PHYSICAL REVIEW A, VOLUME 65, 060301

Finding optimal strategies for minimum-error quantum-state discrimination

M. Ježek, J. Řeháček,* and J. Fiurášek
Department of Optics, Palacký University, 17. listopadu 50, 77200 Olomouc, Czech Republic

(Received 1 February 2002; revised manuscript received 30 April 2002; published 19 June 2002)

We propose a numerical algorithm for finding optimal measurements for quantum-state discrimination. The theory of the semidefinite programming provides a simple check of the optimality of the numerically obtained results. With the help of our algorithm we calculate the minimum attainable error rate of a device discriminating between three particularly chosen pure qubit states.

DOI: 10.1103/PhysRevA.65.060301 PACS number(s): 03.67.Hk

Nonorthogonality of quantum states is one of the basic features of quantum mechanics. Its deep consequences are reflected in all quantum protocols. For instance, it is well known that perfect discrimination between two nonorthogonal states cannot be made. This has important implications for the information processing at the microscopic level since it sets a limit on the amount of information that can be encoded into a quantum system. Although perfect decisions between nonorthogonal quantum states are impossible, it is of importance to study measurement schemes performing this task in the optimum, though imperfect, way.

Two conceptually different models of decision tasks have been studied. The first one is based on the minimization of the Bayesian cost function, which is nothing but a generalized error rate [1]. In the special case of linearly independent pure states, the second model—unambiguous discrimination of quantum states—makes an interesting alternative. The latter scheme combines the error-less discrimination with a certain fraction of inconclusive results [2–4].

Ambiguous as well as unambiguous discrimination schemes have been intensively studied over the past few years. As a consequence, the optimal measurements distinguishing between pair, trine, tetrad states, and linearly independent symmetric states are now well understood [5–17]. Many of the theoretically discovered optimal devices have already been realized experimentally, mainly with polarized light [18–21]. As an example of the practical importance of the optimal decision schemes let us mention their use for the eavesdropping on quantum cryptosystems [22,23].

The purpose of this paper is to develop a universal method for optimizing ambiguous discrimination between generic quantum states. Assume that Alice sets up M different sources of quantum systems living in p-dimensional Hilbert space. The complete quantum-mechanical description of each source is provided by its density matrix. Alice chooses one of the sources at random using a chance device and sends the generated quantum system to Bob. Bob is also given M numbers $\{\xi_i\}$ specifying probabilities that the ith source is selected by the chance device. Bob is then required to tell which of the M sources $\{\rho_i\}$ generated the quantum system he had obtained from Alice. In doing this he should make as few mistakes as possible.

*Electronic address: rehacek@phoenix.inf.upol.cz

It is well known [1] that each of Bob's strategies can be described in terms of an M-component probability operator measure (POM) $\{\Pi_j\}$, $0\leqslant\Pi_i\leqslant 1$, $\Sigma_j\Pi_j=1$. Each POM element corresponds to one output channel of Bob's discriminating apparatus. The probability that Bob points his finger at the kth source while the true source is j is given by the trace rule: $P(k|j)=\mathrm{Tr}\,\rho_j\Pi_k$. Taking the prior information into account, the average probability of Bob's success in repeated experiments is

$$P_s=\sum_{j=1}^{M}\xi_j\mathrm{Tr}\,\rho_j\Pi_j. \tag{1}$$

Since the objective is to keep Bob's error rate as low as possible we should maximize this number over the set of all M-component POMs. In compact form the problem reads

$$\text{Maximize}\quad P_s\quad\text{subject to constraints:}$$

$$\Pi_j\geqslant 0,\quad j=1,\ldots,M, \tag{2}$$

$$\sum_j\Pi_j=1.$$

Unfortunately, attacking this problem by analytical means has a chance to succeed only in the simplest cases ($M=2$) [1], or cases with symmetric or linearly independent states [5,10,15,17,25]. In most situations one must resort to numerical methods. In the following we will use the calculus of variations to derive an iterative algorithm [36] that provides a convenient way of dealing with the problem (2). This approach has already found its use in the optimization of teleportation protocols [26], the optimization of completely positive maps that approximate some unphysical operations [27], and the maximum-likelihood estimation of quantum states [28] and measurements [29]. We are going to seek the global maximum of the success functional P_s subject to the constraints given in Eq. (2). To take care of the first constraint we will decompose the POM elements as follows $\Pi_j=A_j^{\dagger}A_j$, $j=1,\ldots,M$. The other constraint (completeness) can be incorporated into our model using the method of uncertain Lagrange multipliers. Putting all things together, the functional to be maximized becomes

