Nonlinear operations in quantum-information theory

Daniel R. Terno*
Department of Physics, Technion–Israel Institute of Technology, 32 000 Haifa, Israel

(Received 10 November 1998)

Quantum-information theory is used to analyze various nonlinear operations on quantum states. The universal disentanglement machine is shown to be impossible, and partial (negative) results are obtained in the state-dependent case. The efficiency of the transformation of nonorthogonal states into orthogonal ones is discussed. [S1050-2947(99)08705-3]

PACS number(s): 03.67.−a

The rules of a quantum mechanics make certain processes impossible. Nonorthogonal quantum states cannot be cloned [1]. This is one of the fundamental theorems of the quantum information theory. On other hand, there are explicit constructions of quantum circuits that would perform many interesting transformations of quantum states [2,3], provided that certain devices, like quantum XOR gates [4], can be actually built. Between these two extremes there is a gray area of operations that are not obviously forbidden by the elementary laws of quantum mechanics, but may be ruled out by more careful considerations. The most interesting of them are nonlinear operations, and their analysis from the point of view of quantum-information theory [5,6] is the subject of this paper.

The basic procedures of quantum mechanics — unitary transformations and projections — are linear. However, nonlinear operations with quantum states are common, even if they are not always regarded as such. Nonlinearity naturally enters via selective operations, namely those that involve a filtering at one of their steps. These operations have a finite probability to fail and may look quite exotic, e.g., the transformation [2]

$$\rho^{\rm in}=\begin{pmatrix}\rho_{11} & \rho_{12}\\ \rho_{21} & \rho_{22}\end{pmatrix}\rightarrow\rho^{\rm out}=\begin{pmatrix}\rho_{11}^2 & \rho_{12}^2\\ \rho_{21}^2 & \rho_{22}^2\end{pmatrix}. \tag{1}$$

Nevertheless, the possibility of failure is not a necessary condition of nonlinearity. An operation may be nonselective, i.e., it is always completed successfully, but there may be a demand for certain inputs to be transformed into certain outputs. Here we are interested only in the processing of these particular input states. For example, if we want to process the output of a two-state cloning machine [3], the transformation of the two resulting pure entangled states (labeled 1 or 2) into the direct product of their reduced density matrices,

$$\rho_{1,2}\rightarrow {\rm Tr}_A\,\rho_{1,2}\otimes{\rm Tr}_B\,\rho_{1,2} \tag{2}$$

would be useful. What happens to all other states is irrelevant and unspecified; everything that makes this process work will be accepted. If we are interested only in these two states, the transformation looks nonlinear.

The most general physical operation

$$\rho\rightarrow\rho'=\frac{\mathsf{T}\rho}{{\rm Tr}\,\mathsf{T}\rho} \tag{3}$$

is represented by a superoperator T: $\mathcal{B}(\mathcal{H})_1\rightarrow\mathcal{B}(\mathcal{H})_1$, which is (i) positive, (ii) satisfies $\mathsf{T}^\dagger\mathsf{T}\leqslant\mathsf{1}$, and (iii) is completely positive [7,8]. The validity of the last condition [9] has recently become a hot issue in numerous discussions, but it certainly holds for systems which are initially unentangled with their environment.

According to the ''first representation theorem'' [8]

$$\mathsf{T}\rho=\sum_k \mathsf{A}_k\rho\mathsf{A}_k^\dagger, \tag{4}$$

where the set of A_k's satisfies

$$\sum_k \mathsf{A}_k^\dagger\mathsf{A}_k\leqslant\mathsf{1}. \tag{5}$$

In the problem that is considered here, the transformation in Eq. (2) is nonselective, so that $\Sigma_k\mathsf{A}_k^\dagger\mathsf{A}_k=\mathsf{1}$.

When a ''quantum black box'' [10] is specified, the superoperator T and the corresponding matrices A_k can be inferred from the complete set of input and output data. Similarly, if we extend Eq. (2) to a larger set of states and then use the approach of Chuang and Nielsen [10], we can confirm or refute different realizations of the desired transformation. If the answer is positive the problem is solved. On the contrary, a negative answer tells us nothing. Another realization may still work.

