Quantum-state filtering applied to the discrimination of Boolean functions

János A. Bergou and Mark Hillery
Department of Physics, Hunter College of the City University of New York, 695 Park Avenue, New York, New York 10021, USA

(Received 26 December 2004; published 5 July 2005)

Quantum-state filtering is a variant of the unambiguous state discrimination problem: the states are grouped in sets and we want to determine to which particular set a given input state belongs. The simplest case, when the N given states are divided into two subsets and the first set consists of one state only while the second consists of all of the remaining states, is termed quantum-state filtering. We derived previously the optimal strategy for the case of N nonorthogonal states, $\{|\psi_1\rangle,\ldots,|\psi_N\rangle\}$, for distinguishing $|\psi_1\rangle$ from the set $\{|\psi_2\rangle,\ldots,|\psi_N\rangle\}$ and the corresponding optimal success and failure probabilities. In a previous paper [Phys. Rev. Lett. **90**, 257901 (2003)], we sketched an appplication of the results to probabilistic quantum algorithms. Here we fill the gaps and give the complete derivation of the probabilstic quantum algorithm that can optimally distinguish between two classes of Boolean functions, that of the balanced functions and that of the biased functions. The algorithm is probabilistic, it fails sometimes but when it does it lets us know that it did. Our approach can be considered as a generalization of the Deutsch-Jozsa algorithm that was developed for the discrimination of balanced and constant Boolean functions.

DOI: 10.1103/PhysRevA.72.012302 PACS number(s): 03.67.Lx

I. INTRODUCTION

The Deutsch-Jozsa algorithm was one of the first quantum algorithms [1]. It makes it possible to perform on a quantum computer in one step a task that, on a classical computer, would require many steps. In particular, one is given a "black box" that, when given an n bit input, returns either 0 or 1. The box simply evaluates a Boolean function, that is, a function that maps n-bit binary numbers to the set $\{0,1\}$. It is guaranteed that the functions are of one of two types: they are either constant, giving the same value for all inputs, or balanced, giving 0 for half the inputs and 1 for the other half. The task is to determine whether the function we have been given is constant or balanced. On a classical computer, we would, in the worst case, have to evaluate the function $2^{(n-1)}+1$ times to determine this. On a quantum computer the determination can be made with certainty with only one evaluation of the function.

The quantum algorithm works by mapping the function onto a quantum state. The quantum states for constant and balanced functions lie in orthogonal subspaces, and they can therefore be distinguished perfectly. An obvious question to ask is whether the method can be extended to distinguish functions whose vectors do not lie in orthogonal subspaces. Here we shall show that this is indeed the case, and to do so we shall make use of a procedure known as quantum-state filtering [2–4].

Quantum-state filtering considers the following problem. We are given a system that is in one of a known set of N quantum states $\{\psi_1,\ldots\psi_N\}$. Our task is to determine whether the system is in ψ_1 or not, that is, whether it is in the set $\{\psi_1\}$ or the set $\{\psi_2,\ldots\psi_N\}$. This is what quantum-state filtering accomplishes. It is a probabilistic procedure: it may fail, but we know when it does. The failure probability should be as small as possible, and how this is accomplished depends on the *a priori* probabilities with which the states $\{\psi_1,\ldots\psi_N\}$ appear and the overlaps of ψ_1 with the other states. In order to generalize the Deutsch-Jozsa algorithm, we can apply quantum-state filtering to the quantum states into which the Boolean functions are mapped.

Quantum-state filtering is an instance of mixed state discrimination, in particular the discrimination of a pure state from a mixed state. The set of states $\{\psi_2,\ldots\psi_N\}$ along with their *a priori* probabilities can be viewed as an ensemble, which can be represented by a density matrix. There are several different strategies one can adopt in discriminating mixed states. One can minimize the probability of making an error [5–7]. A second strategy is that of unambiguous discrimination in which one never makes a mistake in identifying the state, but one can sometimes fail to obtain any information about the state one was given [8–13]. Quantum-state filtering is an example of this approach. There are also hybrid strategies in which one can make mistakes and one can also fail to obtain an answer [14,15]. The addition of the option of failing to obtain an answer makes it possible to obtain a smaller error probability than is possible when the possibility of failing is not present.

The paper is organized as follows. In the next section we briefly review quantum state filtering, following Ref. [4] but with a notation simplified for the present purposes. In Sec. III we apply state filtering to the problem of distinguishing balanced functions from particular biased functions, i.e., functions that have a preponderance of zeroes or ones. In Sec. IV we provide a specific example of the positive-operator valued measure (POVM) that figures in the state filtering procedure, and we summarize our results in the Conclusion.

II. QUANTUM-STATE FILTERING

As was mentioned in the Introduction, what we wish to do is to determine whether a system we have been given is in the state ψ_1 or not, given that it must be in one of the states $\{\psi_1,\ldots\psi_N\}$ and that the state ψ_j occurs with probability η_j.

This can be accomplished in one of three ways, and which way is the best depends on the *a priori* probabilities and the overlaps of the states. Two of these methods are von Neumann projective measurements, and the third is a POVM.

