Information transfer and fidelity in quantum copiers

P. Deuar* and W. J. Munro
Centre for Laser Science, Department of Physics, University of Queensland, QLD 4072, Brisbane, Australia
(Received 22 December 1999; published 12 May 2000)

We find that very different quantum copying machines are optimal depending on the indicator used to assess their performance. Several quantum copying machine models acting on nonorthogonal input states are investigated, and assessed according to two types of criteria: transfer of (Shannon) information encoded in the initial states to the copies, and fidelity between the copies and the initial states. Transformations that optimize information transfer for messages encoded in qubits are found for three situations: (1) when the message is decoded one state at a time; (2) with simple schemes that allow the message to be encoded using block-coding schemes; and (3) when the copier produces independent copies. If the message is decoded one symbol at a time, information is best copied by a Wootters-Zurek copier.

PACS number(s): 03.67.Hk, 03.65.Bz

I. INTRODUCTION

Quantum copying has attracted considerable interest in recent years, ever since the discovery of the no-cloning theorem [1,2], and the universal quantum copying machine [3] which copies arbitrary unknown qubits with the best fidelity. To date, most treatments have used fidelity to characterize the quality of the copies produced. The fidelity between two quantum states characterized by density operators $\hat{\rho}_1$ and $\hat{\rho}_2$ is

$$F(\hat{\rho}_1,\hat{\rho}_2)=\left\{\mathrm{Tr}\left[\sqrt{(\hat{\rho}_1)^{1/2}\hat{\rho}_2(\hat{\rho}_1)^{1/2}}\right]\right\}^2. \tag{1}$$

A good summary of its properties is given in Ref. [4]. In the case where one of the states is pure, the fidelity is simply the square of the overlap between the two states.

Many authors [5–15] have made use of two fidelity measures for quantum copiers: the *global fidelity* of the combined output (both copies) of the copier, with respect to a product state of (unentangled) perfect copies, and the *local fidelity* of one copy with respect to the original input state. Here, we will concentrate on a different indicator of copying success: mutual information content between the copies and the originals. One finds that which copier is optimal depends greatly on which indicator is used. In practice, this will mean that what sort of quantum copier is best depends on what one wants to do with the copies afterward.

This article proceeds in the following fashion. After commenting on some drawbacks of fidelity, and why one might want to use different indicators, we outline exactly what we mean by information content between copies and originals in Sec. II. General features of the copiers that will be considered are mentioned in Sec. III. Copiers optimized for maximum copied information are given in Sec. IV (and derivations are given in Appendixes A and B) for three cases: (1) when the information is decoded from the copies one state at a time; (2) when efficient block-coding schemes are used to transmit as much information as is allowed by the Holevo bound; and (3) when the copies are an unentangled product state. In Sec. V the performance of these copiers is assessed according to information transfer and fidelity criteria, and compared to the performance of fidelity-optimized copiers known previously.

II. MUTUAL INFORMATION AND FIDELITY MEASURES

A. Fidelity, and some of its drawbacks

Fidelity is used in many fields as an indicator of closeness between two states, and is often quite useful. It is probably also one of the easiest such indicators to calculate. However, it sometimes suffers from a number of drawbacks (examples of which are given below) when used as a measure of closeness over broad classes of systems, so there will be times when one wants to use a different indicator.

While a fidelity of 1 obviously implies identical states, and 0 implies orthogonal states, what intermediate values mean is highly dependent on the particular states that are being compared, particularly if both states are impure. Thus a statement such as ''The fidelity between the two states was x,'' to be unambiguous, often needs considerable additional information on the states that were compared. To give an example: For standard optical coherent states of complex amplitude α, given by

$$|\alpha\rangle=e^{-(1/2)|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{n!}|n\rangle, \tag{2}$$

the fidelity between two pure coherent states $|\alpha\rangle$ and $|\alpha+1\rangle$ is always constant:

$$F(|\alpha\rangle\langle\alpha|,|\alpha+1\rangle\langle\alpha+1|)=\frac{1}{e}. \tag{3}$$

Now if $\alpha=0$, the two states are the vacuum and a low-photon-number coherent state—states with qualitatively different properties. However, if α is large, then $|\alpha\rangle$ and $|\alpha+1\rangle$ are macroscopic, and experimentally indistinguishable, but the fidelity between them is still $1/e$.

Another drawback of fidelity is that it is not directly related to other quantities commonly measured in experiments. While the fidelity is an expectation value of an observable

*Electronic address: deuar@physics.uq.edu.au

(the observable being either one of the two states), it cannot usually be calculated from the results of experiments whose aim is to do something other than measure fidelity. It is not in general directly related to expectation values or measurement probabilities of other quantities, so it does not say much about the usefulness of a copy. In this sense, fidelity characterizes the closeness of the mathematical representation of physical states more than the closeness of the physical properties of those states. Of course, in many situations, these two types of closeness are equivalent, but not always.

For the specific case of quantum copiers, global or local fidelities are not robust to unitary transformations made on the copies individually after all copying has been completed, and also can be very high even though the copies are uncorrelated with the originals. For example, suppose a message is encoded in a binary alphabet of orthogonal states $|0\rangle,|1\rangle$, and sent through a lossless communication channel that interchanges the states, i.e., they undergo the transformation

$$|0\rangle \rightarrow |1\rangle, \tag{4}$$

$$|1\rangle \rightarrow |0\rangle, \tag{5}$$

then the fidelity of the transmitted with respect to the initial state is *zero*, but nothing of interest has been lost. It is sufficient for an observer receiving the message to relabel the states which they receive to recover the original message.

Conversely, consider the situation where very nonorthogonal states $|a\rangle$ and $|b\rangle$ are used to encode a message. Using appropriate error-correction schemes, some information can be reliably transmitted with this encoding. However, now suppose that the message is intercepted by an eavesdropper, who simply sends the same state $\sqrt{1/2}\,(|a\rangle+|b\rangle)$ on to the intended receiver every time. The fidelity between sent and received states is still very high, but the received message carries no information from the sender.

Global fidelity measures are often particularly removed from experimental results, since they compare the combined state of both copies with a perfect copy state that is generally unattainable due to the no-cloning theorem. However, in practice, one usually makes copies so they can subsequently be considered only individually.

B. Mutual information measures

A different, natural, measure of copying efficiency that can be used is the amount of mutual (Shannon) information [16] shared between the original states, and the copies. This mutual information does away with some of the drawbacks of fidelity, as discussed below.

