

Optimum unambiguous discrimination between linearly independent nonorthogonal quantum states and its optical realization

Yuqing Sun,¹ Mark Hillery,¹ and János A. Bergou^{1,2}

¹*Department of Physics, Hunter College, City University of New York, 695 Park Avenue, New York, New York 10021*

²*Institute of Physics, Janus Pannonius University, H-7624 Pécs, Ifjúság útja 6, Hungary*

(Received 24 October 2001; published 13 July 2001)

Unambiguously distinguishing between nonorthogonal but linearly independent quantum states is a challenging problem in quantum information processing. In principle, the problem can be solved by mapping the set of nonorthogonal quantum states onto a set of orthogonal ones, which then can be distinguished without error. Such nonunitary transformations can be performed conditionally on quantum systems; a unitary transformation is carried out on a larger system of which the system of interest is a subsystem, a measurement is performed, and if the proper result is obtained the desired nonunitary transformation has been performed on the subsystem. We show how to construct generalized interferometers (multiports), which when combined with measurements on some of the output ports, implement nonunitary transformations of this type. The input states are single-photon states in which the photon is divided among several modes. A number of explicit examples of distinguishing among three nonorthogonal states are discussed, and the networks that optimally distinguish among these states are presented.

DOI: 10.1103/PhysRevA.64.022311

PACS number(s): 03.67.-a, 03.65.Ta, 42.50.-p

I. INTRODUCTION

The time evolution of a closed quantum system is unitary; hence scalar products (angles between two quantum states) are conserved. However, when measurements are made on the system, it is possible to perform prescribed nonunitary operations with a certain probability of success. In particular, it is possible to alter the value of scalar products and, hence, the angle between state vectors. Such operations can be used to unambiguously discriminate among nonorthogonal quantum states. A set of nonorthogonal states is mapped onto a set of orthogonal ones, and the orthogonal states can be distinguished without error. According to the quantum theory of measurement, such a nonunitary transformation will always have a certain probability of failure, which, for the discrimination of nonorthogonal states, corresponds to the probability that we obtain inconclusive answers. Our aim here is to find the optimum solution that minimizes the average probability of failure.

Considerable work has been done on this problem. The simplest case, distinguishing two nonorthogonal states, was first considered by Ivanovic [1], and then subsequently by Dieks [2] and Peres [3]. These authors found the optimal solution when the two states are being selected from an ensemble in which they are equally likely. The optimal solution for the situation in which the states have different weights was found by Jaeger and Shimony [4]. One can also consider what happens if the discrimination is not completely unambiguous, i.e., if it is possible for errors to occur, and this was done by Chefles and Barnett [5]. The case of three states was examined by Peres and Terno [6]. The general N -state problem has been studied by Chefles [7], by Chefles and Barnett [8], and by Duan and Guo [9]. Chefles and Barnett employed the positive operator valued measure formalism and specifically solved the case in which the probability of the procedure succeeding is the same for each of the states. Duan and

Guo considered general unitary transformations and measurements on a Hilbert space containing the states to be distinguished and an ancilla, which would allow one to discriminate among N states, and derived matrix inequalities that must be satisfied for the desired transformations to exist.

For the experimental realization of quantum information processing, one must choose a physical system to represent a qubit. Some possibilities that have been used are energy levels of ions, the orientation of a nuclear spin, and the presence or absence of a photon in a cavity [10–13]. Another possibility, the so-called dual-rail representation of a qubit, was proposed by Milburn [14], and later by Chuang and Yamamoto [15,16]. A photon is split between two modes which represent 0 and 1. When the elementary carriers of the information are more than two-dimensional objects (qutrits, . . . , qunits, in general, for n dimensions), one needs a more general representation. Here we will show that single-photon states can be used to represent general nonorthogonal states in n dimensions, and how this representation can be used for state discrimination. A photon is now divided among n modes that represent the numbers $0, 1, \dots, n-1$. The method is a straightforward generalization of the dual-rail representation of a qubit for more than two dimensions, and can be called the multiple-rail representation of a qunit. An optical multiport, a kind of generalized interferometer with more than two inputs and outputs, together with measurements made by the photon detectors placed at some or all of the output ports, can conditionally realize the desired nonunitary transformations of the initial, nonorthogonal single-photon states into orthogonal states. Our previous paper [17], which proposed an optical realization to optimally discriminate between two nonorthogonal states, is a special case of the method presented here.

Optical experiments to distinguish between two quantum states have already been carried out, first by Huttner *et al.* [18] and, more recently, by Clarke *et al.* [19]. Both of these used the polarization states of photons to represent qubits.

This paper is divided into six sections. In Sec. II, we present a method for calculating the optimum probabilities of unambiguous discrimination between linearly independent, nonorthogonal states. In Sec. III, the general properties of a quantum system that realizes the optimum nonunitary transformation are found by assuming that the optimum probabilities are known. In Sec. IV, we show how an optical multiport, which is designed to perform a particular unitary transformation, together with measurements at its output ports can realize nonunitary transformations of nonorthogonal input states represented by single-photon states. Reck *et al.* [20] gave a method to decompose any multiport into a series of beam splitters, phase shifters, and mirrors, and we will use this method to construct the desired multiport. The general method is then illustrated by applying it to qutrits in Sec. V. Examples are presented for the realization of transformations that convert three specific nonorthogonal states to orthogonal ones with a maximum probability of success. A brief discussion and conclusions are given in Sec. VI.

II. OPTIMAL PROBABILITIES FOR UNAMBIGUOUS DISCRIMINATION AMONG NONORTHOGONAL QUANTUM STATES

Suppose we are given a quantum system prepared in the state $|\psi\rangle$, which is guaranteed to be a member of the set of nonorthogonal states $\{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_n\rangle\}$, but we do not know which one. We want to find a procedure that will tell us which member of the set we were given. The procedure may fail to give us any information about the state, and if it fails it must let us know that it has, but if it succeeds it should never give us a wrong answer. We shall refer to such a procedure as state discrimination without error [21,22]. Note that this procedure has $n+1$ outcomes; it either tells us which state we were given, or it tells us that it failed (inconclusive outcome).