The first projective measurement projects onto the subspace orthogonal to ψ_1. This is accomplished by the projection operator

$$F^{(1)} = I - |\psi_1\rangle\langle\psi_1|. \tag{2.1}$$

If we obtain the value 1, then the procedure has succeeded, and we know that the system we have was not in the state ψ_1. If we obtain 0, then the procedure has failed. This happens with a probability

$$Q_1 = \eta_1 + S, \tag{2.2}$$

where

$$S = \sum_{j=2}^{N} \eta_j |\langle\psi_1|\psi_j\rangle|^2 \tag{2.3}$$

is the average overlap between the two subsets.

A second possibility is to split $|\psi_1\rangle$ into two components, $|\psi_1\rangle = |\psi_1^{\perp}\rangle + |\psi_1^{\parallel}\rangle$. Here $|\psi_1^{\perp}\rangle$ is orthogonal to the subspace $\mathcal{H}_2$ that is spanned by the vectors $|\psi_2\rangle, \ldots |\psi_N\rangle$, and $|\psi_1^{\parallel}\rangle$ lies in $\mathcal{H}_2$. Their normalized versions are $|\tilde{\psi}_1^{\perp}\rangle = |\psi_1^{\perp}\rangle / \|\psi_1^{\perp}\|$ and $|\tilde{\psi}_1^{\parallel}\rangle = |\psi_1^{\parallel}\rangle / \|\psi_1^{\parallel}\|$, respectively. We now introduce the operator

$$F^{(2)} = |\tilde{\psi}_1^{\perp}\rangle\langle\tilde{\psi}_1^{\perp}| - (I - |\tilde{\psi}_1^{\perp}\rangle\langle\tilde{\psi}_1^{\perp}| - |\tilde{\psi}_1^{\parallel}\rangle\langle\tilde{\psi}_1^{\parallel}|), \tag{2.4}$$

which has eigenvalues 1, 0, and -1. If we measure $F^{(2)}$ and obtain 1, then the vector was $|\psi_1\rangle$, if we obtain -1, then the vector was in the set $\{|\psi_2\rangle, \ldots |\psi_N\rangle\}$, and if we obtain 0, the procedure failed. In this case the probability of failure Q_2 is given by

$$Q_2 = \eta_1 \|\psi_1^{\parallel}\|^2 + \frac{S}{\|\psi_1^{\parallel}\|^2}. \tag{2.5}$$

Which of these two particular strategies is better is determined by which of these two failure probabilities is smaller. In particular, $Q_1 > Q_2$ if $\eta_1 \|\psi_1^{\parallel}\|^2 > S$, and vice versa.

The third measurement is based on a positive-operator valued measure (POVM [16]), and it can do better in an intermediate range of parameters. The POVM can be implemented by a unitary evolution on a larger space and a selective measurement. The larger space consists of two orthogonal subspaces: the original system space and a failure space. The idea is that the unitary evolution transforms the input sets into orthogonal sets in the original system space and maps them onto the same vector in the failure space. A click in the detector measuring along this vector corresponds to failure of the procedure, since all inputs are mapped onto the same output. A no-click corresponds to success since now the nonorthogonal input sets are transformed into orthogonal output sets in the system space. The one-dimensionality of the failure space follows from the requirement that the filtering is optimum, as shown by the following simple considerations. Suppose that ψ_1 is mapped onto some vector in the failure space and the inputs from the other set are mapped onto vectors that have components perpendicular to this vector. Then a single von Neumann measurement along the orthogonal direction could identify the input as being from the second set, i.e., further filtering would be possible, lowering the failure probability and, contrary to our assumption, our original filtering could not have been optimal.

In particular, let $\mathcal{H}_S$ be the D-dimensional system space spanned by the vectors $\{|\psi_1\rangle, \ldots |\psi_N\rangle\}$ where, obviously, $D \leq N$. We now embed this space in a space of $D+1$ dimensions, $\mathcal{H}_{S+A} = \mathcal{H}_S \oplus \mathcal{H}_A$, where $\mathcal{H}_A$ is a one-dimensional auxiliary Hilbert space, the failure space or ancilla. The basis in this space is denoted by $|\phi_A\rangle$, where $\|\phi_A\| = 1$. Thus the unitary evolution on $\mathcal{H}_{S+A}$ is specified by the requirement that for any of the input states $|\psi_j\rangle$ $(j=1, \ldots, N)$ the final state has the structure

$$|\psi_j\rangle_{out} = U|\psi_j\rangle = \sqrt{p_j}|\psi_j'\rangle + \sqrt{q_j}e^{i\theta_j}|\phi_A\rangle, \tag{2.6}$$

where $|\psi_j'\rangle \in \mathcal{H}_S$, and $\|\psi_j\| = 1$. From unitarity the relation $p_j + q_j = 1$ follows. Furthermore, p_j is the probability that the transformation $|\psi_j\rangle \rightarrow |\psi_j'\rangle$ succeeds and q_j is the probability that $|\psi_j\rangle$ is mapped onto the state $|\phi_A\rangle$. In order to identify p_j and q_j with the state-specific success and failure probability for quantum filtering we have to require that

$$\langle\psi_1'|\psi_j'\rangle = 0, \tag{2.7}$$

for $j=2, \ldots N$. We can now set up $D+1$ detectors in the following way. One of them is directed along $|\psi_1'\rangle$, $D-1$ along the remaining $D-1$ orthogonal directions in the original system space and the last one along $|\phi_A\rangle$ in the failure space. If any of the system-space detectors clicks we can uniquely assign the state we were given to one or the other subset and a click in the failure-space detector indicates that filtering has failed. It should be noted that, in general, for the unitary operator given in Eq. (2.6) to exist we must have $D=N$ [3].