Consider two observers: one of them, the sender (labeled A), is sending states chosen from some ensemble, where the *a priori* probability of sending the ith variety of state is P_i^A. The other observer, the receiver (labeled B), makes measurements on one of the copies, obtaining the jth measurement result with probability $P_{j|i}^B$, given that the ith state was sent into the copier. The amount of information (in bits per sent state) that the receiver has obtained from the sender is the Shannon mutual information, given by

$$I(A{:}B)=\sum_{i,j} P_i^A P_{j|i}^B \log_2 \frac{P_{j|i}^B}{P_j^B}, \tag{6}$$

where P_j^B is the overall probability of the receiver obtaining the jth measurement result, averaged over the input states.

To use this measure to characterize a copying machine, rather than the specific message encoding or the ingenuity of the receiver in constructing a measuring apparatus, three points should be noted. First, even if a perfect copier is used, the amount of information that can be transmitted from originals to copy depends on the ensemble of states that is used to encode the message. Thus, the information about the original extractable from the copy $I(A{:}B)$ must be compared to the amount of information extractable from the original $I(A{:}A)$.

Second, if observer B makes a suboptimal (in terms of recovering the original message) set of measurements, then B's stupidity will affect the mutual information. To eliminate the effect of B's ingenuity (or lack of it), it has to be assumed that optimal measurements are made to recover the encoded message.

Third, a characterization of the copier would usually involve examining its information-copying performance for a given set of input states. However, these may occur with various *a priori* probabilities P_i^A. We will take the case where these probabilities are chosen to encode the maximum amount of information in the signal states to be most representative of the behavior of information in the copier. Thus the mutual information quantities that will be used in later sections of this article are $I_m(A{:}B)$ and $I_m(A{:}A)$, given by

$$I_m(A{:}B)=\max_{\{P_i^A\}}[\max_{\{\mathcal{E}_B\}} I(A{:}B)], \tag{7}$$

where $\{P_i^A\}$ denotes the set of *a priori* probabilities of A using the ith state in the encoding of the message, and $\{\mathcal{E}_B\}$ is the set of all positive valued operator measures [17]. We will call I_m the *copied information*.

While this quantity can be more laborious to calculate, it has some advantages over fidelity. It is unchanged by relabeling or by local unitary transformations on the copies after they have left the copier, as well as always being zero if the copies are independent of the originals.

Also, such mutual information is a physical quantity of interest in its own right, and is in fact what one is interested in many fields (such as cryptography, for example). Even where this is not the case, mutual information between originals and copies can often be calculated from probability distributions of experimental measurements. Furthermore, it is clear what the statement ''the mutual information transfer from A to B is x'' means physically, with no further knowledge of the actual quantum states that were sent. It could be said that the information-copying capacity of a quantum cloner quantifies the *practical usefulness*, in many situations, of the copies produced by it.

There is a qualitative difference between information-theoretic quantities such as copied information, and quantities such as fidelity. Fidelity, and similar quantities such as the Hilbert-Schmidt norm, or the Bures distance, are quanti-

fications of relations between two quantum states (or, more precisely, between their mathematical representations), while information-theoretic quantities deal with the relations between *ensembles* of states. This is the reason that they are robust to such postcopying effects as relabeling of the copy states.

C. Ultimate and one-state copied information

Consider the situation discussed in the preceding subsection. Observer A encodes a message into a sequence of quantum states, chosen from a set of states $\{\hat{\rho}_i^A\}$ labeled by the index i. Each of the sent states has an *a priori* probability P_i^A of being the ith one in the set. When the copying machine acts on the signal state $\hat{\rho}_i^A$, it produces a copy state $\hat{\rho}_i^B$, which is usually different from the original. It has been shown [18,19] that the mutual information between A and B can be no more than $I_H(A:B)$, given by

$$I_m(A:B) \leq I_H(A:B) = S\left(\sum_i P_i^A \hat{\rho}_i^A\right) - \sum_i P_i^A S(\hat{\rho}_i^A), \tag{8a}$$

where $S(\hat{\rho})$ is the von Neumann quantum entropy of state $\hat{\rho}$:

$$S(\hat{\rho}) = -\mathrm{Tr}[\hat{\rho}\log_2\hat{\rho}], \tag{8b}$$

a result known as the Holevo theorem.

In practice, the transmitted information will usually be significantly less than $I_H(A:B)$. However, it has been shown [20,21] that if A encodes the message using only certain sequences of states out of all the possible ones (although still respecting the *a priori* probabilities of individual states), and B makes measurements on whole such sequences rather than on individual states, then as the length of these sequences increases, the information capacity per state can approach arbitrarily close to the Holevo bound $I_H(A:B)$. This is called a block-coding scheme, and such a communication setup is analogous to sending and distinguishing only whole ''words'' at a time in the message, rather than individual ''letters.'' In this analogy, letters correspond to individual quantum states, and words to sequences of them. Naturally, only special choices of the ''words'' to be used will approach the Holevo bound, Eq. (8).

With this in mind, there are two obvious candidates for a mutual information quantity with which to characterize copiers: the *ultimate copied information* given by I_H, and the *one-state copied information* I_1, which is the maximum information obtainable if measurements are made on only one state at a time. Both will be considered in what follows.

III. GENERAL PROPERTIES OF THE COPYING SETUPS CONSIDERED

In the interest of clarity and simplicity (and, one must admit, ease of analysis), only the most basic relevant copying setups have been investigated. This should make the principles involved easier to see, without introducing too much complexity.

Thus, we will consider the case where observer A encodes a message into a binary sequence of pure quantum states $\hat{\rho}_i^A = |\psi_i^A\rangle\langle\psi_i^A|$ $(i=1,2)$ with equal *a priori* probabilities of being sent $(P_i^A = \frac{1}{2})$. The P_i^A are chosen to be one-half for two reasons: (1) this is the simplest case; (2) this is the situation where the maximum amount of information is encoded in the input states.

Since there are only two input states, the dimension of the relevant Hilbert space can be reduced to 2 by appropriate unitary transformations, because the states span at most a two-dimensional manifold in Hilbert space. Any such can be written (discarding an irrelevant phase factor) in an orthogonal basis $\{|+\rangle, |-\rangle\}$ as

$$|\psi_1^A\rangle = \cos\theta|+\rangle + e^{i\mu}\sin\theta|-\rangle, \tag{9a}$$

$$|\psi_2^A\rangle = \sin\theta|+\rangle + e^{-i\mu}\cos\theta|-\rangle, \tag{9b}$$

where the parameter θ ranges from 0 to $\pi/4$ (other values of θ are equivalent to a relabeling of the two states). In the rest of the article, μ will be taken to be zero for simplicity, although all results can easily be extended to the nonzero case. This, then, gives a one-parameter family of input states:

$$|\psi_1^A\rangle = \cos\theta|+\rangle + \sin\theta|-\rangle, \tag{10a}$$

$$|\psi_2^A\rangle = \sin\theta|+\rangle + \cos\theta|-\rangle. \tag{10b}$$

These can be fully labeled by the fidelity between them,

$$f = F(\hat{\rho}_1^A, \hat{\rho}_2^A) = \sin^2(2\theta). \tag{11}$$

In a similar fashion, by taking the least complex case, the copiers considered will be unitary, create only two copies, and be symmetric. By symmetric we mean that the reduced quantum states of both copies by themselves are equal.