In order to achieve error-free discrimination, Chefles has shown, in a very clear analysis of the problem, that the set $|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_n\rangle$ must be linearly independent [7]. If the states are not orthogonal (which we shall assume), they cannot be discriminated perfectly. That means that if we are given $|\psi_i\rangle$ we will have some probability p_i of distinguishing it successfully and, correspondingly, some failure probability $q_i = 1 - p_i$ of obtaining an inconclusive answer. Denote by \mathcal{H} the Hilbert space spanned by the initial states $\{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_n\rangle\}$. Since there is a chance to get an inconclusive answer, the number of outcomes of this process is larger than the dimension of \mathcal{H} ; hence this process is a ‘‘generalized measurement’’ which can be represented by a set of operators that form a resolution of the identity [7]

$$\hat{A}_I^\dagger \hat{A}_I + \sum_i \hat{A}_i^\dagger \hat{A}_i = \hat{\mathbf{1}}, \quad (2.1)$$

where \hat{A}_i is the operator that corresponds to the outcome $|\psi_i\rangle$, and \hat{A}_I is the operator that corresponds to the inconclusive outcome. In more detail, if ρ is the density matrix of our given state, then the probability of obtaining the k th outcome, where k can be $1, \dots, n$ or I , is $p_k = \text{Tr}(\rho \hat{A}_k^\dagger \hat{A}_k)$ and if

the outcome is k then the resulting density matrix is $\hat{\rho}_k = \hat{A}_k^\dagger \rho \hat{A}_k / p_k$. The requirement that the discrimination be error-free implies that

$$\langle \psi_i | \hat{A}_k^\dagger \hat{A}_k | \psi_i \rangle = p_i \delta_{ik}. \quad (2.2)$$

From this, by an application of the Schwarz inequality, it follows that

$$\langle \psi_k | \hat{A}_i^\dagger \hat{A}_i | \psi_j \rangle = p_i \delta_{ji} \delta_{jk}. \quad (2.3)$$

If we denote by η_i the *a priori* probability that the system was prepared in the state $|\psi_i\rangle$, the average probabilities of success and of failure to distinguish the states $|\psi_i\rangle$ are, respectively,

$$P = \sum_i \eta_i p_i, \quad (2.4)$$

$$Q = \sum_i \eta_i q_i.$$

Our objective is to find the set of $\{p_i\}$ that maximizes the probability of success P , and the set of operators \hat{A}_k that realize the corresponding generalized measurement.

Define the failure state $|\phi_i\rangle$ as

$$|\phi_i\rangle = \hat{A}_I |\psi_i\rangle. \quad (2.5)$$

This is the state of the system if the input state was $|\psi_i\rangle$ and the outcome was inconclusive. Chefles [7] showed that the states $\{|\phi_i\rangle\}$ are linearly dependent when P is a maximum. The interpretation of this result is the following. Because only linearly independent states can be discriminated without error, the operator corresponding to the inconclusive outcome maps the set of linearly independent states $\{|\psi_i\rangle\}$ onto a linearly dependent set, which then cannot be unambiguously discriminated by any further process. As we shall see, however, this does not mean that some information cannot be extracted from an inconclusive result.

Now consider the inner product $\langle \phi_k | \phi_j \rangle$, and define the matrix C by $C_{kj} = \langle \phi_k | \phi_j \rangle$. Using Eqs. (2.1), (2.5), and (2.3), we find

$$\langle \phi_k | \phi_j \rangle = \langle \psi_k | \psi_j \rangle - p_j \delta_{jk}. \quad (2.6)$$

The matrix C is positive semidefinite. This can be seen by noting that for any n -dimensional vector, whose components we shall denote by b_i , where $i = 1, \dots, n$,

$$\sum_{i,j=1}^n b_i^* C_{ij} b_j = \left\| \sum_{i=1}^n b_i |\phi_i\rangle \right\|^2 \geq 0. \quad (2.7)$$

When the p_i are equal to their optimal values, i.e., the values that maximize P , the linear dependence of the $|\phi_i\rangle$ implies that

$$\det(C) = 0, \quad (2.8)$$

so that C has at least one zero eigenvalue when P is a maximum.

In conclusion, the optimum probabilities p_i can be found by maximizing P subject to the following constraints: (i) $\det(C)=0$; (ii) C is non-negative, or, equivalently, all of the principal minors of C are non-negative [9]. In both of these conditions, the matrix C is considered to be a function of the probabilities p_i , $i=1, \dots, n$, with matrix elements given explicitly by $C_{kj}=\langle\psi_k|\psi_j\rangle-p_j\delta_{jk}$.

When we consider the case of just two nonorthogonal states, the above result immediately gives the following relationship between any two failure probabilities:

$$q_1q_2=|\langle\psi_1|\psi_2\rangle|^2. \quad (2.9)$$

This is the same result that was obtained in our previous paper [19]. In particular, when the two states have equal *a priori* probabilities, $\eta_1=\eta_2=\frac{1}{2}$, we found the maximum probability of success to be

$$P=1-|\langle\psi_1|\psi_2\rangle|. \quad (2.10)$$

This is the well known Ivanovic-Dieks-Peres limit [1–3].

III. REALIZATION OF GENERALIZED MEASUREMENT

Once we know the set of optimum discrimination probabilities $\{p_i\}$, we would like to find a realizable experimental procedure to achieve it. We shall do this first abstractly, and then show how it can be realized by linear optical elements. Let us first summarize the procedure, and subsequently fill in the details. We begin with a total Hilbert space \mathcal{K} , which is the direct sum of two subspaces, $\mathcal{K}=\mathcal{H}\oplus\mathcal{A}$. The space \mathcal{H} is an n -dimensional space that contains the vectors $|\psi_i\rangle$, and \mathcal{A} is the space that will contain the failure vectors $|\phi_i\rangle$. We shall denote the dimension of \mathcal{A} by m . The input state of the system is one of the vectors $|\psi_i\rangle$, which is now a vector in the subspace \mathcal{H} of the total space \mathcal{K} . A unitary transformation U that acts in the entire space \mathcal{K} is now applied to the input vector, resulting in the state $|\psi_i^{\mathcal{K}}\rangle_{out}$. A measurement is performed on the part of $|\psi_i^{\mathcal{K}}\rangle_{out}$ in \mathcal{A} , and, if the proper result is obtained, the vector $|\psi_i^{\mathcal{K}}\rangle_{out}$ is projected onto the vector $|e_i^{\mathcal{H}}\rangle$, which lies in the subspace \mathcal{H} . The probability of this occurring is p_i . The vectors $\{|e_i^{\mathcal{H}}\rangle, i=1, \dots, n\}$ are orthonormal and can be distinguished perfectly. The effect of the unitary transformation on an extended space and the measurement is to map a set of nonorthogonal vectors onto a set of orthogonal ones.