In order to optimize the POVM, we have to determine those values of q_j in Eq. (2.6) that yield the smallest average failure probability Q, where

$$Q = \sum_{j=1}^{N} \eta_j q_j. \tag{2.8}$$

Taking the scalar product of $U|\psi_1\rangle$ and $U|\psi_j\rangle$ in Eq. (2.6), and using Eq. (2.7), gives

$$|\langle\psi_1|\psi_j\rangle|^2 = q_1 q_j \tag{2.9}$$

for $j=2, \ldots, N$, and Eq. (2.8) can be cast in the form $Q(q_1) = \eta_1 q_1 + S/q_1$. Unitarity of the transformation U delivers the necessary condition that q_1 must lie in the range $\|\psi_1^{\parallel}\|^2 \leq q_1 \leq 1$. Details of the derivation, along with a discussion of the sufficient conditions for the existence of U, can be found in Ref. [3]. Provided that a POVM solution exists, the minimum of $Q(q_1)$ is reached for $q_1 = \sqrt{S/\eta_1}$ and is given by

$$Q_{\text{POVM}} = 2\sqrt{\eta_1 S}. \tag{2.10}$$

Thus the failure probability for optimal unambiguous quantum-state filtering can be summarized as

$$Q=\begin{cases} 2\sqrt{\eta_1 S} \quad \text{if } \eta_1\|\psi_1^{\parallel}\|^4 \leqslant S \leqslant \eta_1, \\ \eta_1 + S \quad \text{if } S > \eta_1, \\ \eta_1\|\psi_1^{\parallel}\|^2 + \dfrac{S}{\|\psi_1^{\parallel}\|^2} \quad \text{if } S < \eta_1\|\psi_1^{\parallel}\|^4. \end{cases} \tag{2.11}$$

The first line represents the POVM result, Eq. (2.10), and it gives a smaller failure probability, in its range of validity, than the von Neumann measurements, Eqs. (2.2) and (2.5). Outside of the POVM range of validity we recover the von Neumann results. It should be noted that for these results to hold, unlike for unambiguous state discrimination, linear independence of all states is not required. Instead, the less stringent requirement of the linear independence of the sets is sufficient, in agreement with the findings in Ref. [17].

It is useful to see what happens if the POVM is applied to a vector that lies in the subspace $\mathcal{H}_2$ spanned by the vectors $\{\psi_2, \ldots \psi_N\}$, but is not one of these vectors. Let

$$|\psi\rangle = \sum_{j=2}^{N} c_j |\psi_j\rangle, \tag{2.12}$$

so that

$$U|\psi\rangle = \sqrt{p}|\psi'\rangle + \sqrt{q}e^{i\theta}|\phi_A\rangle, \tag{2.13}$$

where $\|\psi'\|=1$, and

$$\sqrt{p}|\psi'\rangle = \sum_{j=2}^{N} c_j \sqrt{p_j}|\psi_j'\rangle,$$

$$\sqrt{q}e^{i\theta} = \sum_{j=2}^{N} c_j \sqrt{q_j}e^{i\theta_j}. \tag{2.14}$$

We note that p is the probability that we would determine that the state ψ is in $\mathcal{H}_2$, and q is the probability that we would fail to do so. Taking the inner product of the above equation with $U|\psi_1\rangle$ gives us that

$$q_1 q = |\langle\psi_1|\psi\rangle|^2. \tag{2.15}$$

Therefore the POVM can be used to solve a more general problem than the one for which it was designed. In particular, suppose we are given a system that is guaranteed to be either in the state ψ_1 or in a state in the subspace $\mathcal{H}_2$, and we want to determine which is the case. We can do this by applying the POVM presented here, and our probability of failing is given by the above equation. Because the POVM was designed for a particular basis of $\mathcal{H}_2$, it will not necessarily be optimal for the more general problem, but it will work.

III. APPLICATION TO A BASIS FOR BOOLEAN FUNCTIONS

We can now apply this result to distinguishing between sets of Boolean functions. Let $f(x)$, where $0 \leqslant x \leqslant 2^n - 1$, be a Boolean function, i.e., $f(x)$ is either 0 or 1. One of the sets we want to consider is the set of balanced functions. The other will consist of two "biased" functions. We shall call a function biased if it is not balanced or constant, i.e., if it returns 0 on m_0 of its arguments, 1 on $m_1 = 2^n - m_0$, and $m_0 \neq m_1 \neq 0$ or 2^{n-1}. In order to discriminate a particular biased function from an unknown balanced one $2^{(n-1)} + m_1 + 1$ function evaluations are necessary in the worst case classically, where we have assumed, without loss of generality, that $m_1 < m_0$. This number comes from the fact that there are balanced functions that agree with our biased function in $2^{(n-1)} + m_1$ places. Suppose that a balanced function is 1 for every argument for which our biased function is 1. This means that they agree for at least m_1 arguments. Of the remaining arguments, the biased function will be 0 on all of them, and the balanced function will be 0 on $2^{(n-1)}$ of them. This means that the functions agree in a total of $2^{(n-1)} + m_1$ places.