The unitarity of the copying process implies a ''black box'' process: no external disturbance is required during the copying. Probabilistic copiers [10,22] are not considered here.

Physically, there are two subsystems o and c (which can be considered two dimensional for reasons outlined above) put into the unitary copying machine, and two come out. At the input, the subsystem o contains the original state to be copied, while c contains a ''blank'' state that is always the same, irrespective of what enters at o. Both subsystems contain the (usually imperfect) copies when they exit the copier, while an ancillary machine state subsystem (x) is also used in some of the copiers. At the input, all three subsystems are unentangled, while at the output, entanglement is usually present. Due to unitarity, the full entangled output states consisting of all three subsystems o, c, and x are pure, but the states of individual subsystems are in general mixed.

IV. THREE INFORMATION-OPTIMIZED QUANTUM COPIERS

In this section, we present transformations for several copiers optimized for information transfer to the copies, given a binary sequence of equiprobable input states. All these copiers are symmetric. The input states are in general nonorthogonal, and the degree of orthogonality is characterized by f, the square of the overlap between the two input states $\hat{\rho}_1^A$ and $\hat{\rho}_2^A$. These will be compared to known fidelity-optimized copiers in the next section.

A. Copiers that optimize the one-state copied information

Rather than carry out a tedious optimization, it stands to reason that if any unitary copier allows one to extract as much information about the originals from the copies as from the originals themselves, then it achieves the optimum. Is there such a copier?

Perhaps surprisingly, one finds that the Wootters-Zurek (WZ) quantum copying machine [1,3] (used in the original proof of the no-cloning theorem) allows one to extract as much information (using a one state at a time extraction) from either of the copies as from the original. One can imagine that the same information transfer could be achieved by making measurements on the originals, and sending the results classically, but that a simple unitary transformation with no coupling to the external environment can achieve the same is perhaps less obvious. What is more, the WZ copier does much better than any fidelity-optimized copiers, as will be seen later.

Explicitly, the transformation of the input states (10) is given by

$$|\psi_1^A\rangle \rightarrow \sin\theta|++\rangle + \cos\theta|--\rangle, \tag{12a}$$

$$\psi_2^A\rangle \rightarrow \cos\theta|++\rangle + \sin\theta|--\rangle, \tag{12b}$$

where the basis vectors $|+-\rangle$, etc., indicate tensor products $|+\rangle_o|-\rangle_c$ of the basis vectors for the o and c copy subsystems, respectively. The combined state of the copies is highly entangled, but the reduced density matrices of the copies (the full output density matrices traced over all subsystems except one copy) are in the classically mixed states

$$\hat{\rho}_1^B = \begin{pmatrix} \cos^2\theta & 0 \\ 0 & \sin^2\theta \end{pmatrix}, \tag{13a}$$

$$\hat{\rho}_2^B = \begin{pmatrix} \sin^2\theta & 0 \\ 0 & \cos^2\theta \end{pmatrix}. \tag{13b}$$

The one-state copied information, which is the same as can be extracted from the originals, is

$$I_1^{\mathrm{WZ}} = \tfrac{1}{2}[(1+q)\log_2(1+q) + (1-q)\log_2(1-q)], \tag{14a}$$

where q, which we will call the distinguishability parameter, is

$$q = \sqrt{1-f}. \tag{14b}$$

From the purely classical nature of $\hat{\rho}_i^B$, it follows that the ultimate copied information I_H^{WZ} is no bigger than I_1^{WZ}. In fact, applying more WZ copying machines to the copies made by the first one, in a cascade effect, creates larger numbers of copies, each of which still carries the same amount of (one-state) information as the original message. In this way, arbitrary numbers of optimal copies can be made—similarly to how one can make arbitrary numbers of copies of classical information.

The local fidelity between a copy and the originals is

$$F(\hat{\rho}_i^A, \hat{\rho}_i^B) = 1 - \frac{f}{2}. \tag{15}$$

There are other copiers related to the WZ copier which allow the same optimal one-state information transfer. One example is the family of copying transformations created by applying identical local unitary transformations on both copies after they come out of the WZ copier. The particular transformation presented above in Eq. (12) is the one that gives the best local fidelity out of this family of transformations.

B. Copiers without ancilla that optimize the ultimate copied information

It is also of interest how well information can be transmitted when the possibility of complicated block-coding schemes is allowed, as discussed in Sec. II C. To make the calculations relatively tractable analytically, we have made two restrictions on the copiers that we considered for this task.

First, only copiers that do not use an ancillary subsystem x, entangled with the copies, have been considered. It is probably possible to obtain somewhat better performance in ultimate information copying by using such helper subsystems, since discarding x after copying is completed partially relaxes the conditions that the copy states o and c must satisfy to preserve unitarity (since one then has more parameters left to optimize over). It is not clear how much better one could do with such helper states, but we suspect not much better, since from Fig. 2 below, the copier considered here is only marginally better than several others obtained by optimizing over different indicators such as fidelity and one-state copied information.

Second, for similar reasons, we have assumed that since both possible input states $\hat{\rho}_i^A$ are of equal purity $\mathrm{Tr}[(\hat{\rho}_i^A)^2]$ (totally pure, in fact), then both reduced copy states $\hat{\rho}_i^B$ will be of equal purity also:

$$\mathrm{Tr}[(\hat{\rho}_1^B)^2] = \mathrm{Tr}[(\hat{\rho}_2^B)^2]. \tag{16}$$

This is also a property shared by all other copiers mentioned in this article. The usual assumptions of Sec. III, such as both copies being equal, apply also.

So, an ancillaless copier, that produces two identical (usually imperfect) copies of any of two possible pure signal states, that makes copies of the same purity whichever of the two input states is sent, and that (given the above) maximizes the amount of information that can be transmitted to each of the copies by any block-coding scheme when the two input states are equiprobable, is given by the somewhat lengthy characterization below. The details of how this was obtained have been left for Appendix A.

