Universal Programmable Quantum State Discriminator that is Optimal for Unambiguously Distinguishing between Unknown States

János A. Bergou and Mark Hillery

Department of Physics, Hunter College, City University of New York, 695 Park Avenue, New York, New York 10021, USA
(Received 21 September 2004; published 25 April 2005)

We construct a device that can unambiguously discriminate between two unknown quantum states. The unknown states are provided as inputs, or programs, for the program registers, and a third system, which is guaranteed to be prepared in one of the states stored in the program registers, is fed into the data register of the device. The device will then, with some probability of success, tell us whether the unknown state in the data register matches the state stored in the first or the second program register. We show that the optimal device, i.e., the one that maximizes the probability of success, is universal. It does not depend on the actual unknown states that we wish to discriminate.

Given two unknown quantum states, $|\psi_1\rangle$ and $|\psi_2\rangle$, we can construct a device that will unambiguously discriminate between them. If this device is given a system in one of these two states, it will produce one of three outputs: 1, 2, or 0. If the output is 1, then the input was $|\psi_1\rangle$; if the output is 2, then the input was $|\psi_2\rangle$; and if the output is 0, which we call failure, then we learn nothing about the input. The device will not make an error, and it will never give an output of 2 if the input was $|\psi_1\rangle$, and vice versa. This strategy is called unambiguous discrimination. The input states are not necessarily orthogonal; in fact, they can be completely arbitrary within the constraint that they are linearly independent [1]. The cost associated with this condition is that the probability of receiving the output 0 (failure) is not zero. The minimum value of this probability for two known and equally likely states is $|\langle\psi_1|\psi_2\rangle|$ (Refs. [2–4]).

The actual state-distinguishing device for two *known* states depends on the two states, $|\psi_1\rangle$ and $|\psi_2\rangle$; i.e., these two states are "hard wired" into the machine. What we would like to do here is see if we can construct a machine in which the information about $|\psi_1\rangle$ and $|\psi_2\rangle$ is supplied in the form of a program. This machine would be capable, with the correct program, of distinguishing any two quantum states. One such device has been proposed by Dušek and Bužek [5]. This device distinguishes the two states $\cos(\phi/2)|0\rangle \pm \sin(\phi/2)|1\rangle$. The angle ϕ is encoded into a one-qubit program state in a somewhat complicated way. The performance of this device is good; it does not achieve the maximum possible success probability for all input states, but its success probability, averaged over the angle ϕ, is greater than 90% of the optimal value. In a series of recent works Fiurášek *et al.* investigated a closely related programmable device that can perform a von Neumann projective measurement in any basis, the basis being specified by the program. Both deterministic and probabilistic approaches were explored [6], and experimental versions of both the state discriminator and the projective measure-

ment device were realized [7]. Sasaki *et al.* developed a related device, which they called a quantum matching machine [8]. Its input consists of K copies of two equatorial qubit states, which are called templates, and N copies of another equatorial qubit state $|f\rangle$. The device determines to which of the two template states $|f\rangle$ is closest. This device does not employ the unambiguous discrimination strategy, but optimizes an average score that is related to the fidelity of the template states and $|f\rangle$. Programmable quantum devices to accomplish other tasks have recently been explored by a number of authors [9–15].

Here we want to construct a programmable state discriminating machine whose program is related in a simple way to the states $|\psi_1\rangle$ and $|\psi_2\rangle$ that we are trying to distinguish. A motivation for our problem is that the program state may be the result of a previous set of operations in a quantum information processing device, and it would be easier to produce a state in which the information about $|\psi_1\rangle$ and $|\psi_2\rangle$ is encoded in a simple way than one in which the encoding is more complicated.

We shall, therefore, consider the following problem which is perhaps the simplest version of a programmable state discriminator. The program consists of the two qubit states that we wish to distinguish. In other words, we are given two qubits: one in the state $|\psi_1\rangle$ and another in the state $|\psi_2\rangle$. We have no knowledge of the states $|\psi_1\rangle$ and $|\psi_2\rangle$. Then we are given a third qubit that is guaranteed to be in one of these two program states, and our task is to determine, as best as we can, in which one. We are allowed to fail, but not to make a mistake. What is the best procedure to accomplish this?