As has been mentioned, the second set of functions we shall consider has only two members, and we shall call it W_k. A function is in W_k if $f(x)=0$ for $0 \leqslant x < [(2^k - 1)/2^k]2^n$ and $f(x)=1$ for $[(2^k-1)/2^k]2^n \leqslant x \leqslant 2^n - 1$, or if $f(x)=1$ for $0 \leqslant x < [(2^k-1)/2^k]2^n$ and $f(x)=0$ for $[(2^k-1)/2^k]2^n \leqslant x \leqslant 2^n - 1$. The problem we wish to consider is distinguishing between balanced functions and functions in W_k, that is, we are given an unknown function that is in one of the two sets, and we want to find out which set it is in. We note that the two functions in W_k are biased functions, so that this problem is a particular instance of a more general problem of distinguishing a particular set of biased functions from balanced functions. This is by no means the only example the method we are proposing here can handle, but it is a particularly simple one.

This is clearly a variant of the Deutsch-Jozsa problem [1]. In that case one is given an unknown function that is either balanced or constant, and one wants to determine which. Classically, in the worst case one would have to evaluate the function $D/2+1$ times, where we have set $D=2^n$, but in the quantum case only one function evaluation is necessary. The solution of this problem makes use of the unitary mapping

$$|x\rangle|y\rangle \rightarrow |x\rangle|y + f(x)\rangle, \tag{3.1}$$

where the first state, $|x\rangle$, is an n-qubit state, the second state, $|y\rangle$, is a single qubit state, and the addition is modulo 2. The state $|x\rangle$, where x is an n-digit binary number, is a member of the computational basis for n qubits, and the state $|y\rangle$, where y is either 0 or 1, is a member of the computational basis for a single qubit. In solving the Deutsch-Jozsa problem, this mapping is employed in the following way:

$$\frac{1}{\sqrt{2D}}\sum_{x=0}^{D-1} |x\rangle(|0\rangle - |1\rangle) \rightarrow \frac{1}{\sqrt{2D}}\sum_{x=0}^{D-1} (-1)^{f(x)}|x\rangle(|0\rangle - |1\rangle), \tag{3.2}$$

and we shall do the same. This has the effect of mapping Boolean functions to vectors in the D-dimensional Hilbert space $\mathcal{H}_D$; the final qubit is not entangled with the remaining n qubits and can be discarded. The vectors $\Sigma_{x=0}^{D-1}(-1)^{f(x)}|x\rangle$ that are produced by balanced functions are orthogonal to those produced by constant functions. This is why the Deutsch-Jozsa problem is easy to solve quantum mechanically. In our case, the vectors produced by functions in W_k are not orthogonal to those produced by balanced functions. However, we can solve the problem probabilistically by using our unambiguous state filtering procedure.

In order to apply quantum state filtering to this problem, we note that both functions in W_k are mapped, up to an overall sign, to the same vector in $\mathcal{H}_D$, which we shall call $|w_k\rangle$. The vectors that correspond to balanced functions are contained in the subspace $\mathcal{H}_b$ of $\mathcal{H}_D$, where $\mathcal{H}_b=\{|v\rangle \in \mathcal{H}_D | \Sigma_{x=0}^{D-1} v_x=0\}$, and $v_x=\langle x|v\rangle$. This subspace has dimension 2^n-1, and it is possible to choose an orthonormal basis, $\{|v_j\rangle | j=2,\ldots D\}$, for it in which each basis element corresponds to a particular balanced Boolean function. As any state in $\mathcal{H}_b$ is a linear combination of the $|v_j\rangle$ basis vectors, a state filtering procedure that can optimally distinguish $|w_k\rangle$ from the set $\{|v_j\rangle\}$ will also discriminate $|w_k\rangle$ from any set of states in $\mathcal{H}_b$, though not necessarily optimally.

Let us now construct the basis for $\mathcal{H}_D$ that we wish to use. This will be done inductively. For $D=2$, choose the basis vectors to be

$$|v_1^{(1)}\rangle = \frac{1}{\sqrt{2}}(1,1) \quad |v_2^{(1)}\rangle = \frac{1}{\sqrt{2}}(1,-1). \tag{3.3}$$

Note that $|v_2^{(1)}\rangle$ is the vector that corresponds to the two balanced one-bit Boolean functions, and $|v_1^{(1)}\rangle$ corresponds to the two constant functions. These vectors are orthonormal. For $D=2^2$ the basis vectors are

$$|v_1^{(2)}\rangle = \frac{1}{2}(1,1,1,1) \quad |v_2^{(2)}\rangle = \frac{1}{2}(1,1,-1,-1),$$

$$|v_3^{(2)}\rangle = \frac{1}{2}(1,-1,1,-1) \quad |v_4^{(2)}\rangle = \frac{1}{2}(1,-1,-1,1). \tag{3.4}$$

The vector $|v_1^{(2)}\rangle$ corresponds to the two constant functions, and the other basis elements correspond to balanced functions.