Quantum-information theory can discern absolutly impossible processes from tentatively possible ones. We look at the desired operation as a part of some hypothetical decision scheme, which aims at distinguishing between different input states [6,11,12], or as a communication channel [13] with input and output alphabets given by the left- and right-hand sides of an expression like Eq. (2).

Using Eqs. (3), (4) it is easy to show that any chain of the measurements and/or transformations of the system, with or without *ancilla*, can be described as a single measurement [14] represented by a positive operator-valued measure (POVM). Thus if the proposed procedure allows us to im-

*Electronic address: terno@physics.technion.ac.il

prove *any* of the distinguishability criteria of the input states beyond the optimal value, it contains some flaw and therefore is not physical.

One of these distinguishability criteria is the probability of error (P_E), defined as follows [6,11]: an observer is given one of two states ρ_1 and ρ_2. That state is secretly chosen, and since we are interested in the intrinsic difference between the states, the probilities to pick up either of the states are equal. The task of the observer is to decide, after performing anything that is allowed by quantum mechanics, which state was given to him. The probability that the observer makes a wrong guess with the best possible decision scheme gives a measure of the distinguishability. Two orthogonal states can be distinguished perfectly, thus giving $P_E=0$. When the states are identical $P_E=\frac{1}{2}$. In general the optimal result is

$$P_E(\rho_1,\rho_2)=\frac{1}{2}-\frac{1}{4}\mathrm{Tr}|\rho_1-\rho_2|. \quad (6)$$

Any transformation that would give a lower PE certainly violates the laws of quantum mechanics. Unfortunately, the usefulness of this criterion is limited to the case of two states, which is the only one for which a closed solution is known.

Another criterion is the accessible information $I(\rho_1,\rho_2)$, which is defined as a maximal mutual information over all possible decision schemes. For a set of states ρ_i with fixed *a priori* probabilities π_i each possible measurement scheme X gives a probability distribution of the results. We calculate the mutual information between the probability distributions corresponding to different states. Finally the maximum is taken over all possible measurements. In the case of two inputs the accessible information is given by [5,6,12,13]

$$I(\rho_1,\rho_2):=\max_X[I(p(\rho_1,X),p(\rho_2,X))]. \quad (7)$$

This definition is naturally extended to more than two states. Accessible information cannot increase, but there are only a handful of cases where it is explicitly known, in particular two pure states and two spin-$\frac{1}{2}$ states [15].

When I is unknown we can use different unequalities that relate the accessible information to other distinguishability criteria [6,12]. The most useful of them involves a new notion, which is called entropy defect or relative entropy [5] and is given by

$$\Delta S(\rho_1,\ldots,\rho_n)=S(\overline{\rho})-\sum_i \pi_i S(\rho_i), \quad (8)$$

where

$$S(\rho)=-\mathrm{Tr}\,\rho\ln\rho, \quad (9)$$

is the von Neumann entropy, $\overline{\rho}=\Sigma_i\pi_i\rho_i$, and π_i is the *a priori* probability distribution. In order to measure the intrinsic difference between two of the states, we set $\pi_1=\pi_2=\frac{1}{2}$, as in the definition of P_E. The Levitin-Holevo inequality states that the entropy defect is an upper bound of the accessible information [5]

$$I(\rho_1,\rho_2)\leqslant\Delta S(\rho_1,\rho_2). \quad (10)$$

Moreover, the entropy defect does not increase under trace preserving completely positive maps [13],

$$\Delta S(\mathsf{T}\rho_1,\mathsf{T}\rho_2)\leqslant\Delta S(\rho_1,\rho_2). \quad (11)$$

It is obvious that the compliance with known information bounds is only a necessary condition that the proposed transformation must satisfy. However, it is a powerful tool, as the following examples illustrate.