We shall consider the first two qubits we are given as a program. They are fed into the program register of some device, called the programmable state discriminator, and the third, unknown qubit is fed into the data register of this device. The device then tells us, with optimal probability of success, to which of the two program states the unknown state of the qubit in the data register corresponds. We can

design such a device by viewing our problem as a task in measurement optimization. We want to find a measurement strategy that, with maximal probability of success, will tell us which one of the two program states, stored in the program register, matches the unknown state, stored in the data register. Our measurement is allowed to return an inconclusive result but never an erroneous one. Thus, it will be described by a POVM (positive-operator-valued measure) that will return 1 (the unknown state stored in the data register matches $|\psi_1\rangle$), 2 (the unknown state stored in the data register matches $|\psi_2\rangle$), or 0 (we do not learn anything about the unknown state stored in the data register).

Our task is then reduced to the following measurement optimization problem. One has two input states

$$|\Psi_1^{\text{in}}\rangle = |\psi_1\rangle_A |\psi_2\rangle_B |\psi_1\rangle_C, \qquad |\Psi_2^{\text{in}}\rangle = |\psi_1\rangle_A |\psi_2\rangle_B |\psi_2\rangle_C, \tag{1}$$

where the subscripts A and B refer to the program registers (A contains $|\psi_1\rangle$ and B contains $|\psi_2\rangle$), and the subscript C refers to the data register. Our goal is to unambiguously distinguish between these inputs, keeping in mind that one has no knowledge of $|\psi_1\rangle$ and $|\psi_2\rangle$; i.e., we want to find a POVM that will accomplish this.

Let the elements of our POVM be Π_1, corresponding to unambiguously detecting $|\Psi_1^{\text{in}}\rangle$, Π_2, corresponding to unambiguously detecting $|\Psi_2^{\text{in}}\rangle$, and Π_0, corresponding to failure. The probabilities of successfully identifying the two possible input states are given by

$$\langle \Psi_1^{\text{in}} | \Pi_1 | \Psi_1^{\text{in}} \rangle = p_1 \langle \Psi_2^{\text{in}} | \Pi_2 | \Psi_2^{\text{in}} \rangle = p_2, \tag{2}$$

and the condition of no errors implies that

$$\Pi_2 | \Psi_1^{\text{in}} \rangle = 0 \Pi_1 | \Psi_2^{\text{in}} \rangle = 0. \tag{3}$$

In addition, because the alternatives represented by the POVM exhaust all possibilities, we have that

$$I = \Pi_1 + \Pi_2 + \Pi_0. \tag{4}$$

The fact that we know nothing about $|\psi_1\rangle$ and $|\psi_2\rangle$ means that the only way we can guarantee satisfying the above conditions is to take advantage of the symmetry properties of the states, i.e., that $|\Psi_1^{\text{in}}\rangle$ is invariant under interchange of the first and third qubits, and $|\Psi_2^{\text{in}}\rangle$ is invariant under interchange of the second and third qubits. That unknown states can be unambiguously compared with a nonzero probability of success, using symmetry considerations only, has been first pointed out by Barnett *et al.* [16]. In our case, we require that Π_1 give zero when acting on states that are symmetric in qubits B and C, while Π_2 give zero when acting on states that are symmetric in qubits A and C. Defining the antisymmetric states for the corresponding pairs of qubits,

$$|\psi_{BC}^{(-)}\rangle = \frac{1}{\sqrt{2}} (|0\rangle_B |1\rangle_C - |1\rangle_B |0\rangle_C),$$

$$|\psi_{AC}^{(-)}\rangle = \frac{1}{\sqrt{2}} (|0\rangle_A |1\rangle_C - |1\rangle_A |0\rangle_C), \tag{5}$$

we introduce the projectors to the antisymmetric subspaces of the corresponding qubits as

$$P_{BC}^{\text{as}} = |\psi_{BC}^{(-)}\rangle\langle\psi_{BC}^{(-)}|, \qquad P_{AC}^{\text{as}} = |\psi_{AC}^{(-)}\rangle\langle\psi_{AC}^{(-)}|. \tag{6}$$

We can now take for Π_1 and Π_2 the operators

$$\Pi_1 = c_1 I_A \otimes P_{BC}^{\text{as}}, \qquad \Pi_2 = c_2 I_B \otimes P_{AC}^{\text{as}}, \tag{7}$$

where I_A and I_B are the identity operators on the spaces of qubits A and B, respectively, and c_1 and c_2 are as yet undetermined non-negative real numbers. The no-error condition dictates that $\Pi_1 = Q_A \otimes P_{BC}^{\text{as}}$ and $\Pi_2 = Q_B \otimes P_{AC}^{\text{as}}$, and it can be shown that the unknown operators Q_A and Q_B can be chosen to be proportional to the identity [17]. Using the above expressions for Π_j, where $j = 1, 2$ in Eq. (2), we find that

$$p_j = \langle \Psi_j^{\text{in}} | \Pi_j | \Psi_j^{\text{in}} \rangle = c_j \frac{1}{2} (1 - |\langle\psi_1|\psi_2\rangle|^2). \tag{8}$$

The average probability, P, of successfully determining which state we have, assuming that the input states occur with a probability of η_1 and η_2, respectively, is given by

$$P = \eta_1 p_1 + \eta_2 p_2 = \frac{1}{2} (\eta_1 c_1 + \eta_2 c_2)(1 - |\langle\psi_1|\psi_2\rangle|^2), \tag{9}$$

and we want to maximize this expression subject to the constraint that $\Pi_0 = I - \Pi_1 - \Pi_2$ is a positive operator.