Now let us construct a basis for arbitrary $D=2^n$. First, suppose that we have all of the bases $\{v_j^m | j=1,\ldots 2^m\}$ for $m \leq n-1$. We now want to construct the basis for $D=2^n$. Denote a string of p 1's by $u(p)$. Our first basis vector is just

$$v_1^{(n)} = \frac{1}{\sqrt{D}}(u(D)), \tag{3.5}$$

which just corresponds to a constant function on n bits, and our second is

$$v_2^{(n)} = \frac{1}{\sqrt{D}}(u(D/2),-u(D/2)). \tag{3.6}$$

The general basis vector is composed of blocks of 1's and -1's. It can be expressed as

$$\frac{1}{2^{(n-p+1)/2}}[(v_j^{(p-1)})_1(u(D/2^p),-u(D/2^p)),\ldots(v_j^{(p-1)})_{2^{(p-1)}} \times(u(D/2^p),-u(D/2^p))], \tag{3.7}$$

where $(v_j^{(k-1)})_q$ is just the qth component of $v_j^{(k-1)}$. Here we have that j runs from 1 to 2^{p-1}, and p goes from 2 to $n-1$. That these vectors are orthogonal for different values of p can be seen rather easily. Let the first vector have $p=p_1$ and the second have $p=p_2$, where $p_1 < p_2$. When taking the inner product, each block of the form $(u(D/2^{p_1}),-u(D/2^{p_1}))$, in the first vector will be paired with a block of uniform 1's or -1's in the second vector. This means that each $(u(D/2^{p_1}), -u(D/2^{p_1}))$ block in the first vector will contribute a zero to the total inner product, so that the entire inner product is itself just zero. When taking the inner product of two vectors for which the values of p are the same, we see that the inner product is just proportional to $\langle v_{j_1}^{(p-1)} | v_{j_2}^{(p-1)}\rangle$, where j_1 and j_2 are the j values for the two vectors. Therefore if $j_1 \neq j_2$ the vectors are orthogonal. Henceforth we shall denote the vector in Eq. (3.7) by $v_{p,j}$ where j runs from 1 to 2^{p-1}, and p goes from 2 to $n-1$. The vector in Eq. (3.5) will be denoted as $v_{0,1}$, and the vector in Eq. (3.6) as $v_{1,1}$.

Let us first see how the filtering procedure performs when applied to the problem of distinguishing $|w_k\rangle$ from the $D-1$ orthonormal basis states $|v_{p,j}\rangle$, for $p>0$, in $\mathcal{H}_b$. We assume that their *a priori* probabilities are equal, and we shall denote this probability by η, where $\eta=(1-\eta_1)/(2^n-1)$ and η_1 is the *a priori* probability for the function to be in W_k. Together with the state $|w_k\rangle$, the total number of states is D, the dimension of the system Hilbert space.

In order to find which of the three measurement procedures is optimal, we need to calculate both S and $\|w_k^{\|}\|$. Now $|w_k\rangle$ is a unit vector, so that the sum of the square of its component along $|v_{0,1}\rangle$ and $\|w_k^{\|}\|^2$ is 1. Therefore we have that

$$\|w_k^{\|}\|^2 = 1 - |\langle v_{0,1}|w_k\rangle|^2 = \frac{2^k-1}{2^{(2k-2)}}. \tag{3.8}$$

The calculation of S is particularly simple in this case. We have

$$\begin{aligned} S &= \eta \sum_{p=1}^{n-1} \sum_{j=1}^{2^{p-1}} |\langle v_{p,j}|w_k\rangle^2 \\ &= \eta(1-|\langle v_{0,1}|w_k\rangle|^2) \\ &= \frac{\eta(2^k-1)}{2^{(2k-2)}} \\ &= \eta\|w_k^{\|}\|^2. \end{aligned} \tag{3.9}$$

Substituting these quantities into the conditions in Eq. (2.11), we find that the first von Neumann measurement is optimal if $\eta_1 > \zeta_1$, the POVM is optimal if $\zeta_2 \leq \eta_1 \leq \zeta_1$, and the second von Neumann measurement is optimal if $\eta_1 < \zeta_2$ where

$$\zeta_1 = \left[1 + \frac{(2^n-1)(2^k-1)}{2^{2(k-1)}} \right]^{-1} \simeq 2^{-(n-k+2)},$$

$$\zeta_2 = \frac{2^k-1}{2^{2(k-1)}(2^n-1)+2^k-1} \simeq 2^{-(n+k-2)}, \qquad (3.10)$$

where, in the approximate expressions, we have assumed that $2^n \gg 2^k \gg 1$.