We now need to specify U and the measurement, and let us discuss the latter first. The measurement has two outcomes, one of them corresponding to the operator $P_{\mathcal{H}}$ that projects onto the subspace \mathcal{H} , and the other to the operator $P_{\mathcal{A}}=I-P_{\mathcal{H}}$ that projects onto the subspace \mathcal{A} . The first outcome corresponds to the successful transformation of $|\psi_i^{\mathcal{K}}\rangle_{out}$ into $|e_i^{\mathcal{H}}\rangle$, and its probability of occurrence is p_i . This implies that

$$|\psi_i^{\mathcal{K}}\rangle_{out}=\sqrt{p_i}|e_i^{\mathcal{H}}\rangle+|\phi_i^{\mathcal{A}}\rangle, \quad (3.1)$$

where $|\phi_i^{\mathcal{A}}\rangle$ is a failure state, and we have added a superscript \mathcal{A} to denote the fact that it is in the subspace \mathcal{A} . The other outcome corresponds to obtaining an inconclusive answer and transforms $|\psi_i^{\mathcal{K}}\rangle_{out}$ into $|\phi_i^{\mathcal{A}}\rangle$.

Equation (3.1) and the fact that U is unitary imply that

$$\langle\phi_k^{\mathcal{A}}|\phi_j^{\mathcal{A}}\rangle=\langle\psi_k^{\mathcal{H}}|\psi_j^{\mathcal{H}}\rangle-p_j\delta_{jk}, \quad (3.2)$$

which is just Eq. (2.6). If (and only if) the matrix C is non-negative, we can always find vectors $|\phi_i^{\mathcal{A}}\rangle$ that satisfy this equation. This follows from the fact that a non-negative matrix can be written as the product of a matrix and its adjoint; in particular, we can express C as

$$C=A^\dagger A, \quad (3.3)$$

for some matrix A . If we define $|\phi_j\rangle=A|j\rangle$, where $|j\rangle$ is the vector whose j th component is 1 and all of whose other components are zero, then we have that $C_{jk}=\langle\phi_j|\phi_k\rangle$. Once we have found these vectors and specified the vectors $|e_i^{\mathcal{H}}\rangle$, then the operator U can be found by means of Eq. (3.1). These conditions may not completely determine U ; if they do not, then there is freedom in choosing it. This will be the case if the dimension of \mathcal{A} is greater than 1. U maps vectors in \mathcal{A} to vectors in the subspace \mathcal{S} of \mathcal{K} that consists of the vectors that are orthogonal to all of the vectors $|\psi_i^{\mathcal{K}}\rangle_{out}$. The dimension of \mathcal{S} is m . The freedom in choosing U comes from the fact that Eq. (3.1) does not specify how \mathcal{A} is mapped into \mathcal{S} . If both are one dimensional, then the mapping is determined (up to an overall phase), but if their dimension is greater than 2 it is not.

Once the measurement and the operator U have been specified, our realization of the generalized measurement is completely determined. The next task is to find a physical system with which to implement it.

IV. OPTICAL REALIZATION OF NONUNITARY TRANSFORMATION

We now want to propose an experimental procedure to achieve our nonunitary transformation by using optical devices. We shall show how this can be accomplished by using a single-photon representation of the states $|\psi_i\rangle$ and an optical multipoint together with photodetectors at the output ports to carry out the desired nonunitary transformation.

Our Hilbert space will consist of a single photon, which is divided among $n+m$ modes. The modes themselves could be distinguished by having different wave vectors or they might be modes of different optical fibers. A basis for this space consists of the single-photon states $\{a_j^\dagger|0\rangle|j=1, \dots, n+m\}$, where $|0\rangle$ is the vacuum state and a_j^\dagger is the creation operator for the j th mode. The states $\{a_j^\dagger|0\rangle|j=1, \dots, n\}$ form a basis for the space \mathcal{H} , and the states $\{a_j^\dagger|0\rangle|j=n+1, \dots, n+m\}$ form a basis for the space \mathcal{A} . The initial states $|\psi_i\rangle$ can be represented as single-photon states in \mathcal{H} , which can be written as

$$|\psi_i\rangle=\sum_{j=1}^n d_{ij}|e_j^{\mathcal{H}}\rangle=\sum_{j=1}^n d_{ij}\hat{a}_j^\dagger|0\rangle, \quad (4.1)$$

where we have chosen the states $|e_j^{\mathcal{H}}\rangle$ to be $|e_j^{\mathcal{H}}\rangle = \hat{a}_j^\dagger |0^{\mathcal{H}}\rangle$.

An optical $2N$ -port is a lossless linear device with N input ports and N output ports. Its action on the input states can be described by a unitary operator U_{2N} , and physically it consists of an arrangement of beam splitters, phase shifters, and mirrors. Choosing $N = n + m$, we send the single-photon state $|\psi_i\rangle$ into the first n input ports, which correspond to \mathcal{H} , and the vacuum into the remaining m input ports, which correspond to \mathcal{A} . Photodetectors are placed at the last m output ports (the ones corresponding to \mathcal{A}), and if there is no photon detected the desired nonunitary transformation will have been carried out. In particular, with $|\psi_i^{\mathcal{K}}\rangle_{out} = U_{2N}|\psi_i\rangle$, where $|\psi_i^{\mathcal{K}}\rangle_{out}$ is given by Eq. (3.1), the action of the measurement, if successful, is to project the output state onto $|e_i^{\mathcal{H}}\rangle$, and the probability of achieving this is p_i .

This procedure is conceptually clear, but in practice there are problems in requiring that a photon not be detected. For example, the photon may be incident on the detector but not be registered due to finite detection efficiency. Because we are dealing with only one photon, however, it is possible to circumvent this problem. Suppose now that our $2N$ -port has detectors placed at its first n output ports, the ones corresponding to \mathcal{H} . These detectors are necessary in order to tell us which output state in \mathcal{H} we have, and this, in turn, tells what the input state was, if the procedure is successful. Now, because there is only one photon, if it emerged from one of the first n output ports, it did not emerge from one of the last m ports. That means we do not need detectors at the last m ports. If the photon emerges from one of the first n ports and is detected, then the procedure has been successful.