Since the exact cloning of quantum states, $\rho\rightarrow\rho\otimes\rho$ is impossible [1], let us consider the best approximate cloners. The optimal two-state cloning machine [3] is a device which has one of the two possible pure spin-$\frac{1}{2}$ states as the input and produces an entangled pair on the output. Two identical reduced density matrices are close to the cloned state with exceptionally high fidelty, $F=\mathrm{Tr}[(\mathrm{Tr}_A\rho_i^{\mathrm{out}})|\psi_i^{\mathrm{in}}\rangle\langle\psi_i^{\mathrm{in}}|]>0.985$. Although the fidelity of the reduced density matrices is very high, it will be shown that it is impossible to separate the output according to the prescription of Eq. (2). If the two input states are parametrized as

$$|u\rangle=\cos\theta|0\rangle+\sin\theta|1\rangle,$$
$$|v\rangle=\sin\theta|0\rangle+\cos\theta|1\rangle, \quad (12)$$

the output states are

$$|u'\rangle=(a\cos\theta+c\sin\theta)|00\rangle+b(\cos\theta+\sin\theta)(|01\rangle+|10\rangle)+(c\cos\theta+a\sin\theta)|11\rangle,$$
$$|v'\rangle=(c\cos\theta+a\sin\theta)|00\rangle+b(\cos\theta+\sin\theta)(|01\rangle+|10\rangle)+(a\cos\theta+c\sin\theta)|11\rangle, \quad (13)$$

where $a(\theta)$, $b(\theta)$, and $c(\theta)$ are given in the Appendix.

Before looking at the state-dependendent disentanglement machine, it is easy to see why the universal device is impossible. A transformation

$$\rho\rightarrow\mathrm{Tr}_A\,\rho\otimes\mathrm{Tr}_B\,\rho, \quad (14)$$

for all input states is essentially nonlinear. Thus it cannot correspond to any physical process [16]. More explicitly, if Eq. (14) holds for two arbitrary states ρ_1 and ρ_2, it also holds for their convex combination [7]

$$\rho_x=x\rho_1+(1-x)\rho_2,\quad 0\leqslant x\leqslant 1. \quad (15)$$

If Eq. (14) represents a linear transformation, then

$$\mathrm{Tr}_A\,\rho_x\otimes\mathrm{Tr}_B\,\rho_x=x\mathrm{Tr}_A\,\rho_1\otimes\mathrm{Tr}_B\,\rho_1+(1-x)\mathrm{Tr}_A\,\rho_2\otimes\mathrm{Tr}_B\,\rho_2. \quad (16)$$

On the other hand,

$$\mathrm{Tr}_A\,\rho_x\otimes\mathrm{Tr}_B\,\rho_x=x^2\mathrm{Tr}_A\,\rho_1\otimes\mathrm{Tr}_B\,\rho_1+(1-x)^2\,\mathrm{Tr}_A\,\rho_2\otimes\mathrm{Tr}_B\,\rho_2+x(1-x)(\mathrm{Tr}_A\,\rho_1\otimes\mathrm{Tr}_B\,\rho_2+\mathrm{Tr}_A\rho_2\otimes\mathrm{Tr}_B\,\rho_1), \quad (17)$$

which is clearly a contradiction.

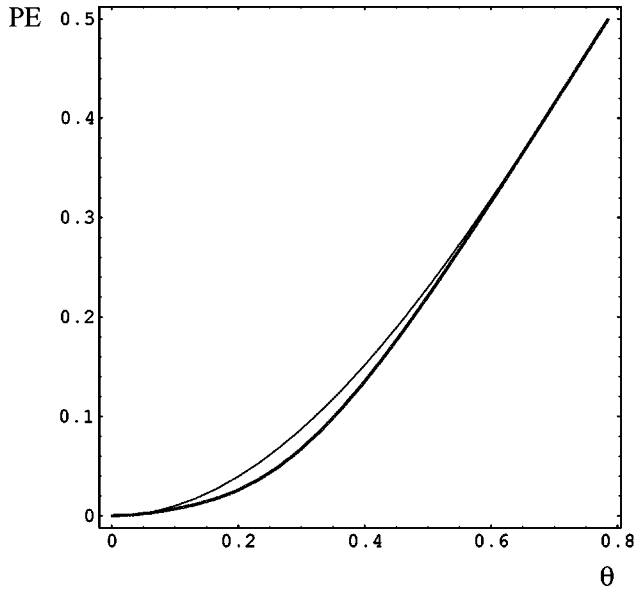

FIG. 1. Probabilities of error: thin line is in the optimal measurement, thick line is in the optimal measurement after disentanglement.