It is useful to get an idea of how much the failure probabilities of the different procedures differ. To do so, we will look in the range in which the POVM is optimal, and compare the failure probabilities of the three different measurements. In this range, the failure probabilities are given by

$$Q_{\text{POVM}} = \frac{1}{2^{k-2}} \left[\frac{\eta_1(1-\eta_1)(2^k-1)}{2^n-1} \right]^{1/2},$$

$$Q_1 = \eta_1 + \frac{(1-\eta_1)(2^k-1)}{2^{2k-2}(2^n-1)},$$

$$Q_2 = \frac{\eta_1(2^k-1)}{2^{2k-2}} + \frac{1-\eta_1}{2^n-1}. \qquad (3.11)$$

It should be noted that while Q_{POVM} in the above expression is valid only if $\zeta_2 \leqslant \eta_1 \leqslant \zeta_1$, the expressions for Q_1 and Q_2 are valid for any value of η_1. Assuming that $2^n \gg 2^k \gg 1$ we find that

$$\frac{Q_{\text{POVM}}}{Q_1} \simeq \frac{4}{\sqrt{2^{n+k}\eta_1}} \quad \frac{Q_{\text{POVM}}}{Q_2} \simeq 4\sqrt{2^{n-k}\eta_1}, \qquad (3.12)$$

so that the POVM result represents a considerable improvement over either of the von Neumann measurements when it is optimal. For example, in the case in which all of the *a priori* probabilities are equal, i.e., $\eta_1 = 1/2^n$, we have that

$$Q_{\text{POVM}} = \frac{(2^k-1)^{1/2}}{2^{n+k-2}} \simeq \frac{1}{2^{n-2+(k/2)}}, \qquad (3.13)$$

both of the ratios in Eq. (3.12) are $4/2^{k/2}$. For $k \gg 1$, this implies that the difference in performance between the POVM and the von Neumann measurements can be significant.

IV. EXAMPLE OF POVM

So far, our discussion of the POVM has been rather abstract; we know when it exists and when it does not, and we have described how it works. It is useful, however, to see explicitly how the unitary operator [see Eq. (2.6)] that plays a prominent role in the scheme can be constructed. Our task is to find explicit expressions for the vectors $|\psi'{}_j\rangle$. In order to do this, we shall consider an explicit example, the case $k = 2$.

We first note that $\langle w_2 | v_{p,j} \rangle = 0$ for all $p > 2$, which means that all of these vectors can be perfectly distinguished from $|w_2\rangle$. Therefore we define $U|v_{p,j}\rangle = |v_{p,j}\rangle$ for $p > 2$. U will map the remaining four-dimensional subspace, spanned by the vectors $|v_{p,j}\rangle$, with $p \leqslant 2$, into itself, and we can henceforth confine our attention to this subspace.

Let us now set

$$|\psi_1\rangle = |w_2\rangle \quad |\psi_2\rangle = |v_{1,1}\rangle,$$

$$|\psi_3\rangle = |v_{2,1}\rangle \quad |\psi\rangle = |v_{2,2}\rangle. \qquad (4.1)$$

The condition that guarantees the existence of U is that the matrix M, given by

$$M_{jk} = \langle \psi'_j | \psi'_k \rangle = \langle \psi_j | \psi_k \rangle - \sqrt{q_j q_k} e^{i(\theta_k - \theta_j)}, \qquad (4.2)$$

be positive. Taking the inner product of $U|\psi_j\rangle$ and $U|\psi_1\rangle$, $j \neq 1$, we find that

$$\langle \psi_j | \psi_1 \rangle = \sqrt{q_1 q_j} e^{i(\theta_1 - \theta_j)}, \qquad (4.3)$$

so that we can express M_{jk}, for $j,k > 1$, as

$$M_{jk} = \langle \psi_j | \psi_k \rangle - \frac{1}{q_1} \langle \psi_j | \psi_1 \rangle \langle \psi_1 | \psi_k \rangle. \qquad (4.4)$$

Setting $x = 1/(4q_1)$, we find that

$$M = \begin{pmatrix} 1-q_1 & 0 & 0 & 0 \\ 0 & 1-x & -x & x \\ 0 & -x & 1-x & x \\ 0 & x & x & 1-x \end{pmatrix}. \qquad (4.5)$$

The eigenvalues of this matrix are $1-q_1$, $1-[3/(4q_1)]$, and 1, which is doubly degenerate. The matrix is positive if $(3/4) \leqslant q_1 \leqslant 1$.