If we denote the annihilation operators corresponding to the input modes of the $2N$ -port by $a_{j_{in}}$, $j = 1, 2, \dots, N$, then the output operators are given by

$$a_{j_{out}} = U_{2N}^{-1} a_{j_{in}} U_{2N} = \sum_{k=1}^N M_{jk} a_{k_{in}}, \quad (4.2)$$

where M_{jk} are the elements of an $N \times N$ unitary matrix M . In the Schrödinger picture, the ‘‘in’’ and ‘‘out’’ states are related by

$$|\psi^{\mathcal{K}}\rangle_{out} = U_{2N} |\psi^{\mathcal{K}}\rangle_{in}. \quad (4.3)$$

In general, for an ‘‘in’’ state that contains a single-photon

$$|\psi^{\mathcal{K}}\rangle_{in} = \sum_{j=1}^N c_j a_j^\dagger |0\rangle, \quad (4.4)$$

where $\sum_{j=1}^N |c_j|^2 = 1$, the ‘‘out’’ state is given by

$$\begin{aligned} |\psi^{\mathcal{K}}\rangle_{out} &= U_{2N} |\psi^{\mathcal{K}}\rangle_{in} \\ &= U_{2N} \sum_{j=1}^N c_j a_j^\dagger U_{2N}^{-1} |0\rangle \\ &= \sum_{j,k=1}^N c_j M_{jk}^T a_{k_{in}}^\dagger |0\rangle. \end{aligned} \quad (4.5)$$

Note that we have made use of the fact that the vacuum is invariant under the transformation U_{2N} . This implies that the matrix element M_{il} is the same as the matrix element of U_{2N} between the single-particle states $|i\rangle = a_{i_{in}}^\dagger |0\rangle$ and $|l\rangle = a_{l_{in}}^\dagger |0\rangle$. Choosing $c_j = \delta_{jl}$ in the above equation and then taking the inner product of the result with $|i\rangle$, we find that

$$\langle i | U_{2N} | l \rangle = M_{il}. \quad (4.6)$$

The desired matrix M can be found from Eq. (3.1), and our next task is to decompose it in such a way that it corresponds to a linear optical network.

This problem has been solved by Reck *et al.* [20], and we shall summarize their method. They gave an algorithmic procedure to factorize any $N \times N$ unitary matrix into a product of two-dimensional $U(2)$ transformations, and it is this procedure that we shall adopt here to construct our $2N$ -port, which is characterized by the matrix M of Eq. (4.2).

It is well known that a lossless beam splitter and a phase shifter with appropriate parameters can implement any $U(2)$ transformation; a beam splitter with a phase shifter at one output port transforms the input operators into output operators as

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}_{out} = \begin{pmatrix} e^{i\phi} \sin \omega & e^{i\phi} \cos \omega \\ \cos \omega & -\sin \omega \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}_{in} \quad (4.7)$$

where a_1 and a_2 are the annihilation operators of modes 1 and 2 respectively. In their paper, Reck *et al.* considered the use of a Mach-Zehnder interferometer to simulate the effect of a beam splitter that does not split the incoming beam equally, in which case ω describes the reflectivity and transmittance of the effective beam splitter with $\sqrt{R} = \sin \omega$, $\sqrt{T} = \cos \omega$, and ϕ describes the effect of the phase shifter. If the matrix describes an actual beam splitter, then $\sqrt{R} = \cos \omega$ and $\sqrt{T} = \sin \omega$. Any $N \times N$ unitary matrix $U(N)$ can be reduced to an $(N-1) \times (N-1)$ unitary matrix, $U(N-1)$, by multiplying from the right by a succession of two-dimensional unitary matrices,

$$U(N) \cdot R(1) = \begin{pmatrix} e^{i\alpha_1} & 0 \\ 0 & U(N-1) \end{pmatrix}. \quad (4.8)$$

Here $R(1) = T_{1,2} \cdot T_{1,3} \cdots T_{1,N}$, and $T_{p,q}$ is defined as an N -dimensional identity matrix with elements I_{pp} , I_{pq} , I_{qp} , I_{qq} replaced by the corresponding elements of a $U(2)$ matrix. It performs a unitary transformation on a two-dimensional subspace of the full N -dimensional space, and can be implemented by attaching a beam splitter and a phase shifter to ports p and q .

We can repeat the above transformation, decreasing the dimension of the remaining unitary matrix by 1 at each step. Applying this procedure to the matrix M of Eq. (4.2), we have that

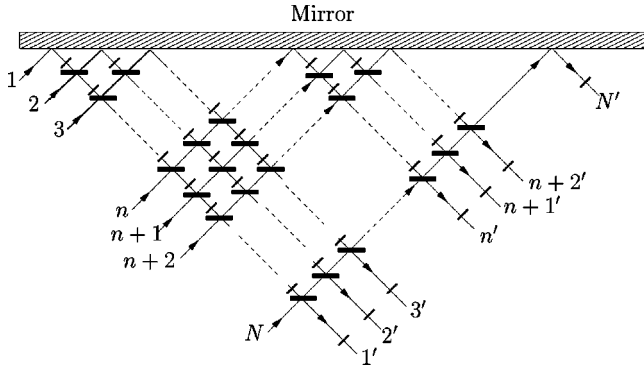


FIG. 1. The implementation of an optical multiport that performs the unitary transformation $M(N)$ described in Eq. (4.12). The beams are straight lines, a suitable beam splitter is at each crossing point of the beams, phase shifters are at one input of each beam splitter and at the outputs of the multiport. Each diagonal line of the multiport reduces the dimension of $M(N)$ by 1.

$$M \cdot R(1) \cdot R(2) \cdots R(n+m) = \begin{pmatrix} e^{i\alpha_1} & 0 & & 0 \\ 0 & e^{i\alpha_2} & & \\ 0 & 0 & \ddots & \\ 0 & 0 & & e^{i\alpha_{n+m}} \end{pmatrix}. \quad (4.9)$$

Denoting by $D(\alpha_1, \alpha_2, \dots, \alpha_n)$ the diagonal matrix

$$D = \begin{pmatrix} e^{-i\alpha_1} & & & 0 \\ & e^{-i\alpha_2} & & \\ & & \ddots & \\ 0 & & & e^{-i\alpha_{n+m}} \end{pmatrix}, \quad (4.10)$$

we have

$$M \cdot R(1) \cdot R(2) \cdots R(n+m-1) \cdot D = \mathbf{1}, \quad (4.11)$$

i.e.,

$$M = D^{-1} \cdot R(n+m-1)^{-1} \cdots R(1)^{-1}. \quad (4.12)$$

Since the product of matrices is equivalent to setting up experimental devices in sequence, Eq. (4.12) implies that to get M the actual experimental setup is made of a series of $U(2)$ blocks to achieve $R(n+m)^{-1} \cdots R(1)^{-1}$, and $n+m$ appropriate phase shifters attached to the output ports to produce D^{-1} . Figure 1 gives a picture of the practical implementation of M .

