Maximum Confidence Quantum Measurements

Sarah Croke,[1,2,*] Erika Andersson,[1] Stephen M. Barnett,[1] Claire R. Gilson,[2] and John Jeffers[1]
[1]*Department of Physics, University of Strathclyde, Glasgow G4 0NG, United Kingdom*
[2]*Department of Mathematics, University of Glasgow, Glasgow G12 8QW, United Kingdom*
(Received 16 November 2005; published 22 February 2006)

We consider the problem of discriminating between states of a specified set with maximum confidence. For a set of linearly independent states unambiguous discrimination is possible if we allow for the possibility of an inconclusive result. For linearly dependent sets an analogous measurement is one which allows us to be as confident as possible that when a given state is identified on the basis of the measurement result, it is indeed the correct state.

DOI: 10.1103/PhysRevLett.96.070401 PACS numbers: 03.65.Ta, 03.67.Hk

One of the problems in exploiting the capability of a quantum system for carrying information is the difficulty in extracting the information encoded in a quantum state. It is not possible simply to measure the state of a quantum system in a single shot measurement, as the state is not itself an observable. Thus, without some prior knowledge, the state cannot be determined with certainty and without error. In fact, this is the case unless the state is known to be one of a mutually orthogonal set. In quantum communications, however, the receiving party has to discriminate between a known set of states $\{\hat{\rho}_i\}$ with known prior probabilities p_i [1]. In general, the states will not be orthogonal so that perfect discrimination is not possible, and we have to settle for the best that can be done. This means optimizing a figure of merit, with the simplest being to minimize the probability of incorrectly identifying the state. Necessary and sufficient conditions that the operators describing this minimum error measurement must satisfy are known [2,3], but the optimal measurement itself is known only in certain special cases [3–8]. A second possibility, unambiguous discrimination, is possible between two nonorthogonal states if we are prepared to accept the possibility of an inconclusive result [9–11]. When the inconclusive result is not obtained, it is possible to identify the initial state with certainty. This strategy is optimized by minimizing the probability of obtaining an inconclusive result [12]. Unambiguous discrimination can be extended to higher dimensions [13], but it is applicable only to sets of linearly independent states [14]. Other figures of merit include the mutual information shared by the transmitting and receiving parties [15,16] and the fidelity between the state received and one transmitted on the basis of the measurement result [17,18]. Examples of optimal minimum error, mutual information, and unambiguous discrimination measurement strategies have been demonstrated in experiments on optical polarization [19–24].

For linearly dependent states the analogue of unambiguous discrimination would be a measurement which allows us to be as confident as possible that the state we infer from our measurement result is the correct one. We take this criterion as the basis of maximum confidence measurements. A related problem was considered by Kosut *et al.* [25], who posed the question "if the detector declares that a specific state is present, what is the probability of that state actually being present?" That work used a worst-case optimality criterion; i.e., it considered the measurement that maximizes the smallest possible value of this probability for a given set of states. Here we consider the construction of a measurement that achieves the maximum possible value of this probability for each state in a set. In order to do this we sometimes have to accept the possibility of an inconclusive outcome, just as we need to for unambiguous discrimination.

Any measurement can be described mathematically by a probability operator measure (POM) [4], also known as a positive operator valued measure [26]. Each possible measurement outcome ω_i is associated with a probability operator, or POM element $\hat{\Pi}_i$. In order to form a physically realizable measurement, these elements must satisfy the conditions

$$\hat{\Pi}_i \geq 0, \qquad \sum_i \hat{\Pi}_i = \hat{I}. \tag{1}$$

The probability of obtaining outcome ω_j as a result of measurement on a system in state $\hat{\rho}$ is given by $\mathrm{Tr}(\hat{\rho}\hat{\Pi}_j)$.

Suppose that a measurement is made on a quantum system known to have been prepared in one of N possible states $\{\hat{\rho}_i\}$, with associated *a priori* probabilities $\{p_i\}$. Suppose further that the outcome of the measurement, denoted ω_j, is taken to imply that the state of the system was $\hat{\rho}_j$. No restrictions are placed on the number or interpretation of other possible outcomes; for the moment we are concerned only with outcome ω_j. How confident can we be that this outcome leads us to correctly identify the state prepared? The quantity of interest is the probability that the prepared state was $\hat{\rho}_j$, *given* that the outcome ω_j was obtained, that is, $P(\hat{\rho}_j|\omega_j)$. Using Bayes Rule we can write