$$\mathcal{L}=\sum_j \xi_j \operatorname{Tr}\{\rho_j A_j^\dagger A_j\} - \operatorname{Tr}\left\{\lambda \sum_j A_j^\dagger A_j\right\}, \tag{3}$$

where λ is a Hermitian Lagrange operator. This expression is now to be varied with respect to M independent variables A_j to yield a necessary condition for the extremal point in the form of a set of M extremal equations for the unknown POM elements:

$$\xi_j \rho_j \Pi_j = \lambda \Pi_j, \quad j=1,\ldots,M, \tag{4}$$

originally derived by Holevo in [30]. For our purposes it is advantageous to bring these equations to an explicitly positive semidefinite form,

$$\Pi_j = \xi_j^2 \lambda^{-1} \rho_j \Pi_j \rho_j \lambda^{-1}, \quad j=1,\ldots,M. \tag{5}$$

The Lagrange operator λ is obtained by summing Eq. (5) over j,

$$\lambda = \left(\sum_j \xi_j^2 \rho_j \Pi_j \rho_j\right)^{1/2}. \tag{6}$$

The iterative algorithm comprised of the $M+1$ equations (5) and (6) is the main formal result of this paper. One usually starts from some "unbiased" trial POM $\{\Pi_j^0\}$. After plugging it in Eq. (6) the first guess of the Lagrange operator λ is obtained. This operator is, in turn, used in Eq. (5) to get the first correction to the initial-guess strategy $\{\Pi_j^0\}$ [31]. The procedure gets repeated, until, eventually, a stationary point is attained. Notice that both the positivity and completeness of the initial POM are preserved in the course of iterating.

In the many tests we did a monotonic convergence to the true global maximum of the success rate (1) always had been observed, though we have no analytic proof of this behavior in general. However, by slight modification of the iteration (5) one can make sure that the Bayes cost (error rate) is reduced in each step. Let us rewrite the right-hand side of Eq. (5) as a perturbation of the old POM element:

$$\Pi_j^{i+1} = (1+D_j^\dagger)\Pi_j^i(1+D_j) = \Pi_j^i + \delta\Pi_j, \tag{7}$$

where $D_j = (\xi_j \rho_j - \lambda)\lambda^{-1}$. The success rate is changed by the amount of $\delta P_s = \Sigma_j \operatorname{Tr} \rho_j \delta\Pi_j$, which is required to be positive. If we found that δP_s was negative we would then change the sign of the perturbation [37]: $\delta\Pi_j \to -\delta\Pi_j$. This ensures that the Bayes cost is reduced in each step. Further notice that the operators D_j are closely related to the gradient of the functional (3). Our algorithm is an example of a more general class of gradient-type algorithms of the form $a_k^{i+1} = [\partial P(a^i)/\partial a_k^i] a_k^i [\partial P(a^i)/\partial a_k^i]$. Such algorithms are known to behave well and their applications cover many inverse and optimization problems such as image processing [32], positron emission tomography [33], or optimization over completely positive maps [26,27]. Finally notice that the norm of the perturbation $\delta\Pi_j$ decreases as the POM $\{\Pi_j\}$ approaches the POM of the optimal discrimination apparatus. In this way the length of the iteration step self-adapts in the course of iterating—it gets progressively smaller in the neighborhood of the solution.