Now let us suppose that there is a disentanglement procedure which is described by Eq. (2). We apply the Helstrom's decision scheme to the output of the state-dependent cloner and to the direct product of the disentangled copies, $\mathrm{Tr}_A\,\rho_i^{\mathrm{out}}\otimes\mathrm{Tr}_B\,\rho_i^{\mathrm{out}}$. In both cases the analytical expressions for PE can be found explicitly. They are given in the Appendix and their graphs are plotted in Fig. 1. We see that the disentanglement of the copies decreases PE and, as a result, this process is impossible. A closer look into the construction of the optimal measurement reveals that the proposed transformation is realized by an operator which is not *positive* and, consequently, can represent no physical process.

Now, let us again consider the states

$$|u\rangle=a|00\rangle+b(|01\rangle+|10\rangle)+c|11\rangle,$$
$$|v\rangle=c|00\rangle+b(|01\rangle+|10\rangle)+a|11\rangle, \tag{18}$$

where the coefficients a, b, and c are arbitrary real numbers subject only to the normalization, $a^2+2b^2+c^2=1$. They can be parametrized by spherical coordinates as

$$a=\sin\vartheta\cos\varphi,\quad b=\sin\vartheta\sin\varphi/\sqrt{2},\quad c=\cos\vartheta. \tag{19}$$

It is possible to derive analytical expressions for P_E and P_{Ed}, which are given in the Appendix. $P_E(\vartheta,\varphi)$ is the actual optimal result, while the decision process which leads to $P_{Ed}(\vartheta,\varphi)$ includes the hypothetical disentanglement procedure. Obviously, the regions of the (ϑ,φ) plane where $P_E<P_{Ed}$ are forbidden, i.e., the disentanglement procedure cannot be realized. Figure 2 shows these areas together with the line that corresponds to the disentanglement of the output of the optimal cloner, which lies in one of the forbidden domains.

Another bound can be obtained by using the entropy defect. The explicit expressions for ΔS and ΔS_{d} will not be given here, because they are too cumbersome. Proceeding

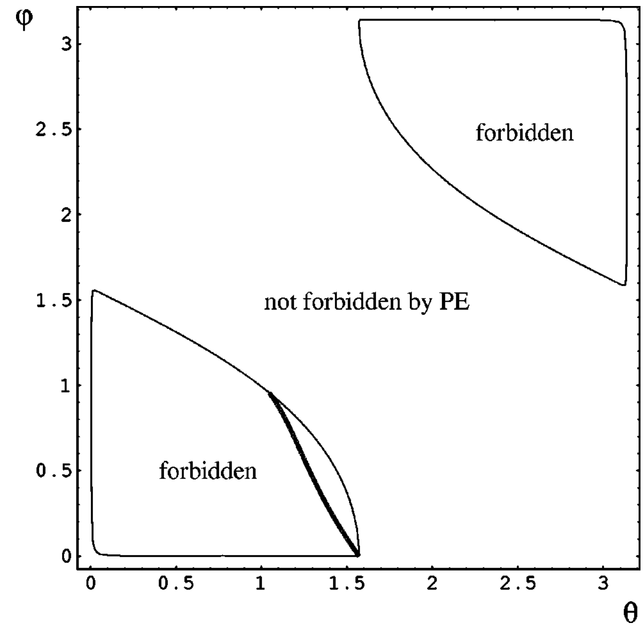

FIG. 2. Domains of the parameters for the disentanglement transformation according to the P_E criterion. The states lying on the border are not forbidden.

exactly as in the previous case, we look for the regions where $\Delta S\leqslant\Delta S_{\mathrm{d}}$, i.e., where the disentanglement (2) is impossible.These domains are presented in Fig. 3. It is instructive to compare it with Fig. 2. The boundary of the regions forbidden by P_E coincides with some parts of the boundary obtained by ΔS. However, the differences are clear. This partial agreement requires further investigation and may give some new insights on the relationship between different distinguishability criteria.