There is a whole family of copying transformations, related by local unitary transformations on the copies after they have stopped interacting with each other, which give the same ultimate information copied I_H^u. Of these, we will specify that particular one in this family which gives the greatest local fidelity between the copies and originals. The transformation can be written in terms of the parameters r_m and ϕ_m, which have to be determined numerically. In terms of the initial states (10),

$$|\psi_1^A\rangle \to \sqrt{\frac{1+r_m}{2}}|b_1\rangle + \sqrt{\frac{1-r_m}{2}}|b_2\rangle, \tag{17a}$$

$$|\psi_2^A\rangle \to \sqrt{\frac{x}{2}}|b_1\rangle + \sqrt{\frac{x}{2} - r_m \cos\phi_m}|b_2\rangle + \sqrt{\frac{1-x+r_m\cos\phi_m}{2}}(|b_3\rangle + |b_4\rangle), \tag{17b}$$

where

$$x = \tfrac{1}{2}(1+\cos^2\phi_m + 2r_m\cos\phi_m + \sqrt{1-r_m^2}\sin^2\phi_m), \tag{17c}$$

and the four $|b_j\rangle$ are orthogonal basis states, given in terms of the usual $|+\rangle$ and $|-\rangle$ basis states used in Eqs. (10) and (12) by the matrix equation

$$\begin{pmatrix} |b_1\rangle \\ |b_2\rangle \\ |b_3\rangle \\ |b_4\rangle \end{pmatrix} = U \begin{pmatrix} |++\rangle \\ |-+\rangle \\ |+-\rangle \\ |--\rangle \end{pmatrix}, \tag{18}$$

where the unitary matrix U is

$$U = \frac{1}{2}\begin{pmatrix} 1+\sin\phi_m/2 & 1-\sin\phi_m/2 & \cos\phi_m/2 & \cos\phi_m/2 \\ 1-\sin\phi_m/2 & 1+\sin\phi_m/2 & -\cos\phi_m/2 & -\cos\phi_m/2 \\ -\cos\phi_m/2 & \cos\phi_m/2 & 1+\sin\phi_m/2 & \sin\phi_m/2-1 \\ -\cos\phi_m/2 & \cos\phi_m/2 & \sin\phi_m/2-1 & 1+\sin\phi_m/2 \end{pmatrix}. \tag{19}$$

As can be seen from the above, the basis states $|b_j\rangle$ are entangled over the two copies.

The parameter ϕ_m is actually the angle between the Bloch vectors of the two possible reduced copy states $\hat{\rho}_i^B$, which can be written

$$\hat{\rho}_1^B = \frac{1}{2}\begin{pmatrix} 1+q & q_H \\ q_H & 1-q \end{pmatrix}, \tag{20a}$$

$$\hat{\rho}_2^B = \frac{1}{2}\begin{pmatrix} 1-q & q_H \\ q_H & 1+q \end{pmatrix}, \tag{20b}$$

where the parameters q and q_H are

$$q = r_m \sin\frac{\phi_m}{2}, \tag{21a}$$

$$q_H = r_m \cos\frac{\phi_m}{2}, \tag{21b}$$

and appear in the expressions for I_1 and I_H.

Now $\cos\phi_m$ is dependent on r_m, and is given in terms of it as the second largest [23] real root of the following quartic polynomial in $\cos\phi_m$:

$$\begin{aligned} 0 = {} & \cos^4\phi_m[r_m^2(2-r_m^2-2\sqrt{1-r_m^2})] + \cos^3\phi_m[4r_m^2(1 \\ & -\sqrt{1-r_m^2})] + \cos^2\phi_m\{2[r_m^4+2r_m^2+4f(\sqrt{1-r_m^2}-1)]\} \\ & + \cos\phi_m[4r_m^2(1+\sqrt{1-r_m^2}-4f)] + [(4f-1)^2-(1 \\ & -r_m^2)^2 + 2(r_m^2-4f)\sqrt{1-r_m^2}]. \end{aligned} \tag{22}$$

The ultimate copied information is given by

$$\begin{aligned} I_H^u = {} & \tfrac{1}{2}[(1+r_m)\log_2(1+r_m) + (1-r_m)\log_2(1-r_m)] \\ & - \tfrac{1}{2}[(1+q_H)\log_2(1+q_H) + (1-q_H)\log_2(1-q_H)], \end{aligned} \tag{23}$$

which can be made a function of r_m only, using Eq. (22). To obtain the optimum copier, we find numerically the value of r_m that maximizes I_H^u on $r_m \in [\sqrt{1-f}, 1]$.

The one-state copied information I_1^u is given by the same expression in the distinguishability parameter q as for the WZ copier [Eq. (14a)], with q now given by Eq. (21a).

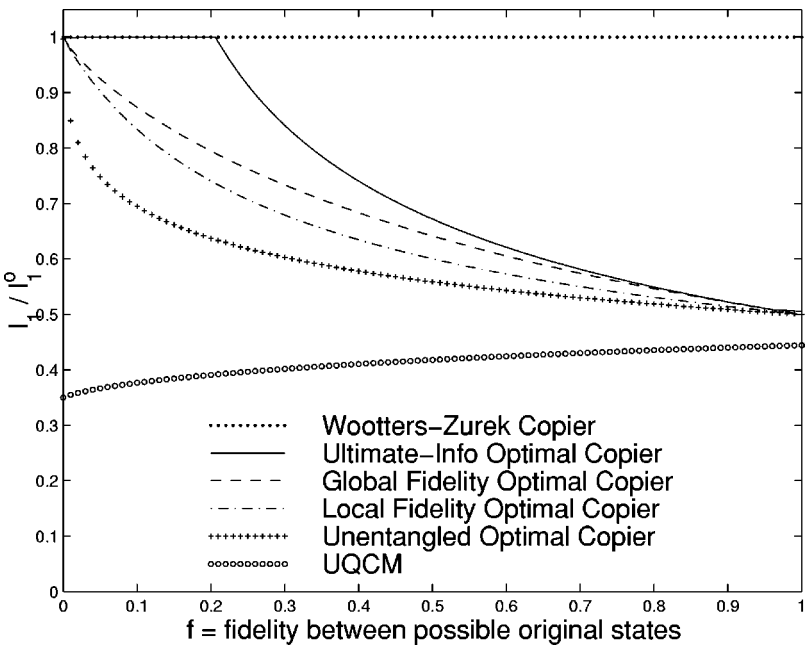

FIG. 1. One-state copied information (in bits per signal state) I_1 for the copying machines discussed in Secs. IV and V and Appendix C, as a fraction of the maximum one-state information I_1^o extractable from the input states (10), plotted as a function of the fidelity f between the two pure input signal states.