It is possible to save some steps by modifying this procedure. As was mentioned earlier, the matrix M is not always completely determined. In particular, there is freedom in choosing the matrix elements M_{jk} for $k > n$. Let us now see what happens if we apply the procedure of Reck *et al.* to the transpose of M , M^T , instead of M itself. It is now the matrix elements $(M^T)_{jk}$ for $j > n$ that are not completely determined, and we shall leave them that way for now. The matrices making up $R(1)$ are chosen to make all of the elements, except the first, of the first row of $M^T R(1)$ zero. In finding each of the matrices $T_{1,q}$, we only need to use the

matrix elements that are completely determined (this is not true if we start with M instead of M^T). Now if the first row of $M^T R(1)$ is zero except for the first element, then by unitarity the first column is also zero, except for its first element. Continuing in this way we have that

$$M^T \cdot R(1) \cdot R(2) \cdots R(n) = \begin{pmatrix} e^{i\alpha_1} & 0 & & 0 \\ 0 & e^{i\alpha_2} & & \\ 0 & 0 & \ddots & \\ 0 & 0 & & M_m \end{pmatrix}, \quad (4.13)$$

where M_m is an $m \times m$ unitary matrix that contains the information about the matrix elements in M that are not completely specified. At this point, we can choose M_m to be any unitary matrix, and the simplest choice is the $m \times m$ identity matrix, I_m . Defining

$$D' = \begin{pmatrix} e^{-i\alpha_1} & 0 & & 0 \\ 0 & e^{-i\alpha_2} & & \\ 0 & 0 & \ddots & \\ 0 & 0 & & I_m \end{pmatrix}, \quad (4.14)$$

we have that

$$M = [R(1)R(2) \cdots R(n)D']^*. \quad (4.15)$$

V. APPLICATION TO THREE STATES AND EXAMPLES

In this section we first apply the above considerations to the problem of realizing optimal discrimination among three nonorthogonal but linearly independent quantum states, in general. Then we illustrate the method on specific examples. For simplicity, we assume that the *a priori* probabilities are all equal, $\eta_1 = \eta_2 = \eta_3 = 1/3$.

From Eq. (2.4), the probability of failure is

$$Q = \frac{1}{3} \sum_{i=1}^3 q_i. \quad (5.1)$$

The requirement of the linear dependence of the $|\phi_i\rangle$ vectors ($i=1,2,3$) leads to the constraint given by Eq. (2.8). For the case of three vectors it can be written as

$$\begin{aligned} \Delta &= \det(C) \\ &= q_1 q_2 q_3 - q_1 |O_{23}|^2 - q_2 |O_{13}|^2 - q_3 |O_{12}|^2 \\ &\quad + O_{12} O_{23} O_{13}^* + O_{12}^* O_{23}^* O_{13} \\ &= 0, \end{aligned} \quad (5.2)$$

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

Employing the Lagrange multiplier method, we wish to minimize the quantity

$$Q' = \frac{1}{3} \sum_i q_i + \lambda \Delta, \quad (5.3)$$

which immediately leads to the conditions

$$\begin{aligned}\frac{\partial Q'}{\partial q_1} &= \frac{1}{3} + \lambda \Delta_{23} = 0, \\ \frac{\partial Q'}{\partial q_2} &= \frac{1}{3} + \lambda \Delta_{13} = 0, \\ \frac{\partial Q'}{\partial q_3} &= \frac{1}{3} + \lambda \Delta_{12} = 0,\end{aligned}\quad (5.4)$$

where λ is a Lagrange multiplier, and Δ_{12} , Δ_{13} , Δ_{23} are subdeterminants of C , $\Delta_{12} = q_1 q_2 - |O_{12}|^2$, etc. Equation (5.4) implies that

$$\Delta_{12} = \Delta_{13} = \Delta_{23} = -\frac{1}{3\lambda}. \quad (5.5)$$

This means that all three subdeterminants are equal. Let $\delta = -1/3\lambda$ denote this common value and recall that all subdeterminants of C must be non-negative, so that $\delta \geq 0$.

From Eq. (5.5) we can solve for the q_i 's, yielding

$$\begin{aligned}q_1 &= \sqrt{\frac{(|O_{12}|^2 + \delta)(|O_{13}|^2 + \delta)}{(|O_{23}|^2 + \delta)}}, \\ q_2 &= \sqrt{\frac{(|O_{12}|^2 + \delta)(|O_{23}|^2 + \delta)}{(|O_{13}|^2 + \delta)}}, \\ q_3 &= \sqrt{\frac{(|O_{13}|^2 + \delta)(|O_{23}|^2 + \delta)}{(|O_{12}|^2 + \delta)}}.\end{aligned}\quad (5.6)$$

Finally, we can substitute Eq. (5.6) into Eq. (5.2) to solve for δ and then use the above equations to find the corresponding q_i values. When we solve for δ , there are often a number of different solutions. However, we need only consider solutions that are greater than or equal to zero, and which give values of q_i that are between 0 and 1. If there are several solutions that satisfy these conditions, we must determine which one gives the actual minimum. If there are none, then we must examine the boundary of the allowed region to find the minimum. The point (q_1, q_2, q_3) lies inside or on the surface of a unit cube whose vertices lie on the points (j, k, l) , where $j, k, l = 0$ or 1. If the Lagrange multiplier approach does not yield a valid solution the minimum of Q subject to the constraint $\Delta = 0$ must lie on the surface of the cube.

Note that if the overlaps are real and positive, a situation we shall consider shortly, then $\delta = 0$ is always a solution of Eq. (5.2). In this case, if all the corresponding q_i for $\delta = 0$ are between 0 and 1, then this set of $\{q_i\}$ is a possible solution to our problem, i.e., a minimum of Q that satisfies $\Delta = 0$. If it is, in fact, the solution, we see that $\Delta_{12} = \Delta_{13} = \Delta_{23} = 0$, which implies that each possible pair of the states $|\phi_i\rangle$, $i = 1, 2, 3$, is linearly dependent, so that all three states ϕ_i are in a line, i.e., the dimensionality of the auxiliary Hilbert space \mathcal{A} is 1. If the solution to the problem is one for

which $\delta > 0$, no pair of failure states is linearly dependent. However, the three failure states together are linearly dependent, so that in this case the dimensionality of the auxiliary Hilbert space \mathcal{A} is 2.