In the examples that we considered above, these criteria give identical predictions for the states with an equal degree of entanglement. For pure states the degree of entanglement

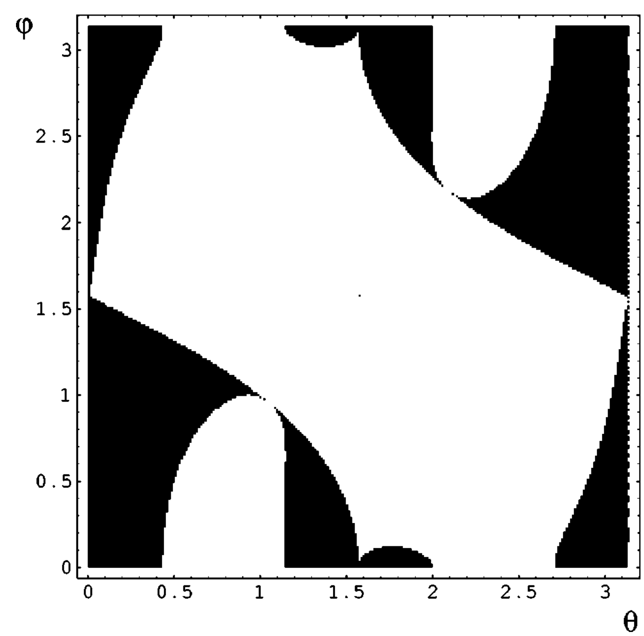

FIG. 3. Domains of the parameters for the disentanglement transformation according to the ΔS criterion. The states lying on the border are not forbidden. The forbidden domain is drawn in black.

is measured by a von Neumann entropy of either subsystem [17],

$$E(\rho)=S(\mathrm{Tr}_A\rho)=S(\mathrm{Tr}_B\rho). \tag{20}$$

The eigenvalues of either of $\mathrm{Tr}_A|u\rangle\langle u|$ or $\mathrm{Tr}_A|v\rangle\langle v|$ are

$$\lambda_\pm=\frac{1}{2}(1\pm(a+c)\sqrt{1+2b^2-2ac}), \tag{21}$$

These eigenvalues (and, as a result, the degrees of the entanglement of the corresponding states) are unchanged under the transformation $\vartheta\rightarrow\pi-\vartheta$, $\varphi\rightarrow\pi-\varphi$. It is also a symmetry of Eqs. (A6)–(A9) and the expressions for ΔS.

A more intricate question [18] is whether there is a transformation that takes an entangled state into a separable one [19] and preserves the reduced density operators,

$$\rho\rightarrow\tilde{\rho}=\sum_i w_i\rho_i^{\mathrm{A}}\otimes\rho_i^{\mathrm{B}}, \mathrm{Tr}_A\,\tilde{\rho}=\mathrm{Tr}_A\,\rho, \mathrm{Tr}_B\,\tilde{\rho}=\mathrm{Tr}_B\,\rho. \tag{22}$$

Some time after this paper was submitted for publication, it was shown by Mor [20] that a universal disentanglement machine into separable states is impossible.