We can find the vectors $|\psi'_j\rangle$ by noting that, because M is positive, it can be expressed as $M = B^\dagger B$, where $B = U_0 \sqrt{M}$, and U_0 is an arbitrary unitary operator. If we set

$$|\psi'_j\rangle = B_{1j}|v_{0,1}\rangle + B_{2j}|v_{1,1}\rangle + B_{3j}|v_{2,1}\rangle + B_{4j}|v_{2,2}\rangle, \qquad (4.6)$$

then we find that $M_{jk} = \langle \psi'_j | \psi'_k \rangle$, and we have completely specified U. Note that there is considerable arbitrariness, because we are free to choose the unitary operator U_0. A particular choice yields

$$|\psi'_1\rangle = \sqrt{1-q_1}|v_{0,1}\rangle,$$

$$|\psi'_2\rangle = \frac{1}{\sqrt{2}}|v_{1,1}\rangle + \frac{1}{\sqrt{6}}|v_{2,1}\rangle + \sqrt{1-\frac{x}{3}}|v_{2,2}\rangle,$$

$$|\psi'_3\rangle = -\frac{1}{\sqrt{2}}|v_{1,1}\rangle + \frac{1}{\sqrt{6}}|v_{2,1}\rangle + \sqrt{1-\frac{x}{3}}|v_{2,2}\rangle,$$

$$|\psi'_4\rangle = \sqrt{\frac{2}{3}}|v_{2,1}\rangle - \sqrt{1-\frac{x}{3}}|v_{2,2}\rangle. \qquad (4.7)$$

These vectors, plus the choice of q_1, which one finds by minimizing the failure probability and checking to see whether it is in the allowed range (in this case between 3/4 and 1), completely specify the POVM.

V. APPLICATION TO ALL BALANCED FUNCTIONS

Now that we know how this procedure performs on the basis vectors in $\mathcal{H}_b$, we shall examine its performance on any even function. What we shall do is compare the performance of the two von Neumann measurements to that of the POVM that is optimal for the basis vectors. We cannot apply the filtering procedure directly to the set of balanced functions, unless we want the dimension of the space in which the POVM is defined to be equal to the number of balanced functions. We are instead interested in a POVM acting in the space $\mathcal{H}_D \oplus \mathcal{H}_A$, where $\mathcal{H}_A$ is a one-dimensional auxiliary Hilbert space. This can be accomplished by using the POVM that was derived for the more restricted problem of distinguishing the basis vectors from $|w_k\rangle$.

First let us see how the von Neumann measurements perform. We shall assume that $|w_k\rangle$ has an *a priori* probability of η_1 and that each of the vectors corresponding to a balanced function has the same probability. There are

$$M_{bal} = \binom{D}{D/2} \tag{5.1}$$

vectors corresponding to balanced functions, so that the probability of each of them is $(1-\eta_1)/M_{bal}$. In order to calculate the average error probabilities we need to calculate

$$S_b = \frac{1-\eta_1}{M_{bal}} \sum_{v_b} |\langle w_k | v_b \rangle|^2, \tag{5.2}$$

where the sum is over all vectors corresponding to balanced functions. This sum is calculated in the Appendix, and we find that

$$S_b = \frac{1-\eta_1}{D-1} f_k, \tag{5.3}$$

where $f_k = (2^k - 1)/2^{2k-2}$. This expression is identical to the one calculated for the basis vectors alone. That means that the failure probabilities for the two von Neumann measurements, in the case of all balanced functions and not just the basis vectors, are still given by the expressions in Eq. (3.11).

Now let us turn our attention to the POVM. As discussed at the end of Sec. II, even for the POVM designed to distinguish the basis vectors from $|w_k\rangle$, the failure probability for any balanced function vector, $|v_b\rangle$, is given by

$$q_{v_b} = \frac{|\langle w_k | v_b \rangle|^2}{q_1}, \tag{5.4}$$

so that the total failure probability is

$$Q_{\mathrm{POVM}} = \eta_1 q_1 + \frac{S_b}{q_1}. \tag{5.5}$$

Choosing q_1 to minimize the right-hand side, we obtain the expression for Q_{POVM} given in Eq. (3.11). Again, we must have $\zeta_2 \leqslant \eta_1 \leqslant \zeta_1$ for this expression to be valid.

Now let us look at some choices for η_1. If all of the functions are equally probable, then $\eta_1 = 1/(M_{bal}+1)$, and the first von Neumann measurement is the optimal one. The measurement fails if the vector is $|w_k\rangle$, but this event is so unlikely that it has a negligible effect on the average failure probability. If $\eta_1 = 1/D$, that is, the probabilities of getting $|w_k\rangle$ or a balanced-function vector are proportional to the dimensions of the subspaces in which they lie, then the POVM is optimal. If $\eta_1 = 1/2$, so that we are equally likely to be given $|w_k\rangle$ or a balanced-function vector, then the second von Neumann measurement is optimal. This illustrates how the best strategy is influenced by the *a priori* probabilities.

VI. CONCLUSION

If one is given a Boolean function that is promised to be either even or in W_k, classically, in the worst case, one would have to evaluate it $2^n[(1/2)+(1/2^k)]+1$ times to determine to which set it belongs. Using quantum information processing methods, one has a very good chance of determining this with only one function evaluation. This shows that Deutsch-Jozsa-type algorithms need not be limited to constant functions; certain kinds of biased functions can be discriminated as well.

Unambiguous state discrimination is a procedure that is of fundamental interest in quantum information theory. Its only application so far has been to quantum cryptography. The results presented here suggest that related methods can also serve as a tool in the development of quantum algorithms.