Let S be the four-dimensional subspace of the entire eight-dimensional Hilbert space of the qubits A, B, and C, that is spanned by the vectors $|0\rangle_A |\psi_{BC}^{(-)}\rangle$, $|1\rangle_A |\psi_{BC}^{(-)}\rangle$, $|0\rangle_B |\psi_{AC}^{(-)}\rangle$, and $|1\rangle_B |\psi_{AC}^{(-)}\rangle$. In the orthogonal complement of S, $S^\perp$, the operator Π_0 acts as the identity, so that, in $S^\perp$, Π_0 is positive. Therefore, we need to investigate its action in S. First, let us construct an orthonormal basis for S. Applying the Gram-Schmidt process to the four vectors, given above, that span S, we obtain the orthonormal basis

$$|\Phi_1\rangle = |0\rangle_A |\psi_{BC}^{(-)}\rangle,$$

$$|\Phi_2\rangle = \frac{1}{\sqrt{3}} (2|0\rangle_B |\psi_{AC}^{(-)}\rangle - |0\rangle_A |\psi_{BC}^{(-)}\rangle),$$

$$|\Phi_3\rangle = |1\rangle_A |\psi_{BC}^{(-)}\rangle,$$

$$|\Phi_4\rangle = \frac{1}{\sqrt{3}} (2|1\rangle_B |\psi_{AC}^{(-)}\rangle - |1\rangle_A |\psi_{BC}^{(-)}\rangle). \tag{10}$$

In this basis, the operator Π_0, restricted to the subspace S, is given by the 4×4 matrix

$$\Pi_0 = \begin{pmatrix} 1 - c_1 - \frac{1}{4}c_2 & -\frac{\sqrt{3}}{4}c_2 & 0 & 0 \\ -\frac{\sqrt{3}}{4}c_2 & 1 - \frac{3}{4}c_2 & 0 & 0 \\ 0 & 0 & 1 - c_1 - \frac{1}{4}c_2 & -\frac{\sqrt{3}}{4}c_2 \\ 0 & 0 & -\frac{\sqrt{3}}{4}c_2 & 1 - \frac{3}{4}c_2 \end{pmatrix}.$$

(11)

Because of the block diagonal nature of Π_0, the characteristic equation for its eigenvalues, λ, is given by the biquadratic equation

$$[\lambda^2 - (2 - c_1 - c_2)\lambda + 1 - c_1 - c_2 + \frac{3}{4}c_1c_2]^2 = 0.$$

(12)

It is easy to obtain the eigenvalues explicitly. For our purposes, however, the conditions for their nonnegativity are more useful. These can be read out from the above equation, yielding

$$2 - c_1 - c_2 \geq 0, \qquad 1 - c_1 - c_2 + \frac{3}{4}c_1c_2 \geq 0. \quad (13)$$

The second is the stronger of the two conditions. When it is satisfied the first one is always met but the first one can still be used to eliminate nonphysical solutions. We can use the second condition to express c_2 in terms of c_1,

$$c_2 \leq \frac{1 - c_1}{1 - (3/4)c_1}.$$

(14)

For maximum probability of success, we chose the equal sign. Inserting the resulting expression into (9) gives

$$P = \frac{1}{2}\left(\eta_1 c_1 + \eta_2 \frac{1 - 1c_1}{1 - (3/4)c_1}\right)(1 - |\langle\psi_1|\psi_2\rangle|^2). \quad (15)$$

We can easily find $c_1 = c_{1,\text{opt}}$ where the right-hand side of this expression is maximum, and using this together with Eq. (14) we obtain

$$c_{1,\text{opt}} = \frac{2}{3}\left(2 - \sqrt{\frac{\eta_2}{\eta_1}}\right) c_{2,\text{opt}} = \frac{2}{3}\left(2 - \sqrt{\frac{\eta_1}{\eta_2}}\right). \quad (16)$$

Inserting these optimal values into (9) gives

$$P_{\text{POVM}} = \frac{2}{3}(1 - \sqrt{\eta_1\eta_2})(1 - |\langle\psi_1|\psi_2\rangle|^2). \quad (17)$$