Since Eqs. (5) and (6) represent only a necessary condition for the extreme, one should always check the optimality of the stationary point by verifying the following set of conditions [1,30]:

$$\lambda - \xi_j \rho_j \geqslant 0, \quad j=1,\ldots,M. \tag{8}$$

It is worth mentioning that this condition can also be derived from the theory of the semidefinite programming (SDP) [24]. SDP tools also provide an alternative means of solving the problem (2) numerically [34,35]. Recently, it has been pointed out [35] that many problems of the quantum-information processing can be formulated as SDP problems. To see the link between the quantum discrimination and SDP theory let us note that the dual problem of SDP is defined as follows: Maximize $-\operatorname{Tr} F_0 Z$ subject to constraints:

$$Z \geqslant 0, \tag{9}$$

$$\operatorname{Tr} F_i Z = c_i, \quad i=1,\ldots,m,$$

where data are $m+1$ Hermitian matrices F_i and a complex vector $c \in \mathbb{C}^m$, and Z is a Hermitian variable. Our problem (2) reduces to this dual SDP problem upon the following substitutions:

$$F_0 = -\bigoplus_{j=1}^m \xi_j \rho_j, \quad Z = \bigoplus_{j=1}^m \Pi_j, \tag{10}$$

$$F_i = \bigoplus_{j=1}^m \Gamma_i, \quad c_i = \operatorname{Tr} \Gamma_i, \quad i=1,\ldots,p^2.$$

Here operators $\{\Gamma_i, i=1,\ldots,p^2\}$ comprise an orthonormal operator basis in the p^2-dimensional space of Hermitian operators acting on the Hilbert space of our problem: $\operatorname{Tr} \Gamma_j \Gamma_k = \delta_{jk}$, $j,k=1,\ldots,p^2$. For simplicity, let us take Γ_1 proportional to the unity operator, then all c_i apart from c_1 vanish. In SDP the necessary condition (4) for the maximum of the functional (3) is called the complementary slackness condition. When inequalities in Eq. (8) hold it can be shown to be also sufficient.

The advantage of the SDP formulation of the quantum-state discrimination problem is that there are strong numerical tools designed for solving SDP problems, for their review see [34]. They make use of the duality of SDP problems. The optimal value is bracketed between the trial maximum of the dual problem and trial minimum of the primal problem. One then hunts the optimal value down by making this interval gradually smaller. SDP tools are much more complicated than the proposed algorithm but they are guaranteed to converge to the real solution.

We have carried out extensive tests of our numerical algorithm on discriminations involving up to four pure or mixed states in Hilbert spaces of dimensions two, three, and four. In all these cases a monotonic convergence to the global optimum has been observed.

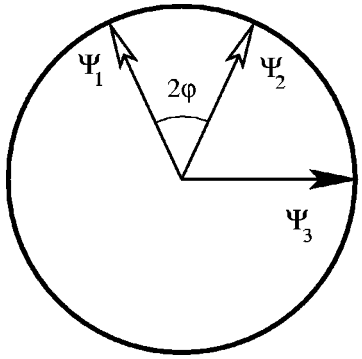

FIG. 1. A cut through the Bloch sphere showing the states to be discriminated.

Let us illustrate the utility of our algorithm on a simple, albeit nontrivial example of discriminating between three coplanar pure qubit states. The geometry of this problem is shown in Fig. 1. Ψ_1 and Ψ_2 are equal-prior states, $\xi_1=\xi_2=\xi/2$, symmetrically placed around the z axis; the third state lies in the direction of x or y. A similar configuration (with Ψ_3 lying along z) has recently been investigated by Andersson *et al.* [17]. Exploiting the mirror symmetry of their problem the authors derived analytic expressions for POMs minimizing the average error rate. For a given angle φ the optimum POM turned out to have two or three nonzero elements depending on the amount of the prior information ξ.

Our problem is a bit more complicated one due to the lack of the mirror symmetry. Let us see whether the transition from the mirror-symmetric configuration to a nonsymmetric one has some influence on the qualitative behavior of the optimal POMs. Minimal error rates calculated using the proposed iterative procedure [Eqs. (5) and (6)] for the fixed angle of $\varphi=\pi/16$ are summarized in Fig. 2. The conclusions that can be drawn from the numerical results partly coincide with that of Ref. [17]. For large ξ (region III) the optimum strategy consists of the optimal discrimination between states Ψ_1 and Ψ_2. When ξ becomes smaller than a certain φ-dependent threshold (region II), state Ψ_3 can no longer be