As another example, we analyze the information approach to the transformation of pure nonorthogonal states into orthogonal ones. Recently [2] a quantum circuit that does this operation for spin-$\frac{1}{2}$ states was proposed. Since nonorthogonal quantum states cannot be distinguished with certainty, such a transformation has only a limited probability of success. It consists of applying a XOR gate [4] to the pair of identically prepared particles in either of the states and measuring the spin of the second particle in z direction. If the spin is ''down'' the transformation succeeds. In this case the outputs are $|\phi^{\mathrm{out}}\rangle_1\otimes|\downarrow\rangle$ or $|\phi_2^{\mathrm{out}}\rangle\otimes|\downarrow\rangle$; and we have $\langle\phi_2^{\mathrm{out}}|\phi_1^{\mathrm{out}}\rangle=0$ even if initially there was an overlap between the states.

Leaving the explicit calculations aside, let us look at the bounds on this operation. Orthogonal states can be, in principle, distinguished unambiguously. Thus it is possible to use this procedure as part of an error-free scheme of discrimination between two nonorthogonal states [2]. The exact solution of this problem [21] is given by a POVM with two outputs which correspond to the unambiguous results, A_1 and A_2, and an output which is a failure of the measurement, $\mathsf{A}_?=1-\mathsf{A}_2-\mathsf{A}_1$. The optimal POVM is explicitly known and the probability of having a definite answer is

$$P=1-|\langle\phi_1|\phi_2\rangle|. \tag{23}$$

Since the direct products of two copies of the same states have an overlap equal to $|\langle\phi_1|\phi_2\rangle|^2$, the probability of success of any transformation of two nonorthogonal states into orthogonal ones is bounded by

$$P(\mathrm{success})\leqslant 1-|\langle\phi_1^{\mathrm{in}}|\phi_2^{\mathrm{in}}\rangle|^2. \tag{24}$$

Moreover, there should be a transformation that achieves this upper bound: if the state is known, any number of its copies can be produced. As a result, a transformation which consists of the unambiguous state identification and the corresponding preparation achieves the bound of Eq. (24). However, the more reasonable (and physically fruitful) transformation of [2] has this maximal efficiency too. The importance of Eq. (24) is that generalizations of this result [22,23] are valid for an arbitrary number of linearly independent states. For example, if m copies of three pure nonorthogonal states are used to produce the orthogonal ones, the efficiency of the operation is bounded by [22]

$$P(\mathrm{success})\leqslant 1-\sum_{i=1}^{3} k_i|[\phi_1^{\mathrm{in}}\phi_2^{\mathrm{in}}\phi_3^{\mathrm{in}}]|^{2m}/3, \tag{25}$$

where coefficients k_i depend on the relative orientation of the state vectors and $[uvw]$ stands for the triple product of the vectors, i.e., the determinant of their components, in any basis.

Part of this work was performed during the Elsag-Bailey-ISI research meeting on quantum computation, Turin. It is a pleasure to thank Dagmar Bruss, Isaac Chuang, and Lev Levitin for useful discussions. The help of Chris Fuchs and Asher Peres is gratefully acknowledged. This work was supported by a grant from the Technion Graduate School.

APPENDIX

Parameters of the cloning machine:

$$a=\frac{1}{\cos 2\theta}[\cos\theta(P+Q\cos 2\theta)-\sin\theta(P-Q\cos 2\theta)], \tag{A1}$$

$$b=\frac{1}{\cos 2\theta}P\sin 2\theta(\cos\theta-\sin\theta), \tag{A2}$$

$$c=\frac{1}{\cos 2\theta}[\cos\theta(P-Q\cos 2\theta)-\sin\theta(P+Q\cos 2\theta)], \tag{A3}$$

$$P=\frac{1}{2}\frac{\sqrt{1+\sin 2\theta}}{\sqrt{1+\sin^2 2\theta}}, \tag{A4}$$

$$Q=\frac{1}{2}\frac{\sqrt{1-\sin 2\theta}}{\cos 2\theta}. \tag{A5}$$

Probabilities of error for the output of the cloning machine:

$$P_E(\theta)=\frac{1}{2}-\frac{1}{2}|\cos^2\theta-\sin^2\theta|, \tag{A6}$$

$$P_{Ed}(\theta)=\frac{1}{2}-\frac{1}{\sqrt{2}}\frac{\sqrt{\cos^2 2\theta(13-10\cos 4\theta+\cos 8\theta+6\sin 2\theta-2\sin 6\theta)}}{\sqrt{(3-\cos 4\theta)^3}}. \tag{A7}$$