It is interesting to note that, for input states which are sufficiently nonorthogonal ($f \lesssim 0.206$), the copier given here is just the WZ copier described in Sec. IV A. In these cases, $\phi_m = \pi$ and $r_m = \sqrt{1-f}$. This sudden change in behavior (particularly evident in Fig. 1 and Fig. 3 below) may be due to excluding the use of ancillary subsystems. Allowing these may make the I_H optimal copier consistently better (although possibly not by much) than the Wootters-Zurek for all values of f, even the small ones.

The local fidelity between copies and originals for this copier is

$$F(\hat{\rho}_i^A, \hat{\rho}_i^B) = \tfrac{1}{2}(1 + q\sqrt{1-f} + q_H\sqrt{f}). \tag{24}$$

C. An optimal copier that gives unentangled copies

As has been remarked by many previously, optimal quantum copiers typically produce highly entangled copies. This also applies to the two quantum copiers given in Secs. IV A and IV B. Nevertheless, copies of some quality can be made without entanglement between them. This might be desirable in some situations.

Once again two simplifying assumptions have been made to make the calculation easier. It has been assumed that the copies are, again, unentangled with ancillary machine states, and that the output state of the copier is simply a product state of the two identical copies, rather than a classical mixture of several such product states. The case with additional machine states present might allow somewhat higher infor-

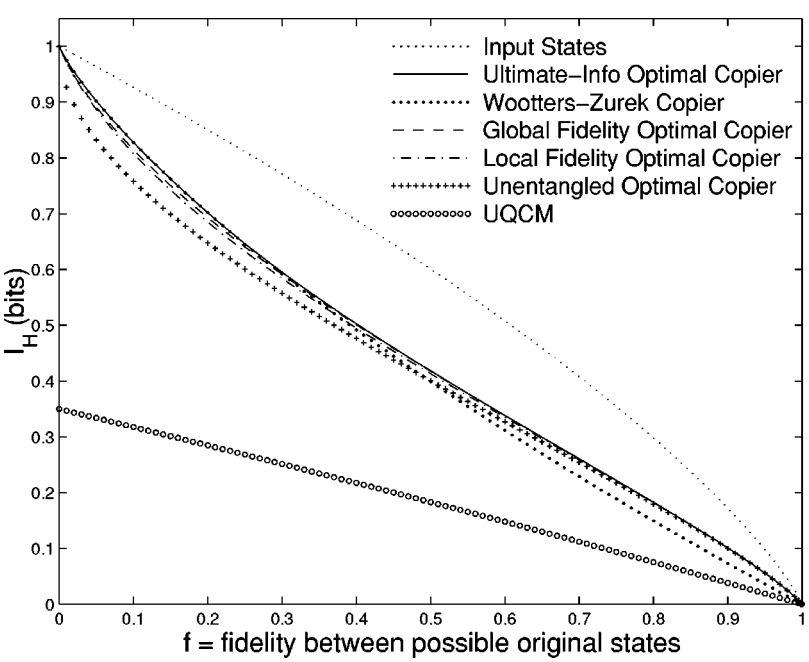

FIG. 2. Ultimate (Holevo bound) copied information (in bits per signal state) I_H for the copying machines discussed in Secs. IV and V and Appendix C, depending on the fidelity f between the two pure input signal states. The Holevo bound on information extractable from the originals is also given under the name ''Input States.''

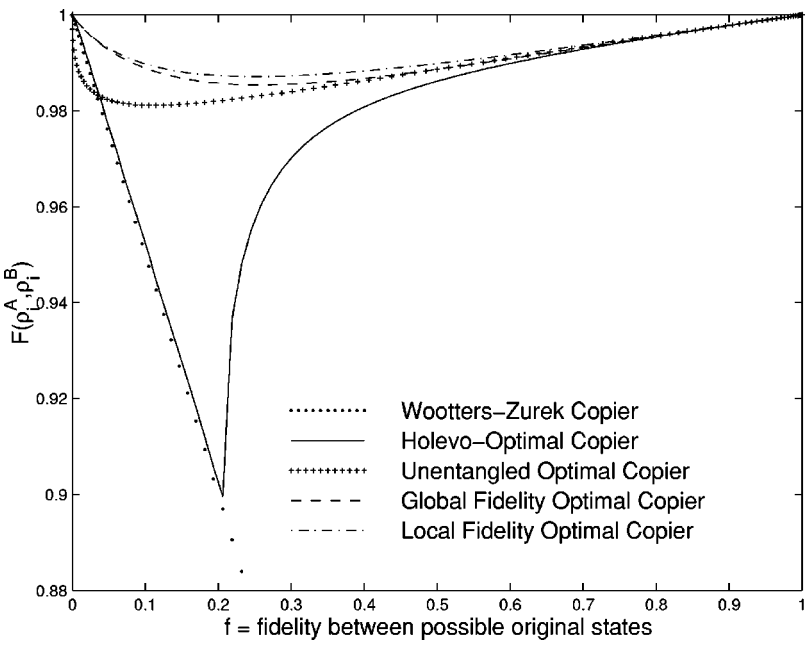

FIG. 3. Local fidelity $F(\hat{\rho}_i^A, \hat{\rho}_i^B)$ between a copy and the original, for the copying machines discussed in Secs. IV and V and Appendix C, as a function of f, the fidelity between the two input signal states.

mation transmission I_H with block-coding methods, for the same reasons as in Sec. IV B. This would be interesting to check, but we have not done this to date. Allowing classical correlations between copies and a machine state subsystem x does not, however, improve information transmission.

Given the above two restrictions, a copier that optimizes both the one-state and ultimate copied information, while keeping the copies unentangled, is given by

$$|\psi_1^A\rangle \to \frac{1+\sqrt{1-\sqrt{f}}}{2}|++\rangle + \frac{1-\sqrt{1-\sqrt{f}}}{2}|--\rangle + \tfrac{1}{2}f^{1/4}(|+-\rangle + |-+\rangle), \tag{25a}$$

$$|\psi_2^A\rangle \to \frac{1-\sqrt{1-\sqrt{f}}}{2}|++\rangle + \frac{1+\sqrt{1-\sqrt{f}}}{2}|--\rangle + \tfrac{1}{2}f^{1/4}(|+-\rangle + |-+\rangle), \tag{25b}$$

with notation identical to Eqs. (12). See Appendix B for details of the optimization.