$$P(\hat{\rho}_j|\omega_j) = \frac{P(\hat{\rho}_j)P(\omega_j|\hat{\rho}_j)}{P(\omega_j)} = \frac{p_j\,\mathrm{Tr}(\hat{\rho}_j\hat{\Pi}_j)}{\mathrm{Tr}(\hat{\rho}\hat{\Pi}_j)}, \tag{2}$$

where $\hat{\rho} = \sum_i p_i \hat{\rho}_i$ is the *a priori* density operator for the system. By maximizing $P(\hat{\rho}_j|\omega_j)$ with respect to the probability operator $\hat{\Pi}_j$, we can put a limit on how well the state $\hat{\rho}_j$ can be identified from the others in the set.

The process of maximizing $P(\hat{\rho}_j|\omega_j)$ is greatly facilitated by means of the ansatz

$$\hat{\Pi}_j = c_j \hat{\rho}^{-1/2} \hat{Q}_j \hat{\rho}^{-1/2}, \tag{3}$$

where $\hat{Q}_j$ is a positive, trace 1 operator, and thus the weighting factor $c_j \geq 0$ represents the probability of occurrence of outcome ω_j, $P(\omega_j)$. Hence

$$\begin{aligned} P(\hat{\rho}_j|\omega_j) &= p_j \operatorname{Tr}(\hat{\rho}^{-1/2}\hat{\rho}_j\hat{\rho}^{-1/2}\hat{Q}_j) \\ &= p_j \operatorname{Tr}(\hat{\rho}_j\hat{\rho}^{-1})\operatorname{Tr}(\hat{\rho}'_j\hat{Q}_j), \end{aligned} \tag{4}$$

where $\hat{\rho}'_j = \hat{\rho}^{-1/2}\hat{\rho}_j\hat{\rho}^{-1/2}/\operatorname{Tr}(\hat{\rho}_j\hat{\rho}^{-1})$. The operators $\hat{\rho}'_j$ and $\hat{Q}_j$ are both positive, with unit trace, and can be thought of as density operators. It follows, therefore, that $P(\hat{\rho}_j|\omega_j)$ is maximized if $\hat{Q}_j$ is a projector onto the pure state that has the largest overlap with $\hat{\rho}'_j$:

$$\hat{Q}_j = |\lambda_j'^{\max}\rangle\langle\lambda_j'^{\max}|, \tag{5}$$

where $|\lambda_j'^{\max}\rangle$ is the eigenket of $\hat{\rho}'_j$ corresponding to the largest eigenvalue $\lambda_j'^{\max}$. The limit is then given by

$$[P(\hat{\rho}_j|\omega_j)]_{\max} = p_j \operatorname{Tr}(\hat{\rho}_j\hat{\rho}^{-1})\lambda_j'^{\max} \tag{6}$$

and is realized by the POM element

$$\hat{\Pi}_j = c_j \hat{\rho}^{-1/2}|\lambda_j'^{\max}\rangle\langle\lambda_j'^{\max}|\hat{\rho}^{-1/2}. \tag{7}$$

If the state $\hat{\rho}_j$ is pure, then this simplifies to

$$\hat{\Pi}_j \propto \hat{\rho}^{-1}\hat{\rho}_j\hat{\rho}^{-1}. \tag{8}$$

As multiplying the POM element by a constant has no effect on the expression in Eq. (2), we have some freedom in choosing the constants of proportionality, and each choice will correspond to a distinct maximum confidence strategy. In some cases it is possible to choose the c_j such that we can form a complete measurement from operators independently optimized in this way. In other cases an inconclusive outcome is necessary, and the constants may be chosen, for example, to minimize the probability of occurrence of the inconclusive outcome.

As an example we consider the case of three equiprobable ($p_i = \frac{1}{3}$, $i = 0, 1, 2$) symmetric qubit states that lie on the same latitude of the Bloch sphere (see Fig. 1). For pure states $\hat{\rho} = |\Psi\rangle\langle\Psi|$, and we can describe our three states by the kets

$$\begin{aligned} |\Psi_0\rangle &= \cos\theta|0\rangle + \sin\theta|1\rangle, \\ |\Psi_1\rangle &= \cos\theta|0\rangle + e^{2\pi i/3}\sin\theta|1\rangle, \\ |\Psi_2\rangle &= \cos\theta|0\rangle + e^{-2\pi i/3}\sin\theta|1\rangle, \end{aligned} \tag{9}$$