Probabilities of error in a more general case:

$$P_E(\vartheta,\varphi)=\frac{1}{2}-\frac{1}{2}|\cos^2\vartheta-\cos^2\varphi\sin^2\vartheta|\times\sqrt{\cos^2\vartheta+\cos^2\varphi\sin^2\vartheta+2\sin^2\varphi\sin^2\vartheta+\cos\varphi\sin 2\vartheta}, \quad \text{(A8)}$$

$$P_{Ed}(\vartheta,\varphi)=\frac{1}{2}-\frac{1}{2}|\cos^2\vartheta-\cos^2\varphi\sin^2\vartheta|\sqrt{F(\vartheta,\varphi)}, \quad \text{(A9)}$$

where F is

$$F(\vartheta,\varphi)=\cos^4\vartheta+\cos^2\vartheta\sin^2\vartheta(3-\cos 2\varphi)+4\cos\varphi\cos\vartheta\sin^2\varphi\sin^3\vartheta+(5-\cos 4\varphi)\sin^4\vartheta/4.$$

[1] W. K. Wootters and W. H. Zurek, Nature (London) **299**, 802 (1982); D. Dieks, Phys. Lett. **92A**, 271 (1982).

[2] H. Bechman-Pasquinucci, B. Huttner, and N. Gisin, Phys. Lett. A **242**, 198 (1998).

[3] D. Bruss, D. P. DiVinzenzo, A. Ekert, C. A. Fuchs, C. Macchiavello, and J. A. Smolin, Phys. Rev. A **57**, 2368 (1998).

[4] A. Barenco, D. Deutch, A. Ekert, and R. Jozsa, Phys. Rev. Lett. **74**, 4083 (1995).

[5] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, Dordrecht, 1993), Chap. 9.

[6] C. A. Fuchs, Ph.D. thesis, University of New Mexico, 1995, e-print quant-ph/9601020.

[7] E. B. Davies, *Quantum Theory of Open Systems* (Academic, New York, 1976), Chap. 1.

[8] K. Kraus, *States, Effects and Operations* (Springer-Verlag, Berlin, 1983).

[9] M. Czachor and M. Kuna, Phys. Rev. A **58**, 128 (1998), and references therein.

[10] I. L. Chuang and M. A. Nielsen, J. Mod. Opt. **44**, 2455 (1997).

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

[12] C. A. Fuchs and C. M. Caves, Open Systems and Information Dynamics **3**, 1 (1995); C. A. Fuchs and J. van de Graaf, IEEE Trans. Inf. Theory (to be published), e-print quant-ph/9712042.

[13] A. S. Holevo, *Coding Theorems for Quantum Channels*, in *Tamagawa University Research Review* (1998), e-print quant-ph/9809023.

[14] G. M. D'Ariano and H. P. Yuen, Phys. Rev. Lett. **76**, 2832 (1996).

[15] L. B. Levitin, Open Systems and Information Dynamics **2**, 319 (1994).

[16] This conclusion was also reached independently by C. H. Bennett and C. A. Fuchs (private communications).

[17] C. H. Bennett, D. P. DiVincenzo, J. S. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996); V. Vedral and M. Plenio, *ibid.* **57**, 1619 (1998).

[18] C. A. Fuchs (private communication).

[19] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996); M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).

[20] T. Mor, e-print quant-ph/9812020.

[21] D. Dieks, Phys. Lett. A **126**, 303 (1988); A. Peres, *ibid.* **128**, 19 (1988).

[22] A. Peres and D. R. Terno, J. Phys. A **31**, 7105 (1998).

[23] M. Ban, K. Kurokawa, R. Monmose, and O. Hirota, Int. J. Theor. Phys. **36**, 1269 (1997); A. Chefles and S. M. Barnett, J. Phys. A **31**, 10 097 (1998).