This gives pure state copies (they must be pure from the unitarity of the transformation, since the input states are pure, and the output state is $\hat{\rho}_i^B \otimes \hat{\rho}_i^B$)

$$\hat{\rho}_1^B = \frac{1}{2}\begin{pmatrix} 1+\sqrt{1-\sqrt{f}} & f^{1/4} \\ f^{1/4} & 1-\sqrt{1-\sqrt{f}} \end{pmatrix}, \tag{26a}$$

$$\hat{\rho}_2^B = \frac{1}{2}\begin{pmatrix} 1-\sqrt{1-\sqrt{f}} & f^{1/4} \\ f^{1/4} & 1+\sqrt{1-\sqrt{f}} \end{pmatrix}. \tag{26b}$$

A family of copiers which do as well in the information measures, but worse in local fidelity between originals and copies, is given by making unitary transformations on the copies individually.

The one-state copied information I_1^{NE} is given by the same expression in q as for the WZ copier (14a), with q now given by

$$q = \sqrt{1-\sqrt{f}}. \tag{27}$$

The ultimate copied information is

$$I_H^{\mathrm{NE}} = 1 - \frac{1+f^{1/4}}{2}\log_2(1+f^{1/4}) - \frac{1-f^{1/4}}{2}\log_2(1-f^{1/4}). \tag{28}$$

The local fidelity of copies with respect to originals is

$$F(\hat{\rho}_i^A, \hat{\rho}_i^B) = \tfrac{1}{2}[f^{3/4} + 1 + \sqrt{(1-f)(1-\sqrt{f})}]. \tag{29}$$

It turns out that this copier also gives the best local fidelity out of such unentangling copiers (see Appendix B).

V. A COMPARISON OF THE COPIERS

To see how well the copiers rate in terms of the information measures I_H and I_1, we first need to determine how much information could be extracted from the input states if they were not copied. Since the input states are not orthogonal for $f>0$, then a full bit of information cannot be extracted from each state even though they are equiprobable.

One finds that the information extractable one state at a time is

$$I_1^o = \tfrac{1}{2}[(1+q)\log_2(1+q) + (1-q)\log_2(1-q)], \tag{30}$$

with the distinguishability parameter $q=\sqrt{1-f}$. This is the same as with the Wootters-Zurek copier (14a). The ultimate information extractable from the signal if block-coding methods are used is, however, unlike that for the WZ copier, much larger:

$$I_H^o = 1 - \frac{1+\sqrt{f}}{2}\log_2(1+\sqrt{f}) - \frac{1-\sqrt{f}}{2}\log_2(1-\sqrt{f}). \tag{31}$$

It is interesting to compare the performance of the copiers given in Sec. IV to previously known fidelity-optimized ones. Three will be considered here, and a brief summary of the copies they produce is given in Appendix C in terms of the input state overlap parameter *f*.

These three copiers are as follows. (1) The universal quantum copying machine [3] (UQCM), which copies arbitrary qubits with a local fidelity of 5/6. This is the maximum possible if it is to copy all with equal fidelity. (2) A copier found by Bruß *et al.* [8] that optimizes the global fidelity when copying one of two nonorthogonal input states. (3) A copier also found by Bruß *et al.* [8,24] that optimizes the local fidelity when copying one of two nonorthogonal input states. So let us see how they compare in performance.

A. One-state copied information

The one-state copied information is a good indicator of the efficiency of communicating classical data to the two copies. The recovery and coding of the information in this case relies only on measurement of one-qubit states, and classical error-correction schemes.

Looking at Fig. 1, one sees that the Wootters-Zurek copier, apart from achieving the optimum and transmitting as much one-state information to both copies as was encoded originally, is also far better at it than any of the other copiers shown (except for the small-*f* region, where the ultimate-information optimized copier becomes the WZ). The WZ copier has by far the simplest transformation out of these copiers, so it seems that for basic information transmission the simplest copier is the best.

The fidelity-optimized copiers do not do as well as the WZ, which in itself is to be expected, as after all they were optimized for fidelity, not information transfer. However, they do very much worse, causing the loss of much information that could be regained if better copiers were used. This shows quite clearly that fidelity is not necessarily a good measure of the quality of the copies for all situations. It is perhaps also surprising that, even though we are considering information transmitted to *one* copy here, the copier that has been optimized for global fidelity between the combined output state and perfect copies, does significantly better than the one that has been optimized for local fidelity between a single copy and original.

The UQCM gives much less information transfer than the other copiers, since all the others have been specifically tailored for the two signal states, whereas the UQCM must handle any arbitrary states with equal fidelity.

The copiers that give optimum unentangled copies do generally significantly worse than the other copiers apart from the UQCM, but one sees that all the copiers apart from the WZ copier and UQCM converge to the same efficiency (much worse than the optimum) for high values of *f*, i.e., when the signal states are not very orthogonal.

Note that a plot of the actual (rather than relative) amount of information extractable from the original signal I_1 is shown in Fig. 2 as the Wootters-Zurek curve, since $I_H^{\mathrm{WZ}} = I_1^o$.

B. Ultimate copied information

The ultimate (Holevo bound) copied information I_H gives an absolute maximum on how much information could possibly be transmitted by a given copier, with the best signaling scheme that is possible. In general, to achieve this bound, the encoding/decoding scheme has to be very elaborate, and it is often not achievable in practice due to complexity. In the case of qubit systems being transmitted here, this would entail making measurements of many-qubit observables to decode the information: a difficult task at present.

As can be seen in Fig. 2, most of the copiers cluster just below the optimal capacity achieved by the copier of Sec. IV B. While this is not necessarily the absolute optimum that can be achieved, as there remains the possibility that introducing helper machine states may increase this bound, this bunching makes it seem plausible that no large gains can be achieved beyond this. This ultimate-information optimal copier is quantitatively not much better than the Wootters-Zurek copier. Its greatest gains, which are still quite modest, come when the overlap between signal states is high, where the absolute information content in the signal is small.

It can be seen that, while the no-cloning theorem did not stop one from perfectly copying information contained in one state at a time, its effect is strong where block-coding schemes are allowed. This is because, if we restrict ourselves to the one case at a time situation, we are not utilizing those properties of the states that are affected by the no-cloning theorem. The difference between what can be extracted from a copy and from the originals is quite striking, and for highly overlapping input states, over 60% of the information in the originals is unavailable from a copy.

The behavior of the copiers for high overlap between states is as one would expect. That is, the Wootters-Zurek copier becomes much less efficient than the others when block-coding schemes are used, as the other copiers do not fully entangle the copies with each other, thus allowing one to extract some extra information by looking at several sequential states together.

Since the Wootters-Zurek copier has $I_1 = I_H$, by comparing the values of I_H for the local and global-fidelity-optimized copiers to the WZ copier, one can see that for these fidelity-optimal copiers, much more information than I_1 can be sent to the copies by allowing complicated block-coding schemes which use correlations between subsequent signal states. This approach, however, is unhelpful with the Wootters-Zurek copier, and is of very little help when using the UQCM.

