sum_{i=1}^n|M_i\varphi\rangle|^2$$
$$=\langle\varphi|\varphi\rangle-\sum_{i=1}^n\langle\varphi|M_i^\dagger M_i|\varphi\rangle\geqslant 0,$$

which indicates that $I_n-\Sigma_{i=1}^n\tilde{M}_i^\dagger\tilde{M}_i$ is positive and concludes the proof.

For a diagonal matrix $\Gamma=\mathrm{diag}\{\gamma_1,\gamma_2,\ldots,\gamma_n\}$, we say a measurement $\{M_i:i=0,1,\ldots,n\}$ can unambiguously dis-

*Email address: sun_x_m@hotmail.com
†Email address: yingmsh@tsinghua.edu.cn

criminate states $|\varphi_1\rangle,|\varphi_2\rangle,\ldots,|\varphi_n\rangle$ with efficiency Γ if $\langle\varphi_i|M_j^\dagger M_j|\varphi_i\rangle=0$ $(i\neq j)$, and $\langle\varphi_i|M_i^\dagger M_i|\varphi_i\rangle=\gamma_i$. The following Lemma 2 reduces the problem to an algebra one with a simple form.

Lemma 2. For linearly independent states $|\varphi_1\rangle,|\varphi_2\rangle,\ldots,|\varphi_n\rangle$, there is a measurement to unambiguously discriminate them with efficiency Γ if and only if $X-\Gamma$ and Γ are both positive, where $X=(\langle\varphi_i|\varphi_j\rangle)_{n\times n}$.

Proof. Based on Lemma 1, we can consider the measurement $\{M_i: i=0,1,\ldots,n\}$ in V^n. Because $|\varphi_1\rangle,|\varphi_2\rangle,\ldots,|\varphi_n\rangle$ are linearly independent, it is easy to see that $M_i^\dagger M_i$ has the form of $\alpha_i|\omega_i\rangle\langle\omega_i|$ where $|\omega_i\rangle$ is in the one-dimensional subspace vertical to each $|\varphi_j\rangle(j\neq i)$. We set the length of $|\omega_i\rangle$ such that $\langle\omega_i|\varphi_i\rangle=1$. Then the ith efficiency $\gamma_i=\langle\varphi_i|\alpha_i|\omega_i\rangle\langle\omega_i|\varphi_i\rangle=\alpha_i$. Now, $\exists M_i$, such that $\{M_0,M_1,\ldots,M_n\}$ is a measurement if and only if $I_n-\Sigma_{i=1}^n M_i^\dagger M_i$ is positive, i.e., for any $|\varphi\rangle=\Sigma_{i=1}^n c_i|\varphi_i\rangle$,

$$\begin{aligned}0\leqslant\langle\varphi|I_n-\sum_{i=1}^n M_i^\dagger M_i|\varphi\rangle&=\langle\varphi|\varphi\rangle\\&-\sum_{i=1}^n\sum_{j=1}^n c_i^*\gamma_j c_i\langle\varphi_i|\omega_j\rangle\langle\omega_j|\varphi_i\rangle\\&=(c_1^*,\ldots,c_n^*)X(c_j,\ldots,c_n)^T-\sum_{i=1}^n\gamma_i c_i^* c_i\\&=(c_1^*,\ldots,c_n^*)(X-\Gamma)(c_1,\ldots,c_n)^T,\end{aligned}$$

which just says $X-\Gamma$ is positive. ■

It should be noted that this result has been got by Duan and Guo in [10]. They derived the result from the fact [12] that a general measurement on system *A* can be represented by a unitary operation *U* on the composite system *ABP*, succeeded by von Neumann's measurement on probe *P*. Here we do not introduce the auxiliary system and consider the problem by a general measurement. The two ways are equivalent, and our lemma can serve as another proof of the problem.

In the rest of the paper, we give a family of lower bounds to the optimal mean efficiency, i.e., $\Sigma_{i=1}^n p_i\gamma_i$. We denote the ith largest eigenvalue of A by $\lambda_i(A)$, $\lambda_1(A)\geqslant\lambda_2(A)\geqslant\cdots\geqslant\lambda_n(A)$.

Theorem 1. For any $q=(q_1,\ldots,q_n)$, $q_i>0$ $(i=1,\ldots,n)$,

$$\rho=\lambda_n(\mathrm{diag}(\sqrt{q_1},\ldots,\sqrt{q_n})X\mathrm{diag}(\sqrt{q_1},\ldots,\sqrt{q_n}))\Sigma_{i=1}^n p_i/q_i,$$

is a lower bound of the optimal mean efficiency $\Sigma_{i=1}^n p_i\gamma_i$. In particular, $\lambda_n(X)$ is a lower bound.