Next, we shall consider specific examples involving three nonorthogonal but linearly independent state vectors, to illustrate the general considerations of the previous sections. In particular we want to determine explicitly the parameters and dimensionality for the special multiports that optimally discriminate among the three quantum states. For simplicity, we shall assume that the *a priori* probabilities are equal in all of our examples.

Our first case is a simple one; the overlaps of the three states will be assumed to be real and equal

$$\langle \psi_1 | \psi_2 \rangle = \langle \psi_2 | \psi_3 \rangle = \langle \psi_3 | \psi_1 \rangle = s, \quad (5.7)$$

where $0 < s < 1$. The constraint of Eq. (2.8) is, in this case,

$$q_1 q_2 q_3 - s^2 \sum_i q_i + 2s^3 = 0, \quad (5.8)$$

and application of the the Lagrangian multiplier method implies that $q_1 = q_2 = q_3$, and that

$$q_i^3 - 3s^2 q_i + 2s^3 = 0. \quad (5.9)$$

This equation has two solutions, $q_i = s, -2s$, of which only $q_i = s$ is valid. This solution is a minimum and it implies that the optimal value of the total failure probability is $Q = s$.

Our next step is to find the failure vectors. For any 3×3 positive matrix L , we find that we can express its matrix elements as $L_{ij} = \langle \phi_i | \phi_j \rangle$ if

$$\begin{aligned}|\phi_1\rangle &= (\sqrt{L_{11}}, 0, 0), \\ |\phi_2\rangle &= \left(\frac{L_{12}}{\sqrt{L_{11}}}, \sqrt{\frac{\Delta_{12}}{L_{11}}}, 0 \right), \\ |\phi_3\rangle &= \left(\frac{L_{13}}{\sqrt{L_{11}}}, \frac{L_{23}L_{11} - L_{12}^* L_{13}}{\sqrt{L_{11}\Delta_{12}}}, \sqrt{\frac{\Delta}{\Delta_{12}}} \right),\end{aligned}\quad (5.10)$$

where $\Delta_{12} = L_{11}L_{22} - |L_{12}|^2$ and $\Delta = \det L$. Applying this to the matrix C , with $q_i = s$, $i = 1, 2, 3$, we find that the three failure vectors are identical; they all have magnitude \sqrt{s} and point in the same direction. Therefore, our failure space \mathcal{A} is one dimensional, the full Hilbert space $\mathcal{K} = \mathcal{H} \oplus \mathcal{A}$ is four dimensional, and we will need an eight-port to accomplish our unitary transformation.

In order to find the necessary unitary transformation, we must first specify our input states. Let us choose our three states to be (in the full space $\mathcal{K} = \mathcal{H} \oplus \mathcal{A}$)

TABLE I. Arguments of $T_{p,q}$ for case 1. The range of arccos is the interval $[0, \pi]$.

	ω	ϕ
$T_{1,2}$	$\arccos[-1/\sqrt{5}]$	0
$T_{1,3}$	$\arccos[-1/\sqrt{6}]$	0
$T_{2,3}$	$-\arccos[-\sqrt{2/5}]$	0
$T_{2,4}$	$\arccos[\sqrt{3s}/\sqrt{1+2s}]$	0

$$|\psi_1^{\mathcal{K}}\rangle_{in} = \begin{pmatrix} \sqrt{\frac{2}{3}}\sqrt{1-s} \\ \sqrt{1+2s}/\sqrt{3} \\ 0 \\ 0 \end{pmatrix}, \quad |\psi_2^{\mathcal{K}}\rangle_{in} = \begin{pmatrix} -\sqrt{1-s}/\sqrt{6} \\ \sqrt{1+2s}/\sqrt{3} \\ \sqrt{1-s}/\sqrt{2} \\ 0 \end{pmatrix},$$

$$|\psi_3^{\mathcal{K}}\rangle_{in} = \begin{pmatrix} -\sqrt{1-s}/\sqrt{6} \\ \sqrt{1+2s}/\sqrt{3} \\ -\sqrt{1-s}/\sqrt{2} \\ 0 \end{pmatrix}, \quad (5.11)$$

where ψ_i are represented by single-photon states. One can verify that $\langle\psi_1|\psi_2\rangle = \langle\psi_2|\psi_3\rangle = \langle\psi_3|\psi_1\rangle = s$. The output states can be found from Eq. (3.1), and are explicitly given by

$$|\psi_1^{\mathcal{K}}\rangle_{out} = \begin{pmatrix} \sqrt{1-s} \\ 0 \\ 0 \\ \sqrt{s} \end{pmatrix}, \quad |\psi_2^{\mathcal{K}}\rangle_{out} = \begin{pmatrix} 0 \\ \sqrt{1-s} \\ 0 \\ \sqrt{s} \end{pmatrix}, \quad (5.12)$$

$$|\psi_3^{\mathcal{K}}\rangle_{out} = \begin{pmatrix} 0 \\ 0 \\ \sqrt{1-s} \\ \sqrt{s} \end{pmatrix}.$$

The unitary transformation U maps the input states onto the output states, i.e., $|\psi_i^{\mathcal{K}}\rangle_{out} = U|\psi_i^{\mathcal{K}}\rangle_{in}$, for $i=1,2,3$. In addition, it must map the vector that is orthogonal to the three input vectors onto the vector that is orthogonal to the three output vectors,

$$\frac{1}{\sqrt{2s+1}} \begin{pmatrix} \sqrt{s} \\ \sqrt{s} \\ \sqrt{s} \\ -\sqrt{1-s} \end{pmatrix} = U \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (5.13)$$

The action of U on these four vectors completely determines it, and we find that it is given by the matrix $M(4)$, which is

$$M(4) = \begin{pmatrix} \sqrt{\frac{2}{3}} & \sqrt{(1-s)/3(2s+1)} & 0 & \sqrt{s/(2s+1)} \\ -1/\sqrt{6} & \sqrt{(1-s)/3(2s+1)} & 1/\sqrt{2} & \sqrt{s/(2s+1)} \\ -1/\sqrt{6} & \sqrt{(1-s)/3(2s+1)} & -1/\sqrt{2} & \sqrt{s/(2s+1)} \\ 0 & \sqrt{3s/(2s+1)} & 0 & -\sqrt{(1-s)/(2s+1)} \end{pmatrix}. \quad (5.14)$$

Using the method described in Sec. IV, $M(4)$ can be factorized as

$$M(4) = T_{1,2} \cdot T_{1,3} \cdot T_{2,3} \cdot T_{2,4}, \quad (5.15)$$

where the parameters that determine the matrices $T_{p,q}$ are given in Table I (this example is referred to as case 1). Note that because these matrices are real the complex conjugate, which appears in Eq. (4.15), is unnecessary.