This is not the full story, however. The above expression is valid only when $c_{1,\text{opt}}$ and $c_{2,\text{opt}}$ are both non-negative. From Eq. (16) it is easy to see that this holds if

$$\frac{1}{5} \leq \eta_1, \qquad \eta_2 \leq \frac{4}{5}. \quad (18)$$

In order to understand what happens outside this interval, we have to turn our attention to the detection operators. Using $c_{1,\text{opt}}$ and $c_{2,\text{opt}}$ in Eq. (7) yields

$$\Pi_{1,\text{opt}} = \frac{2}{3}\left(2 - \sqrt{\frac{\eta_2}{\eta_1}}\right)I_A \otimes P_{BC}^{\text{as}},$$

$$\Pi_{2,\text{opt}} = \frac{2}{3}\left(2 - \sqrt{\frac{\eta_1}{\eta_2}}\right)I_B \otimes P_{AC}^{\text{as}}.$$

(19)

For $\eta_1 = \frac{4}{5}$ (and $\eta_2 = \frac{1}{5}$), $\Pi_{1,\text{opt}} = I_A P_{BC}^{\text{as}}$ and $\Pi_{2,\text{opt}} = 0$. This structure then remains valid for $\eta_1 \geq \frac{4}{5}$. In other words, when the first input dominates the preparation, it is advantageous to use the full projector that distinguishes it with maximal probability of success, $p_{1,\text{opt}} = \frac{1}{2} \times (1 - |\langle\psi_1|\psi_2\rangle|^2)$, at the expense of sacrificing the second input completely, $p_{2,\text{opt}} = 0$. These values yield the average success probability,

$$P_1 = \frac{1}{2}\eta_1(1 - |\langle\psi_1|\psi_2\rangle|^2), \quad (20)$$

for $\eta_1 \geq \frac{4}{5}$. Conversely, for $\eta_2 = \frac{4}{5}$, $\Pi_{2,\text{opt}} = I_B P_{AC}^{\text{as}}$ and $\Pi_{1,\text{opt}} = 0$. This structure then remains valid for $\eta_2 \geq \frac{4}{5}$. So, when the second input dominates the preparation it is advantageous to use the full projector that distinguishes it with maximal probability of success, $p_{2,\text{opt}} = \frac{1}{2} \times (1 - |\langle\psi_1|\psi_2\rangle|^2)$, at the expense of sacrificing the first input completely, $p_{1,\text{opt}} = 0$. These values yield the average success probability,

$$P_2 = \frac{1}{2}\eta_2(1 - |\langle\psi_1|\psi_2\rangle|^2), \quad (21)$$

for $\eta_2 \geq \frac{4}{5}$. As we see, the situation is fully symmetric in the inputs and *a priori* probabilities. In the intermediate range, neither one of the inputs dominates the preparation, and we want to identify them as best as we can, so the POVM solution will do the job there. Our findings can be summarized as follows

$$P^{\text{opt}} = \begin{cases} P_{\text{POVM}} & \text{if } \frac{1}{5} \leq \eta_1 \leq \frac{4}{5}, \\ P_2 & \text{if } \eta_1 < \frac{1}{5}, \\ P_1 & \text{if } \frac{4}{5} < \eta_1. \end{cases} \quad (22)$$

Equation (22) represents our main result. In the intermediate range of the *a priori* probability, the optimal failure probability, Eq. (17), is achieved by a generalized measurement or POVM. Outside this region, for a very small *a priori* probability, $\eta_1 \leq 1/5$, when the preparation is dominated by the second input, or a very large *a priori* probability, $\eta_1 \geq 4/5$, when the preparation is dominated by the first input, the optimal failure probabilities, Eqs. (20) and (21), are realized by standard von Neumann measurements. For very small η_1 the optimal von Neumann measurement is a projection onto the antisymmetric subspace of the A and C qubits. For very large η_1 the optimal von Neumann measurement is a projection onto the antisymmetric subspace of the B and C qubits. At the boundaries of their respective regions of validity, the optimal measurements transform into one another contin-
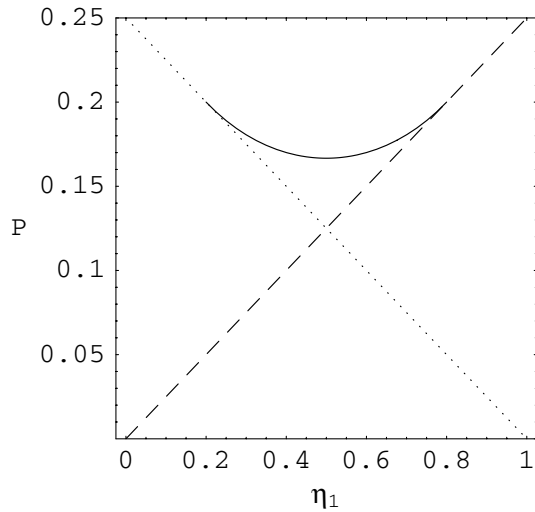
FIG. 1.   Optimal average success probability, P, vs the *a priori* probability, η_1. Dashed line, P_1 from Eq. (20); dotted line, P_2 from Eq. (21); solid line, P_{POVM} from Eq. (17). For the figure we replaced $1 - |\langle\psi_1|\psi_2\rangle|^2$ by its average, $\frac{1}{2}$. The optimal P is given by P_2 for $\eta_1 < 0.2$, by P_{POVM} for $0.2 \leq \eta_1 \leq 0.8$, and by P_1 for $0.8 < \eta_1$.