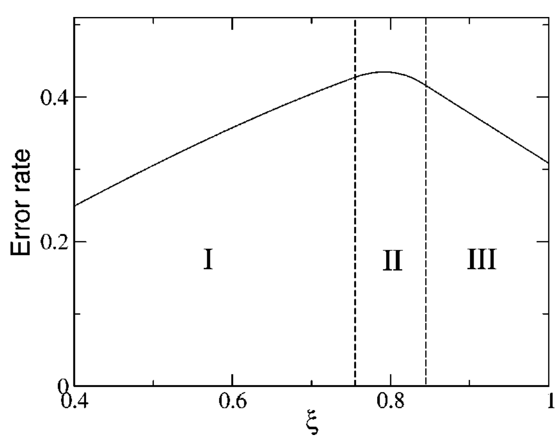

FIG. 2. Average error rate $(1-P_s)$ in dependence on Bob's prior information ξ; $\varphi=\pi/16$. Regions I, II, and III are regions where the optimum discriminating device has two, three, and two output channels, respectively.

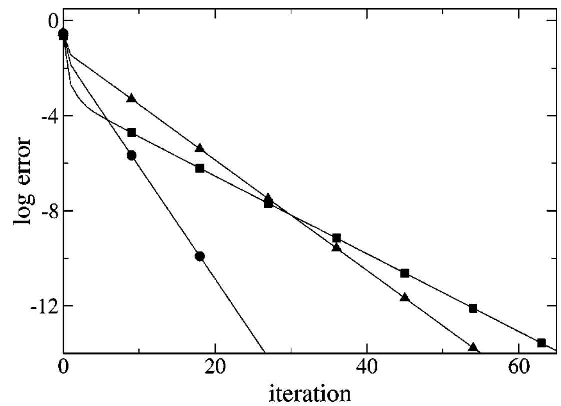

FIG. 3. Accuracy of the calculated error rate of the optimal POM vs the number of iterations. Convergence of the proposed algorithm is shown for three different priors: $\xi=0.6$ (squares), $\xi=0.8$ (triangles), and $\xi=0.9$ (circles). The ordinate is labeled by the precision in decimal digits.

ignored and the optimum POM has three nonzero elements. Simple calculation yields

$$\xi_{\mathrm{II,III}}=\frac{1}{1+\sin\varphi\cos\varphi} \tag{11}$$

for the threshold value of the prior. However, when ξ becomes still smaller (region I), the optimum POM will eventually become a two-element POM again—the optimal strategy now being the optimal discrimination between states Ψ_1 and Ψ_3. This last regime is absent in the mirror-symmetric case. The transition between regions I and II is governed by a much more complicated expression than Eq. (11), and will not be given here.

The convergence properties of the algorithm are shown in Fig. 3 for three typical prior probabilities representing regions I, II, and III of Fig. 2. After a short transient period an exponentially fast convergence sets in. Sixteen-digit precision in the resulting error rate is usually obtained after less than 100 iterations. Let us close the example noting that already a few iterations are enough to determine the optimum discriminating device to the precision the elements of the realistic experimental setup can be controlled with in the laboratory.

In this paper we derived a simple iterative algorithm for finding optimal devices for quantum-state discrimination. Utility of our procedure was illustrated on a nontrivial example of discriminating between three pure qubit states. From the mathematical point of view, the problem of quantum-state discrimination is a problem of the semidefinite programming. Such correspondence is a good news since there exist robust numerical tools designed to deal with SDP problems. These could substitute our iterative algorithm in the very few exceptional cases where it might converge too slowly.

This work was supported by Grant No. LN00A015 and project CEZ:J14/98 ''Wave and particle optics'' of the Czech Ministry of Education.