Proof. We write λ_i for $\lambda_i(\mathrm{diag}(\sqrt{q_1},\ldots,\sqrt{q_n})X\mathrm{diag}(\sqrt{q_1},\ldots,\sqrt{q_n}))$. If we let $\gamma_i=\lambda_n/q_i$, then $\Sigma_{i=1}^n p_i\gamma_i=\lambda_n\Sigma_{i=1}^n p_i/q_i=\rho$. So in the rest we only need to prove $X-\Gamma$ is positive. In fact,

$$\begin{aligned}X-\Gamma&=X-\mathrm{diag}\{\lambda_n/q_1,\ldots,\lambda_n/q_n\}\\&=\mathrm{diag}(\sqrt{q_1},\ldots,\sqrt{q_n})^{-1}\{\mathrm{diag}(\sqrt{q_1},\ldots,\sqrt{q_n})\\&\quad\times X\,\mathrm{diag}(\sqrt{q_1},\ldots,\sqrt{q_n})-\lambda_n I_n\}\mathrm{diag}(\sqrt{q_1},\ldots,\sqrt{q_n})^{-1},\end{aligned}$$

so

$$X-\Gamma\geqslant 0\Leftrightarrow\mathrm{diag}(\sqrt{q_1},\ldots,\sqrt{q_n})X\,\mathrm{diag}(\sqrt{q_1},\ldots,\sqrt{q_n})-\lambda_n I_n\geqslant 0.$$

Because $\lambda_1,\ldots,\lambda_n$ are the eigenvalues of $\mathrm{diag}(\sqrt{q_1},\ldots,\sqrt{q_n})$ X $\mathrm{diag}(\sqrt{q},\ldots,\sqrt{q_n})$, there must exist a unitary matrix U such that $\mathrm{diag}(\sqrt{q_1},\ldots,\sqrt{q_n})$ X $\mathrm{diag}(\sqrt{q_1},\ldots,\sqrt{q_n})=U^\dagger\mathrm{diag}\{\lambda_1,\ldots,\lambda_n\}U$. So

$$\begin{aligned}&\mathrm{diag}(\sqrt{q_1},\ldots,\sqrt{q_n})X\,\mathrm{diag}(\sqrt{q_1},\ldots,\sqrt{q_n})-\lambda_n I_n\\&\quad=U^\dagger\,\mathrm{diag}\{\lambda_1-\lambda_n,\ldots,\lambda_{n-1}-\lambda_n,0\}\,U.\end{aligned}$$

Following the definition of λ_i we have $\lambda_1\geqslant\cdots\geqslant\lambda_n$, and it follows that $\lambda_i-\lambda_n\geqslant 0$. This indicates $X-\Gamma\geqslant 0$ is positive.

In particular, if we let $q=(1,\ldots,1)$, then it follows that $\lambda_n(X)$ is a special lower bound. ■

Note that when $p_1=p_2=1/2$, $n=2$, $\lambda_n=\lambda_2(\left[\begin{smallmatrix}1 & \langle\varphi_1|\varphi_2\rangle\\ \langle\varphi_2|\varphi_1\rangle & 1\end{smallmatrix}\right])=1-|\langle\varphi_1|\varphi_2\rangle|$, which coincides with the IDP limit.

We conclude the paper with some remarks about the mathematical problem reduced in Lemma 2. Mathematically this is a well-known minimal trace problem. The general form is as follows:

$$\text{minimize}\quad c^T x,$$

$$\text{subject to}\quad F(x)\geqslant 0,$$

where $F(x)=F_0+\Sigma_{i=1}^m x_i F_i$ and where the vector $c\in R^m$ and $m+1$ symmetric matrices $F_0,\ldots,F_m\in R^{n\times n}$. The inequality sign in $F(x)\geqslant 0$ means that $F(x)$ is positive.

A special case is the following one:

$$\text{maximize}\quad\sum_{i=1}^n d_i,$$

$$\text{subject to}\quad\hat{\Sigma}-\mathrm{diag}(d)\geqslant 0,\quad d\geqslant 0.$$

It is easy to see that the problem of maximizing the optimal mean efficiency just belongs to the second one. Both the two have been studied for about two decades [13], and have applications in many areas. There are even web sites and softwares for numerical computation, all of which can directly serve as numerical solutions for our unambiguous discrimination problem.

This work is supported by the National Foundation for Distinguished Young Scholars, Grant No. 69725004, the National Key Project for Basic Research, Grant No. 69823001, and the National Foundation of Natural Sciences, Grant No. 69823001.

[1] A. Chefles, Contemp. Phys. **41**, 401 (2000).
[2] I. D. Ivanovic, Phys. Lett. A **123**, 257 (1987).
[3] D. Dieks, Phys. Lett. A **126**, 303 (1988).
[4] A. Peres, Phys. Lett. A **128**, 19 (1988).
[5] G. Jaeger and A. Shimony, Phys. Lett. A **197**, 83 (1995).
[6] M. Ban, Phys. Lett. A **213**, 235 (1996).
[7] A. Peres and D. Terno, J. Phys. A **31**, 7105 (1998).
[8] A. Chefles, Phys. Lett. A **239**, 239 (1998).
[9] A. Chefles and S. M. Barnett, Phys. Lett. A **250**, 223 (1998).
[10] L. M. Duan and G. G. Guo, Phys. Lett. A **80**, 4999 (1998).
[11] S. Y. Zhang, Y. Feng, X. M. Sun, and M. S. Ying, Phys. Rev. A **64**, 062103 (2001).
[12] K. Kraus, *States, Effects, and Operations* (Springer-Verlag, Berlin, 1983).
[13] L. Vandenberghe and S. Boyd, SIAM Rev. **38**(1), 49 (1996).