uously. We also see that the results depend on the overlap of the unknown states only. If we do not know the states but we know their overlap, then Eqs. (17), (20), and (21) immediately give the optimal solutions for this situation. If we know nothing about the states, not even their overlap, then we average these expressions over all input states, which results in the factor, $1 - |\langle\psi_1|\psi_2\rangle|^2$, being replaced by its average value of $\frac{1}{2}$. Then we have the optimum average probabilities of success in the various regions. This situation is depicted in Fig. 1.

In its range of validity the POVM performs better than any von Neumann measurement that does not introduce errors. From the figure it also can be read out that the difference between the performance of the POVM and that of the von Neumann projective measurements is largest for $\eta_1 = \eta_2 = \frac{1}{2}$. For these values $P_{\text{POVM}}^{\text{ave}} = \frac{1}{6}$ while $P_1^{\text{ave}} = P_2^{\text{ave}} = \frac{1}{8}$, so the POVM represents a 33% improvement over the standard quantum measurement.

Finally, we want to point out a striking feature of the programmable state discriminator. Neither the optimal detection operators nor the boundaries for their region of validity, Eqs. (18) and (19), depend on the unknown states. Therefore, our device is *universal*; it will perform optimally for any pair of unknown states. Only the probability of success for fixed but unknown states will depend on the overlap of the states.

This POVM, then, provides us with the best procedure for solving the problem posed at the beginning of this Letter. It also demonstrates the role played by *a priori* information. This device has a smaller success probability than one designed for a case in which we know one of the input states [17], which in turn has a smaller success probability than one designed for the case when we know both possible input states. While its success probability is lower than that for a device that distinguishes known states, the device discussed here is more flexible. All of the information about the states is carried by a quantum program, which means that it works for any two states. Consequently, it can be used as part of a larger device that produces quantum states that need to be unambiguously identified.

[1]   A. Chefles, Phys. Lett. A **239**, 339 (1998).
[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]   M. Dušek and V. Bužek, Phys. Rev. A **66**, 022112 (2002).
[6]   J. Fiurášek, M. Dušek, and R. Filip, Phys. Rev. Lett. **89**, 190401 (2002); J. Fiurášek and M. Dušek, Phys. Rev. A **69**, 032302 (2004).
[7]   J. Soubusta, A. Černoch, J. Fiurášek, and M. Dušek, Phys. Rev. A **69**, 052321 (2004).
[8]   M. Sasaki and A. Carlini, Phys. Rev. A **66**, 022303 (2002); M. Sasaki, A. Carlini, and R. Jozsa, Phys. Rev. A **64**, 022317 (2001).
[9]   M. Nielsen and I. Chuang, Phys. Rev. Lett. **79**, 321 (1997).
[10]   J. Preskill, Proc. R. Soc. London A **454**, 385 (1998).
[11]   M. Hillery, V. Bužek, and M. Ziman, Phys. Rev. A **65**, 022301 (2002).
[12]   G. Vidal, L. Masanes, and I. Cirac, Phys. Rev. Lett. **88**, 047905 (2002).
[13]   M. Hillery, V. Bužek, and M. Ziman, Phys. Rev. A **66**, 042302 (2002).
[14]   A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek, Phys. Rev. Lett. **88**, 217901 (2002).
[15]   J. Paz and A. Roncaglia, Phys. Rev. A **68**, 052316 (2003).
[16]   S. M. Barnett, A. Chefles, and I. Jex, Phys. Lett. A **307**, 189 (2003).
[17]   J. Bergou, M. Hillery, and V. Bužek (to be published).