[1] C.W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
[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] H.P. Yuen, R.S. Kennedy, and M. Lax, IEEE Trans. Inf. Theory **IT-21**, 125 (1975).
[6] A. Chefles, Phys. Lett. A **239**, 339 (1998).
[7] A. Chefles and S.M. Barnett, J. Mod. Opt. **45**, 1295 (1998).
[8] A. Chefles and S.M. Barnett, Phys. Lett. A **250**, 223 (1998).
[9] L.S. Phillips, S.M. Barnett, and D.T. Pegg, Phys. Rev. A **58**, 3259 (1998).
[10] M. Sasaki, K. Kato, M. Izutsu, and O. Hirota, Phys. Rev. A **58**, 146 (1998).
[11] A. Chefles, Contemp. Phys. **41**, 401 (2000); e-print quant-ph/0010114.
[12] J. Walgate, A.J. Short, L. Hardy, and V. Vedral, Phys. Rev. Lett. **85**, 4972 (2000).
[13] S. Virmani, M.F. Sacchi, M.B. Plenio, and D. Markham, Phys. Lett. A **288**, 62 (2001).
[14] S. Zhang, Y. Feng, X. Sun, and M. Ying, Phys. Rev. A **64**, 062103 (2001).
[15] S.M. Barnett, Phys. Rev. A **64**, 030303(R) (2001).
[16] A. Chefles, Phys. Rev. A **64**, 062305 (2001).
[17] E. Andersson, S.M. Barnett, C.R. Gilson, and K. Hunter, Phys. Rev. A **65**, 052308 (2002).
[18] B. Huttner, A. Muller, J.D. Gautier, H. Zbinden, and N. Gisin, Phys. Rev. A **54**, 3783 (1996).
[19] S.M. Barnett and E. Riis, J. Mod. Opt. **44**, 1061 (1997).
[20] R.B.M. Clarke, A. Chefles, S.M. Barnett, and E. Riis, Phys. Rev. A **63**, 040305(R) (2001).
[21] R.B.M. Clarke, V.M. Kendon, A. Chefles, S.M. Barnett, E. Riis, and M. Sasaki, Phys. Rev. A **64**, 012303 (2001).
[22] A.K. Ekert, B. Huttner, G.M. Palma, and A. Peres, Phys. Rev. A **50**, 1047 (1994).
[23] M. Dušek, M. Jahma, and N. Lütkenhaus, Phys. Rev. A **62**, 022306 (2000).
[24] For a slightly different approach based on vector-space optimization theory, see Ref. [5].
[25] M. Ban, K. Kurokawa, R. Momose, and O. Hirota, Int. J. Theor. Phys. **36**, 1269 (1997).
[26] J. Řeháček, Z. Hradil, J. Fiurášek, and Č. Brukner, Phys. Rev. A **64**, 060301(R) (2001).
[27] J. Fiurášek, Phys. Rev. A **64**, 062310 (2001).
[28] J. Řeháček, Z. Hradil, and M. Ježek, Phys. Rev. A **63**, 040303(R) (2001).
[29] J. Fiurášek, Phys. Rev. A **64**, 024102 (2001).
[30] A.S. Holevo, J. Multivariate Anal. **3**, 337 (1973).
[31] It is interesting to notice that when the initial POM is chosen to be the maximally ignorant one, $\Pi_j^0 = 1/M$, $j = 1 \ldots M$, the first correction is quite similar to the "pretty good" measurement introduced by P. Hausladen and W.K. Wootters, J. Mod. Opt. **41**, 2385 (1994). For equally probable states, $\xi_j = 1/M, \forall j$, they are equivalent. Interestingly enough, the pretty good measurement is known to be optimal in certain cases [25].
[32] C.L. Byrne, IEEE Trans. Image Process. **2**, 96 (1993); Y. Vardi and D. Lee, J. R. Statist. Soc. Ser. B. Methodol. **55**, 569 (1993).
[33] B.A. Mair, M. Rao, and J.M.M. Anderson, Inverse Probl. **12**, 965 (1996).
[34] L. Vandenberghe and S. Boyd, SIAM (Soc. Ind. Appl. Math.) Rev. **38**, 49 (1996).
[35] E.M. Rains, IEEE Trans. Inf. Theory **47**, 2921 (2001); K. Audenaert and B. De Moor, Phys. Rev. A **65**, 030302(R) (2002); A.C. Doherty, P.A. Parrilo, and F.M. Spedalieri, Phys. Rev. Lett. **88**, 187904 (2002).
[36] An iterative algorithm for finding the maximum of the average success rate over projection valued POM has been presented in Helstrom, IEEE Trans. Inf. Theory **IT-28**, 359 (1982). In a general case, however, it finds only a lower bound to the optimal P_s.
[37] A little extra care is needed here since the resulting Π_j must be positive. Perturbation with an altered sign might require a smaller step.