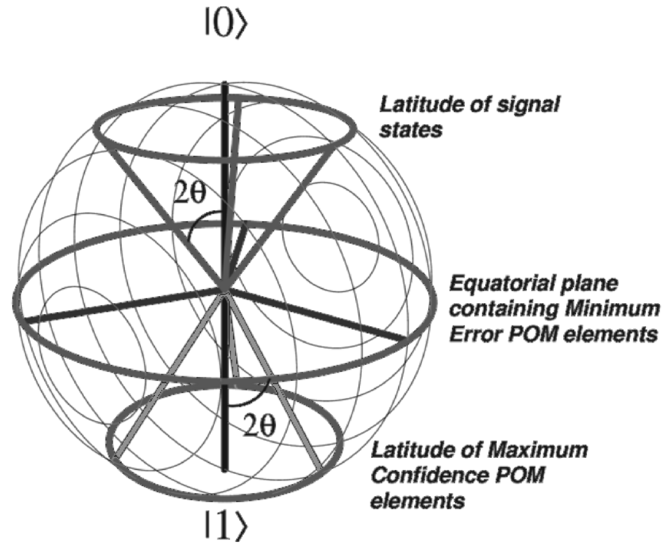

FIG. 1. Bloch sphere representation of states. Any density operator of a two-level system can be written $\hat{\rho} = \frac{1}{2}(\hat{\mathrm{I}} + \mathbf{r}\cdot\hat{\sigma})$ where $\mathbf{r}$ is a real 3 component vector, $\hat{\sigma} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$, and $\hat{\sigma}_{x(yz)}$ are the Pauli spin operators. By plotting the vector $\mathbf{r}$, states of a 2D complex system can be represented graphically in a 3D real system. The signal states used in the example, as well as the directions along which the POM elements lie are illustrated here.

where $|0\rangle$ and $|1\rangle$ form an orthogonal basis for the qubit and, without loss of generality, we set $0 \leq \theta \leq \pi/4$. For this set of states $\hat{\rho} = \cos^2\theta|0\rangle\langle 0| + \sin^2\theta|1\rangle\langle 1|$, and the maximum confidence POM elements are easily calculated using Eq. (8) to be $\hat{\Pi}_i = a_i|\phi_i\rangle\langle\phi_i|$ ($i = 0, 1, 2$), where the a_i are positive constants and

$$\begin{aligned} |\phi_0\rangle &= \sin\theta|0\rangle + \cos\theta|1\rangle, \\ |\phi_1\rangle &= \sin\theta|0\rangle + e^{2\pi i/3}\cos\theta|1\rangle, \\ |\phi_2\rangle &= \sin\theta|0\rangle + e^{-2\pi i/3}\cos\theta|1\rangle. \end{aligned} \tag{10}$$

Our maximum confidence if outcome ω_j is obtained, that is the maximum probability that the state identified is correct, is the same for each possible outcome, and is calculated from Eq. (6) to be

$$P(\hat{\rho}_j|\omega_j) = \tfrac{2}{3}. \tag{11}$$

The above elements do not form a POM for any choice of the constants a_i, and hence an inconclusive result is needed. The POM element corresponding to the inconclusive outcome is given by $\hat{\Pi}_? = \hat{\mathrm{I}} - \hat{\Pi}_0 - \hat{\Pi}_1 - \hat{\Pi}_2$, with a probability of occurrence $P(?) = \operatorname{Tr}(\hat{\rho}\hat{\Pi}_?) = 1 - 2(a_0 + a_1 + a_2)\cos^2\theta\sin^2\theta$. Different choices give competing maximum confidence strategies, and we need an additional criterion to select the best of these. One way to do this is to follow the example of unambiguous state discrimination and to minimize $P(?)$ subject to the constraint $\hat{\Pi}_? \geq 0$. As $P(?)$ is a monotonically decreasing function of a_0, a_1, a_2, the optimal values of these parameters lie on the boundary of the allowed domain, defined by $\hat{\Pi}_? \geq 0$. This leads us to choose a_0, a_1, a_2 to be

$$a_0 = a_1 = a_2 = (3\cos^2\theta)^{-1}. \tag{12}$$

The POM element corresponding to the inconclusive outcome is then of the form

$$\hat{\Pi}_? = (1 - \tan^2\theta)|0\rangle\langle 0|, \tag{13}$$

which gives the inconclusive probability

$$P(?) = \cos 2\theta. \tag{14}$$

For the purposes of comparison, we note that the minimum error POM for this set of states is given by the square root measurement [5,7], and can be written $\hat{\Pi}_i^{\mathrm{ME}} = \frac{2}{3}|\phi_i^{\mathrm{ME}}\rangle\langle\phi_i^{\mathrm{ME}}|$ where

$$|\phi_0^{\mathrm{ME}}\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle),$$
$$|\phi_1^{\mathrm{ME}}\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i/3}|1\rangle), \tag{15}$$
$$|\phi_2^{\mathrm{ME}}\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{-2\pi i/3}|1\rangle).$$