ACKNOWLEDGMENTS

We would like to thank Ulrike Herzog for useful conversations. This research was supported by the Office of Naval Research (Grant No. N00014-92-J-1233), the National Science Foundation (Grant No. PHY-0139692), the Hungarian Science Research Fund (Grant No. T 03061), a PSC-CUNY grant, and a CUNY collaborative grant.

APPENDIX

We want to evaluate the sum in Eq. (5.2). The vector $|w_k\rangle$ is given by

$$|w_k\rangle = \frac{1}{\sqrt{D}}(1,1,\ldots 1,-1,-1,\ldots -1), \tag{6.1}$$

where the first $D-(D/2^k)$ places are 1's and the final $D/2^k$ places are -1's. Now consider a vector corresponding to a balanced function, which has m 1's in its last $D/2^k$ places, where $0 \leqslant m \leqslant D/2^k$, so that it has $(D/2)-m$ 1's in its first $D-(D/2^k)$ places. The overlap of $|w_k\rangle$ with this vector is

$$\frac{1}{2^{k-1}} - \frac{4m}{D}. \tag{6.2}$$

The number of balanced functions of this type, C_m, is given by

$$C_m = \binom{D/2^k}{m}\binom{D(2^k-1)/2^k}{(D/2)-m}, \tag{6.3}$$

so that the sum we have to evaluate is given by

$$S_b = \eta \sum_{m=0}^{D/2^k} C_m \left(\frac{1}{2^{k-1}} - \frac{4m}{D}\right)^2, \tag{6.4}$$

where $\eta = (1-\eta_1)/M_{bal}$ is the *a priori* probability of each balanced function.

In order to find S_b, we have to evaluate three types of sums. We shall discuss one is some detail, and simply give results for the other two. The first sum is

$$s_0 = \sum_{m=0}^{D/2^k} C_m. \tag{6.5}$$

This can be evaluated by noting that

$$(1+x)^{D/2^k} = \sum_{m=0}^{D/2^k} \binom{D/2^k}{m} x^m,$$

$$(1+x)^{D(2^k-1)/2^k} = \sum_{l=0}^{D(2^k-1)/2^k} \binom{D(2^k-1)/2^k}{(D/2)-m} x^l. \tag{6.6}$$

Multiplying these two expressions together we find

$$(1+x)^D = \sum_{m=0}^{D/2^k} \sum_{l=0}^{D(1-2^{-k})} \binom{D/2^k}{m} \binom{D(1-2^{-k})}{(D/2)-m} x^{l+m}. \tag{6.7}$$

Comparing coefficients of $x^{D/2}$ on both sides of this equation, we find that

$$s_0 = \binom{D}{D/2}. \tag{6.8}$$

The remaining two sums are

$$s_1 = \sum_{m=0}^{D/2^k} mC_m = \frac{D}{2^k} \binom{D-1}{(D/2)-1}, \tag{6.9}$$

and

$$s_2 = \sum_{m=0}^{D/2^k} m^2 C_m$$

$$= \frac{D}{2^k}\left(\frac{D}{2^k}-1\right)\binom{D-2}{(D/2)-2} + \frac{D}{2^k}\binom{D-1}{(D/2)-1}. \tag{6.10}$$

Substitution of these expressions into Eq. (6.4) yields the result in Eq. (5.3).

[1] D. Deutsch and R. Josza, Proc. R. Soc. London, Ser. A **439**, 553 (1992).

[2] Y. Sun, J. A. Bergou, and M. Hillery, Phys. Rev. A **66**, 032315 (2002).

[3] J. A. Bergou, U. Herzog, and M. Hillery, Phys. Rev. Lett. **90**, 257901 (2003).

[4] J. Bergou, U. Herzog, and M. Hillery, Phys. Rev. A **71**, 042314 (2005).

[5] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).

[6] A. S. Holevo, J. Multivariate Anal. **3**, 337 (1973).

[7] H. P. Yuen, R. S. Kennedy, and M. Lax, IEEE Trans. Inf. Theory **IT-21**, 125 (1975).

[8] T. Rudolph, R. W. Spekkens, and P. S. Turner, Phys. Rev. A **68**, 010301(R) (2003).

[9] Ph. Raynal, N. Lütkenhaus, and S. J. van Enk, Phys. Rev. A **68**, 022308 (2003).

[10] Y. C. Eldar, M. Stojnic, and B. Hassibi, Phys. Rev. A **69**, 062318 (2004).

[11] Yuan Feng, Runyao Duan, and Mingsheng Ying, Phys. Rev. A **70**, 012308 (2004).

[12] U. Herzog and J. Bergou, Unambiguous discrimination of mixed quantum states, quant-ph/0502117, Phys. Rev. A (to be published).

[13] Ph. Raynal and N. Lütkenhaus, quant-ph/0502165.

[14] J. Fiurašek and M. Ježek, Phys. Rev. A **67**, 012321 (2003).

[15] Y. C. Eldar, Phys. Rev. A **67**, 042309 (2003).

[16] K. Kraus, *States, Effects and Operations: Fundamental Notions of Quantum Theory* (Springer, Berlin, 1983).

[17] Sh. Zhang and M. Ying, Phys. Rev. A **65**, 062322 (2002).