As for the other information measure, the global-fidelity-optimized copier does slightly better than the local fidelity one. The unentangled copier does slightly worse than the rest, except for the UQCM which is consistently worse on all counts, as it is not tailored to the input states like the others.

C. Local fidelity

This is shown for various copiers in Fig. 3. The UQCM is absent from the plot, as its local fidelity lies far below the others shown there. Figures 1 and 3 show quite clearly that fidelity and information transfer quantify quite different properties of the copying transformation, and one has to keep in mind which properties are desired, before deciding on a quantity to characterize efficiency.

As expected, the best local fidelity occurs for the copier that was optimized for this, and the global fidelity optimal copier is almost as good. The WZ copier is no good at fidelity at all for significantly overlapping states. The unentangled copier is once again slightly worse than most of the others. The sharp change in behavior for the ultimate-information optimal copier is particularly evident in this plot.

VI. COMMENTS AND CONCLUSIONS

It was seen in the previous section that quantum copiers optimized for fidelity measures are far from optimal for basic information transmission to the copies, and, vice versa, information-optimized copiers are far from optimized for fidelity between copies and originals. This indicates that various measures of quality should be used for quantum devices, depending on what final use is to be made of the states created.

Some other general trends that were seen for the quantum copying devices that were considered, include the following. The ultimate-copied-information-optimized copier behaves more similarly to the fidelity-optimized ones than to the one-state optimized WZ copier (where it differs from the WZ). The fidelity-optimized copiers are not bad when one allows multiparticle measurements on the copies, but are far from optimal if one does not. This may be because the fidelity-optimized copiers preserve some of the quantum superposition of the input states (as evidenced by the off-diagonal terms in the density matrices of the copies), whereas the WZ copier makes the copies purely classical mixtures when they are considered individually. To get extra information transmission by making measurements on multistate observables, one needs some quantum effects between the successive copy states, and these effects are lacking with the WZ copier.

A small, but perhaps surprising feature was that the global-fidelity-optimized copier gave better performance in the information measures than the local-fidelity-optimized one, even though only information flowing to one copy was considered. Other features seen include the poor performance of the UQCM relative to the other copiers—unsurprising, since the other ones are tailored specifically to the two signal states, and the poorer performance when the copies are made unentangled with each other.

For all copiers considered, when the input signal states are nonorthogonal, the information carrying capacity of a channel between two observers is significantly greater when the receiver gets undisturbed states (I_H^o) than when the receiver gets one copy, even when the copier is highly optimized (I_H^u). This is an information-theoretic manifestation of the no-cloning theorem.

APPENDIX A: DERIVATION OF THE ULTIMATE-INFORMATION OPTIMAL COPIER

The copier sought has the following properties: it takes one of two ($i=1,2$) pure input states [25] $\hat{\rho}_i^A$ of Hilbert space dimension 2, and by a unitary transformation creates a state $\hat{\rho}_i$ consisting of two (possibly entangled) copies ($\hat{\rho}_i^o$ and $\hat{\rho}_i^c$), again of Hilbert space dimension 2. The state of each copy, when the other copy is ignored, is identical, and both possible copy states (corresponding to input states) have equal purity, as measured by their self-fidelity $\mathrm{Tr}[\hat{\rho}^2]$. Assuming all states considered are normalized, these conditions can be written as

$$\text{normalization:} \quad \mathrm{Tr}[\hat{\rho}_i^A]=1, \tag{A1}$$

$$\text{input pure:} \quad \hat{\rho}_i^A=|\psi_i^A\rangle\langle\psi_i^A|, \tag{A2}$$

$$\text{unitarity:} \quad \hat{\rho}_i=|\psi_i\rangle\langle\psi_i|, \tag{A3}$$

$$\mathrm{Tr}[\hat{\rho}_1\hat{\rho}_2]=\mathrm{Tr}[\hat{\rho}_1^A\hat{\rho}_2^A]=f, \tag{A4}$$

$$\text{symmetry:} \quad \hat{\rho}_i^o=\mathrm{Tr}_c[\hat{\rho}_i]=\hat{\rho}_i^c=\mathrm{Tr}_o[\hat{\rho}_i]=\hat{\rho}_i^B, \tag{A5}$$

$$\text{equal purity:} \quad \mathrm{Tr}[(\hat{\rho}_1^B)^2]=\mathrm{Tr}[(\hat{\rho}_2^B)^2]. \tag{A6}$$

And, of course, on top of these conditions, the Holevo bound on ultimate information copied I_H is to be maximized.

The output states can be written in terms of a vector of complex expansion coefficients in some basis as

$$|\psi_j\rangle=\frac{1}{\sqrt{2}}[\alpha_j,\beta_j e^{i\phi_{\beta j}},\gamma_j e^{i\phi_{\gamma j}},\delta_j e^{i\phi_{\delta j}}], \tag{A7}$$

where $\alpha_j,\beta_j,\gamma_j,\delta_j\in[0,\sqrt{2}]$, and the angles $\phi_{\ldots}\in[0,2\pi)$. Normalization gives $\alpha_j^2+\beta_j^2+\gamma_j^2+\delta_j^2=2$. One of the expansion coefficients can be made real and positive, without affecting the final bound, by multiplying by appropriate unphysical phase factors, so let us do this to the α_j.

Now, any two states in a two-dimensional Hilbert space (such as the reduced states of the two possible copies $\hat{\rho}_1^B$ and $\hat{\rho}_2^B$), can be described by two Bloch vectors $\mathbf{r}_i$. The states are then given by

$$\hat{\rho}_i(\mathbf{r}_1)=\tfrac{1}{2}(\hat{I}+\boldsymbol{\sigma}\cdot\mathbf{r}_i) \quad \text{where} \quad \boldsymbol{\sigma}=[\hat{\sigma}_1,\hat{\sigma}_2,\hat{\sigma}_3] \tag{A8}$$

and $\hat{\sigma}_j$ are the Pauli matrices. By an appropriate choice of basis, one of the two Bloch vectors can be chosen to lie in an arbitrary direction, while the other is separated by some angle ϕ_r from the first, both of them lying in a plane of our choosing. Thus there are only three parameters for these two

states that are not arbitrary, depending on the choice of basis: the lengths of the Bloch vectors r_i, and the angle between them ϕ_r. Also, since

$$\mathrm{Tr}[\hat{\rho}_i(\mathbf{r}_i)^2]=\tfrac{1}{2}(1+|\mathbf{r}_i|^2), \tag{A9}$$

and we are assuming equal copy purity (A6), both Bloch vectors are of equal length $r=|\mathbf{r}_i|$. Let us choose these Bloch vectors to be

$$\mathbf{r}_1=r[0,0,1] \quad \text{and} \quad \mathbf{r}_2=r[\sin\phi_r,0,\cos\phi_r]. \tag{A10}$$