If this measurement is performed and outcome ω_j is obtained, then the probability that the state was, indeed, $\hat{\rho}_j$ is $P(\hat{\rho}_j|\omega_j) = \frac{1}{3}(1 + \sin 2\theta)$. This is plotted for comparison purposes alongside the optimal value of $\frac{2}{3}$ as a function of θ in Fig. 2. It can be seen that for all values of θ, except where the two strategies coincide at $\theta = \frac{\pi}{4}$, the new measurement strategy gives a greater confidence than that found for the minimum error strategy.

Analytic expressions were given above for the operators describing this new strategy for an arbitrary set of states [Eqs. (7) and (8)]. In deriving these the ansatz in Eq. (3) was used. The significance of this can be explained by reference to some general transformation described by the

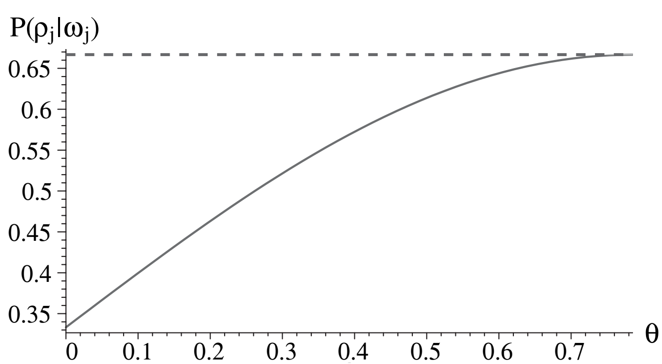

FIG. 2. Comparison of the confidence in the state identified as a result of measurement for minimum error (solid line) and maximum confidence (dashed line) strategies. $P(\hat{\rho}_j|\omega_j)$, the probability that the state identified on the basis of measurement outcome ω_j is correct is plotted as a function of the parameter θ.

invertible operator $\hat{A}$, which transforms $\hat{\rho} \rightarrow \hat{A}\,\hat{\rho}\,\hat{A}^\dagger$. Under this transformation, Eq. (2) becomes

$$P(\hat{\rho}_j|\omega_j) = \frac{p_j \operatorname{Tr}(\hat{A}\hat{\rho}_j\hat{A}^\dagger\hat{\Pi}_j')}{\operatorname{Tr}(\hat{A}\,\hat{\rho}\,\hat{A}^\dagger\hat{\Pi}_j')} \tag{16}$$

for some positive operator $\hat{\Pi}_j'$. It is clear that if we define $\hat{\Pi}_j' = (\hat{A}^\dagger)^{-1}\hat{\Pi}_j\hat{A}^{-1}$, this conditional probability is identical for the original system $\hat{\rho}$ and the transformed system $\hat{A}\,\hat{\rho}\,\hat{A}^\dagger$, under the action of the positive operators $\hat{\Pi}_j$ and $\hat{\Pi}_j'$, respectively. Furthermore, as $\hat{A}$ is invertible, this transformation describes a one-to-one mapping between operators on the original and transformed systems. Thus, if the operator achieving maximum confidence is known for one system, it is easy to find that for the other system, simply by applying the appropriate transformation. The advantage of the transformation used in Eq. (3) is that the operator $\hat{\Pi}_j' \propto \hat{Q}_j$ which maximizes this figure of merit for the transformed set $\{\hat{\rho}_j'\}$ is easily found.

The type of transformation discussed above can be realized as the result of a measurement associated with the POM $\{\hat{A}^\dagger\hat{A}, \hat{I} - \hat{A}^\dagger\hat{A}\}$. Thus the measurement described by the probability operators in Eq. (4) can be viewed as a two-step process. In the first step, a measurement is performed with outcomes ω_{succ}, ω_{fail}, and associated POM elements

$$\hat{\Pi}_{\mathrm{succ}} = \frac{p_{\mathrm{succ}}}{D}\hat{\rho}^{-1}, \qquad \hat{\Pi}_{\mathrm{fail}} = \hat{I} - \hat{\Pi}_{\mathrm{succ}}, \tag{17}$$

where D is the dimension of the state space of the system, and p_{succ} is the probability of occurrence of outcome ω_{succ}. To ensure positivity of both $\hat{\Pi}_{\mathrm{succ}}$ and $\hat{\Pi}_{\mathrm{fail}}$, p_{succ} must satisfy the condition $0 \leq p_{\mathrm{succ}} \leq \alpha_i D$ for all i, where α_i are the eigenvalues of $\hat{\rho}$.