Now let us consider a more general case than the one we have been studying so far. We shall assume that two of the overlaps are the same and the third is different, in particular that

$$\langle\psi_1|\psi_2\rangle = \langle\psi_1|\psi_3\rangle = s_1, \quad (5.16)$$

$$\langle\psi_2|\psi_3\rangle = s_2,$$

where we shall assume, for simplicity, that s_1 and s_2 are real and between 0 and 1. For a fixed value of s_1 there is a restriction on how large s_2 can be. The largest the angle between ψ_2 and ψ_3 can be is twice the angle between ψ_1 and ψ_2 (this maximum is achieved when the vectors are coplanar). This implies that $s_2 \geq 2s_1^2 - 1$. Application of the Lagrange multiplier method to the minimization of Q' gives us $q_2 = q_3$ and

$$q_1 = \frac{q_2^2 + s_1^2 - s_2^2}{q_2}. \quad (5.17)$$

Substituting these results into the constraint equation and defining $y = q_2/s_2$ and $\beta = s_1/s_2$, we have

$$y^4 - (2 + \beta^2)y^2 + 2\beta^2y + 1 - \beta^2 = 0. \quad (5.18)$$

The roots of this equation are $y = 1, -1 \pm \beta$, and two of them, 1 and $\beta - 1$, yield valid solutions, the latter if $\beta \geq 1$. Substitution of these results into Q shows that if $\beta < 2$ then the solution $y = 1$ gives the minimum and if $\beta \geq 2$ then $y = \beta - 1$ gives the minimum. Summarizing, we find that if $\beta < 2$, the minimum value of Q is $[(s_1^2/s_2) + 2s_2]/3$ and (solution 1)

$$q_1 = \frac{s_1^2}{s_2}, \quad (5.19)$$

$$q_2 = q_3 = s_2,$$

and if $\beta \geq 2$, then the minimum value of Q is $2(2s_1 - s_2)/3$ and (solution 2)

$$q_1 = 2s_1, \quad (5.20)$$

$$q_2 = q_3 = s_1 - s_2.$$

Clearly, for these solutions to be valid, all of the probabilities have to be between 0 and 1.

The next step is to find the failure vectors. If solution 1 is the valid one, we find from Eq. (5.10) that the failure space is one dimensional, and if $|u_1^A\rangle$ is the normalized basis vector for this space, then

$$|\phi_1^A\rangle = \frac{s_1}{\sqrt{s_2}} |u_1^A\rangle, \quad (5.21)$$

$$|\phi_2^A\rangle = |\phi_3^A\rangle = \sqrt{s_2} |u_1^A\rangle.$$

If solution 2 is the valid one, then the failure space is two dimensional. If $|u_j^A\rangle$ where $j = 1, 2$ is an orthonormal basis for this space we find that

$$|\phi_1^A\rangle = \sqrt{2s_1} |u_1^A\rangle,$$

$$|\phi_2^A\rangle = \sqrt{\frac{s_1}{2}} |u_1^A\rangle + \sqrt{\frac{s_1}{2} - s_2} |u_2^A\rangle, \quad (5.22)$$

$$|\phi_3^A\rangle = \sqrt{\frac{s_1}{2}} |u_1^A\rangle - \sqrt{\frac{s_1}{2} - s_2} |u_2^A\rangle.$$

Let us look at an example of each solution. If we choose our three states to be $\psi_1 = (1, 0, 0)$, $\psi_2 = (1/\sqrt{3})(1, 1, 1)$, and $\psi_3 = (1/\sqrt{3})(1, 1, -1)$, we find that $s_1 = 1/\sqrt{3}$ and $s_2 = 1/3$, so that solution 1 is valid. The complete treatment of this case (case 2) is given in Table II. We see that we need an eight-port which can be built up by two U(2) blocks. Note that in order to achieve minimum failure probability Q , we need to choose q_1 to be 1, which means that we sacrifice the possibility of distinguishing state $|\psi_1\rangle$.

If we choose our states to be $\psi_1 = (1, 0, 0)$, $\psi_2 = \frac{1}{3}(1, 2, 2)$, and $\psi_3 = \frac{1}{3}(1, 2, -2)$, then we find that solution 2 is valid with $s_1 = 1/3$ and $s_2 = 1/9$. In this case (case 3) we need a ten-port, and the complete results are given in Table III. Note

TABLE II. Summary of parameters and arguments of $T_{p,q}$ for case 2.

Input states	ψ_1	ψ_2	ψ_3
	$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ -1/\sqrt{3} \\ 0 \end{pmatrix}$
Optimal failure probability		$q_1 = 1$ $q_2 = 1/3$ $q_3 = 1/3$	
Output states	ψ_1	ψ_2	ψ_3
	$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0 \\ \sqrt{2/3} \\ 0 \\ 1/\sqrt{3} \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ \sqrt{2/3} \\ 1/\sqrt{3} \end{pmatrix}$
M		$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$	
Factorization of M		$M = T_{1,4} \cdot T_{2,3}$	
Arguments of $T_{p,q}$		$T_{1,4}: \omega = 0, \phi = 0$ $T_{2,3}: \omega = \pi/4, \phi = 0$	

that, if the procedure fails, it is still possible to gain some information about the input state, because the failure space is two dimensional [6]. This is not possible if the failure space has only one dimension.

One possibility is to attach to the failure space outputs (outputs 4 and 5) a network that transforms states $|\phi_2^A\rangle$ and $|\phi_3^A\rangle$ into orthogonal states, which it will do only with a certain probability [19]. In particular, we can construct a network that implements the transformation

$$M(3) = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \sqrt{\frac{2}{3}} & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}, \quad (5.23)$$

where the inputs to the first two ports of this network (we shall call these ports A and B) are the outputs of ports 4 and 5 of the original network, and the input to the third port (port C) is the vacuum. This network has been designed so that, if no photon is detected emerging from output C , then the input state $|\phi_2^A\rangle$ will be transformed into a photon emerging from port A , and the input state $|\phi_3^A\rangle$ will be transformed into a photon emerging from port B . If the input state is $|\phi_1^A\rangle$ and no photon is detected at output C , the probabilities of a pho-

TABLE III. Summary of parameters and arguments of $T_{p,q}$ for case 3.