When this step is performed, any given input state ρ_j is transformed to $\hat{\rho}_j'$, defined as above, with probability $p_{\mathrm{succ}}\operatorname{Tr}(\hat{\rho}_j\hat{\rho}^{-1})/D$ (corresponding to outcome ω_{succ}). This measurement strategy does not require that the operators in Eq. (4) form a complete measurement, and if outcome ω_{fail} is obtained, no further measurement is made, and the result may be interpreted as inconclusive.

The information provided by knowledge of the measurement outcome ω_{succ} also causes the associated probability distribution to be modified as follows:

$$p_j' = P(\hat{\rho}_j|\omega_{\mathrm{succ}}) = \frac{P(\hat{\rho}_j)P(\omega_{\mathrm{succ}}|\hat{\rho}_j)}{P(\omega_{\mathrm{succ}})} = \frac{p_j}{D}\operatorname{Tr}(\hat{\rho}_j\hat{\rho}^{-1}). \tag{18}$$

It is easily verified that

$$\sum_i p_i' = 1, \qquad \sum_i p_i'\hat{\rho}_i' = \frac{1}{D}\hat{I} = \hat{\rho}'. \tag{19}$$

Note that the operators which give maximum confidence for this new set are immediately clear as the operator $\hat{\rho}_j'$

describing any given state commutes with that describing the other states in the set $\hat{\rho}' - p_j\hat{\rho}'_j$. Thus, these two operators are simultaneously diagonalizable. Also, as $\hat{\rho}' \propto \hat{\mathrm{I}}$, the same eigenvector corresponds to the largest eigenvalue of ρ'_j and the smallest eigenvalue of $\hat{\rho}' - p_j\hat{\rho}'_j$. The optimum probability operator $\hat{\Pi}'_j$ is a projection onto this eigenvector [Eq. (5)].

The measurement described by operators $\{\hat{\Pi}'_j = \frac{c_j D}{p_{\mathrm{succ}}}\hat{Q}_j\}$ is thus the second step in the process. The probability of obtaining result ω_j when the system is in any given input state $\hat{\rho}_i$ can be written

$$\begin{aligned} P(\omega_j|\hat{\rho}_i) &= P(\omega_{\mathrm{succ}}|\hat{\rho}_i)P(\omega_j|\omega_{\mathrm{succ}}, \hat{\rho}_i) \\ &= \frac{p_{\mathrm{succ}}\,\mathrm{Tr}(\hat{\rho}_i\hat{\rho}^{-1})}{D}\mathrm{Tr}\Big(\hat{\rho}'_i\frac{c_j D}{p_{\mathrm{succ}}}\hat{Q}_j\Big) \\ &= \mathrm{Tr}(\hat{\rho}_i c_j\hat{\rho}^{-1/2}\hat{Q}_j\hat{\rho}^{-1/2}) = \mathrm{Tr}(\hat{\rho}_i\hat{\Pi}_j). \end{aligned} \quad (20)$$

Thus this two-step process is equivalent to the single step measurement described by probability operators $\{\hat{\Pi}_j, \hat{\Pi}_{\mathrm{fail}}\}$. From the discussion above we know that $\hat{Q}_j$ is a projector onto the eigenstate of $\hat{\rho}'_j$ with the largest eigenvalue. For pure states, as $\hat{Q}_j = \hat{\rho}'_j$, it is possible to choose the constants of proportionality such that $\hat{\Pi}'_j = p'_j D\hat{\rho}'_j$, and these probability operators form a complete measurement. For mixed states this is not possible, and the inconclusive outcome will have an additional component. Thus, for pure states the entire process may be interpreted as a projection of the initial states $\{\hat{\rho}_i\}$ to the transformed set $\{\hat{\rho}'_i\}$, followed by a measurement along these states.