Input states	ψ_1	ψ_2	ψ_3	
	$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1/3 \\ 2/3 \\ 2/3 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 1/3 \\ 2/3 \\ -2/3 \\ 0 \\ 0 \end{pmatrix}$	
Optimal failure probability		$q_1 = 2/3$ $q_2 = 2/9$ $q_3 = 2/9$		
Output states	ψ_1	ψ_2	ψ_3	
	$\begin{pmatrix} 1/\sqrt{3} \\ 0 \\ 0 \\ \sqrt{2/3} \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ \sqrt{7}/3 \\ 0 \\ 1/\sqrt{6} \\ 1/3\sqrt{2} \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ \sqrt{7}/3 \\ 1/\sqrt{6} \\ -1/3\sqrt{2} \end{pmatrix}$	
M	$\begin{pmatrix} 1/\sqrt{3} & -1/2\sqrt{3} & 0 & -\sqrt{7}/12 & 0 \\ 0 & \sqrt{7}/4 & \sqrt{7}/4 & -1/4 & 1/4 \\ 0 & \sqrt{7}/4 & -\sqrt{7}/4 & -1/4 & -1/4 \\ \sqrt{2/3} & 1/2\sqrt{6} & 0 & \sqrt{7}/24 & 0 \\ 0 & 0 & 1/2\sqrt{2} & 0 & -\sqrt{7}/8 \end{pmatrix}$			
Factorization of M	$M = T_{1,4} \cdot T_{2,3} \cdot T_{2,4} \cdot T_{3,5}$			
Arguments of $T_{p,q}$	$T_{1,4} : \omega = \arccos[\sqrt{2/3}], \phi = 0$			
	$T_{2,3} : \omega = \frac{\pi}{4}, \phi = 0$			
	$T_{2,4} : \omega = \arccos[-1/2\sqrt{2}], \phi = 0$			
	$T_{3,5} : \arccos[1/2\sqrt{2}], \phi = 0$			

ton emerging from either port A or port B are the same. Note that, as was discussed earlier, we do not actually need to put a photodetector at output C ; if the photon is detected at either output A or B , that guarantees that it would not be detected at output C . Therefore, if the photon emerges from port A , we can conclude the input to the entire network was either ψ_1 or ψ_2 , and if it emerges from port B then the input was either ψ_1 or ψ_3 . Summarizing, if one of the detectors in ports 1 through 3 clicks, we know what the input state was. If the detector in either port A or B clicks, then we gain partial information about the input state; the number of possibilities has been reduced from three to two. If none of these detectors clicks, then the photon has emerged from output C , and we have gained no information about the input state. This happens with a probability of $1/9$ if the inputs were ψ_2 or ψ_3 and $4/9$ if the input was ψ_1 . The addition of the second network to the failure outputs of the first significantly improves the chances of gaining some information about the input state.

VI. CONCLUSIONS

We have shown that nonorthogonal quantum states, each realized as a photon split among several modes, can be conditionally distinguished by means of a linear optical network.

For three states we have given explicit networks, which give the maximum success probabilities for several sets of states. In addition, it was shown that the addition of a second network to the outputs corresponding to a failure of the initial network to distinguish the states can sometimes provide partial information about the input state.

We believe it should be possible to construct these networks in the laboratory. The networks themselves consist of beam splitters and photodetectors. Producing a single-photon is a problem, but this could perhaps be done by taking a photon pair from a parametric down-converter, detecting one of the photons in order to determine that the pair had indeed been produced, and using the remaining photon in the network.

ACKNOWLEDGMENTS

This research was supported by the Office of Naval Research (Grant No. N00014-92J-1233), by the National Science Foundation (Grant No. PHY-9970507), by the Hungarian Science Research Fund (OTKA, Grant No. T 030671), and by a grant from PSC-CUNY as well as by a CUNY collaborative grant. We would like to thank Gabriel Drobny for useful conversations.

- [1] I.D. Ivanovic, Phys. Lett. A **123**, 257 (1987).
- [2] D. Dieks, Phys. Lett. A **126**, 303 (1988).
- [3] A. Peres, Phys. Lett. A **128**, 19 (1988).
- [4] G. Jaeger and A. Shimony, Phys. Lett. A **197**, 83 (1995).
- [5] A. Chefles and S.M. Barnett, J. Mod. Opt. **45**, 1295 (1998).
- [6] A. Peres and D. Terno, J. Phys. A **31**, 7105 (1998).
- [7] A. Chefles, Phys. Lett. A **239**, 339 (1998).
- [8] A. Chefles and S.M. Barnett, Phys. Lett. A **250**, 223 (1998).
- [9] L-M. Duan and G-C. Guo, Phys. Lett. A **261**, 25 (1999).
- [10] J. Cirac and P. Zoller, Phys. Rev. Lett. **74**, 4094 (1994).
- [11] N. Gershenfeld and I. Chuang, Science **275**, 350 (1997).
- [12] Q. Turchette, C. Hood, W. Lange, H. Mabuichi, and H. Kimble, Phys. Rev. Lett. **75**, 4710 (1995).
- [13] P. Domokos, J. Raimond, M. Brune, and S. Haroche, Phys. Rev. A **52**, 3554 (1995).
- [14] G. Milburn, Phys. Rev. Lett. **62**, 2124 (1988).
- [15] Isaac L. Chuang and Yoshihisa Yamamoto, Phys. Rev. A **52**, 3489 (1995).
- [16] Isaac L. Chuang and Yoshihisa Yamamoto, Phys. Rev. Lett. **76**, 4281 (1996).
- [17] J.A. Bergou, M. Hillery, and Y. Sun, J. Mod. Opt. **47**, 487 (2000).
- [18] B. Huttner, A. Muller, J.D. Gautier, H. Zbinden, and N. Gisin, Phys. Rev. A **54**, 3783 (1996).
- [19] R.B.M. Clarke, A. Chefles, S.M. Barnett, and E. Riis, e-print quant-ph/0007063.
- [20] M. Reck, A. Zeilinger, H.J. Bernstein, and P. Bertani, Phys. Rev. Lett. **73**, 58 (1994).
- [21] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, Dordrecht, 1993), pp. 282–285.
- [22] E.B. Davies, IEEE Trans. Inf. Theory **IT-24**, 239 (1978).