What are the properties of the transformed set? As $\hat{\rho}' = \frac{1}{D}\hat{\mathrm{I}}$, the states span the entire state space. They are "maximally orthogonal" in the sense that the pure state with which any given state ρ'_j has the largest overlap is also that with which the average of the remaining states $\hat{\rho}' - p'_j\hat{\rho}'_j$ has the smallest overlap. In particular, for linearly independent sets, for which this strategy coincides with that of unambiguous discrimination, the initial states are projected onto mutually orthogonal states between which perfect discrimination is possible. This is exactly the way in which unambiguous discrimination between two nonorthogonal states has been realized experimentally [21,22]. For linearly dependent sets, the above may be made clearer by reference to a qubit system. The property $\hat{\rho}' = \frac{1}{2}\hat{\mathrm{I}}$ means that the 3D vector representing the state $\hat{\rho}'_j$ on the Bloch sphere points in the opposite direction to that representing $\hat{\rho}' - p'_j\hat{\rho}'_j$.

Thus, we have constructed a measurement which allows us to be as confident as possible that when a measurement outcome leads us to identify a particular state, that state was indeed present. As different outcomes are treated independently, an inconclusive outcome is sometimes necessary in order to form a physically realizable measurement. We have given analytic expressions for the operators describing this optimal measurement for an arbitrary set of initial states, and have interpreted these expressions in terms of a two-step measurement process. We have illustrated the new strategy by means of an example, and shown that for the set of states considered, when a state is identified, the probability that it was actually present is improved over the minimum error strategy. This strategy is analogous to unambiguous discrimination, but is applicable to linearly dependent states. We plan to demonstrate this strategy experimentally for the example considered here using optical polarization.

This work was supported by the Synergy fund of the Universities of Glasgow and Strathclyde and by the Royal Society.

*Electronic address: sarah@phys.strath.ac.uk

[1] A. Chefles, Contemp. Phys. **41**, 401 (2000).
[2] A. S. Holevo, J. Multivariate Anal. **3**, 337 (1973).
[3] H. P. Yuen, R. S. Kennedy, and M. Lax, IEEE Trans. Inf. Theory **IT-21**, 125 (1975).
[4] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).
[5] M. Ban, K. Kurokawa, R. Momose, and O. Hirota, Int. J. Theor. Phys. **36**, 1269 (1997).
[6] S. M. Barnett, Phys. Rev. A **64**, 030303 (2001).
[7] Y. C. Eldar and G. D. Forney, IEEE Trans. Inf. Theory **IT-47**, 858 (2001).
[8] E. Andersson, S. M. Barnett, C. R. Gilson, and K. Hunter, Phys. Rev. A **65**, 044307 (2002).
[9] I. D. Ivanovic, Phys. Lett. A **123**, 257 (1987).
[10] D. Dieks, Phys. Lett. A **126**, 303 (1988).
[11] A. Peres, Phys. Lett. A **128**, 19 (1988).
[12] G. Jaeger and A. Shimony, Phys. Lett. A **197**, 83 (1995).
[13] A. Peres and D. Terno, J. Phys. A **31**, 7105 (1998).
[14] A. Chefles, Phys. Lett. A **239**, 339 (1998).
[15] E. B. Davies, IEEE Trans. Inf. Theory **IT-24**, 596 (1978).
[16] M. Sasaki, S. M. Barnett, R. Jozsa, M. Ozawa, and O. Hirota, Phys. Rev. A **59**, 3325 (1999).
[17] S. M. Barnett, C. R. Gilson, and M. Sasaki, J. Phys. A **34**, 6755 (2001).
[18] K. Hunter, E. Andersson, C. R. Gilson, and S. M. Barnett, J. Phys. A **36**, 4159 (2003).
[19] S. M. Barnett and E. Riis, J. Mod. Opt. **44**, 1061 (1997).
[20] R. B. M. Clarke, V. M. Kendon, A. Chefles, S. M. Barnett, E. Riis, and M. Sasaki, Phys. Rev. A **64**, 012303 (2001).
[21] B. Huttner, A. Muller, J. D. Gautier, H. Zbinden, and N. Gisin, Phys. Rev. A **54**, 3783 (1996).
[22] R. B. M. Clarke, A. Chefles, S. M. Barnett, and E. Riis, Phys. Rev. A **63**, 040305 (2001).
[23] J. Mizuno, M. Fujiwara, M. Akiba, S. M. Barnett, and M. Sasaki, Phys. Rev. A **65**, 012315 (2002).
[24] S. M. Barnett, Quantum Inf. Comput. **4**, 450 (2004).
[25] R. L. Kosut, I. A. Walmsley, Y. C. Eldar, and H. Rabitz, quant-ph/0403150.
[26] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer Academic Publishers, Dordrecht, 1993).