Thus, without any loss of generality, the copies can be written in an appropriate basis as

$$\hat{\rho}_1^B=\frac{1}{2}\begin{pmatrix}1+r & 0\\ 0 & 1-r\end{pmatrix}, \tag{A11a}$$

$$\hat{\rho}_2^B=\frac{1}{2}\begin{pmatrix}1+r\cos\phi_r & r\sin\phi_r\\ r\sin\phi_r & 1-r\cos\phi_r\end{pmatrix}. \tag{A11b}$$

Using Eqs. (A11),(A7), and conditions (A1),(A5), one obtains the following restrictions on the expansion coefficients of the total output states $\hat{\rho}_i$:

$$\gamma_1=\beta_1, \quad \gamma_2=\beta_2,$$

$$\beta_1^2=1+r-\alpha_1^2, \quad \beta_2^2=1+r\cos\phi_r-\alpha_2^2, \tag{A12a}$$

$$\delta_1^2=\alpha_1^2-2r, \quad \delta_2^2=\alpha_2^2-2r\cos\phi_r,$$

$$\beta_1(\alpha_1 e^{i\phi_{\beta1}}+\delta_1 e^{\phi_{\delta1}-\phi_{\beta1}})=0, \tag{A12b}$$

$$\beta_2(\alpha_2 e^{i\phi_{\beta2}}+\delta_2 e^{\phi_{\delta2}-\phi_{\beta2}})=r\sin\phi_r. \tag{A12c}$$

Now Eq. (A12b) implies that either $\beta_1=0$ or ($\alpha_1=\delta_1$ and $2\phi_{\beta1}=\phi_{\delta1}+\pi$). The second possibility is uninteresting, as it immediately leads to $r=0$, which gives $I_H=0$—certainly not the optimum case, one hopes!

Also, using the unitarity condition (A4) and the equal purity condition (A6), one obtains the restrictions

$$2f=x+r(r-1)\cos\phi_r+C\sqrt{1-r^2}\sqrt{x(x-2r\cos\phi_r)}, \tag{A13}$$

$$r^2(1-\cos^2\phi_r)=2(1+r\cos\phi_r-x)[x-r\cos\phi_r+K\sqrt{x(x-2r\cos\phi_r)}], \tag{A14}$$

respectively. For brevity, the mutually independent parameters x,K,C have been introduced, where

$$x=\alpha_2^2, \tag{A15a}$$

$$K=\cos(\phi_{\beta2}+\phi_{\gamma2}-\phi_{\delta2}), \tag{A15b}$$

$$C=\cos(\phi_{\gamma2}-\phi_{\gamma1}). \tag{A15c}$$

Note that the condition (A14) is equivalent to Eq. (A12c).

Using Eqs. (8), (A8), and (A11) leads to I_H being given by the expression

$$I_H=\tfrac{1}{2}[(1+r)\log_2(1+r)+(1-r)\log_2(1-r)]-\tfrac{1}{2}[(1+q_H)\log_2(1+q_H)+(1-q_H)\log_2(1-q_H)], \tag{A16}$$

with

$$q_H=r\cos\frac{\phi_r}{2}. \tag{A17}$$

One finds that $I_H(r,\cos\phi_r)$ is a monotonically decreasing function of $\cos\phi_r$—thus, to maximize I_H for a given value of $r=r_o$, it suffices to minimize $\cos\phi_r$ (i.e., make the angle between the possible copy Bloch vectors as close to π as possible). $I_H(r,\cos\phi_r)$ is also a monotonically increasing function of r.

For any particular values of r and $\cos\phi_r$, there are three parameters left to vary to try to satisfy Eqs. (A13) and (A14), after the relations (A12) have been used: x, K, and C. Each of the two Eqs. (A13),(A14) will give an allowable range for x (exactly which point in these ranges is satisfied by the copier then depends on C and K). The ends of these ranges are given by

$$\frac{\partial\cos\phi_r}{\partial C}=0 \quad \text{or} \quad C=\pm1 \tag{A18a}$$

for Eq. (A13), and

$$\frac{\partial\cos\phi_r}{\partial K}=0 \quad \text{or} \quad K=\pm1 \tag{A18b}$$

for Eq. (A14). Only those values of $\cos\phi_r$ for which the two x ranges partially overlap give allowable copiers. Now, for any particular $r=r_o$, if we vary $\cos\phi_r$, the x ranges will vary also. In particular, at that value of $\cos\phi_r$ which lies at the boundary of allowed $\cos\phi_r(r_o)$ values, at least one extremity of the first x range, due to Eq. (A13), will coincide with an extremity of the second x range due to Eq. (A14). Of course, not all cases where x range extremities coincide will correspond to a $\cos\phi_r(r_o)$ extremity, but any parameters for which such x extremities coincide will give viable copiers [they could be well within a region of allowed $\cos\phi_r(r_o)$ values]. Hence, if we look at all the parameters [given by Eqs. (A13), (A14), and (A18)] where x range extremities occur, then one of them will give the desired minimum $\cos\phi_r(r_o)$ value. It turns out that this $\cos\phi_r(r_o)$ minimum corresponds to $K=C=1$ when $r\in[\sqrt{1-f},1]$. For $r<\sqrt{1-f}$, $\cos\phi_r(r_o)$ can reach its absolute minimum value of -1, but since I_H is also monotonically increasing in r, the optimum I_H copier must have $r\geqslant\sqrt{1-f}$, so these low values of r can be ignored. This leads to the second largest real root of polynomial (22) as the expression for $\cos\phi_r(r)$ that maximizes I_H for a given $r\geqslant\sqrt{1-f}$. The final value of r which maximizes I_H out of all the copiers considered, r_H, is given now by a straightforward, one-parameter maximization of $I_H(r,\cos\phi_r(r))$ over $r\in[\sqrt{1-f},1]$. Because this calculation is simple, straightforward, and accurate numerically, but not so simple analytically, an analytical solution has not been attempted.

Now, to find the particular transformation which, given input states (10), not only maximizes I_H but also makes the local copy-original fidelity as large as possible, first make the Bloch vectors of the copies be in the same plane as the Bloch vectors of the input states, and then make both pairs symmetric about a common axis. The Bloch vectors of the input states are

$$\mathbf{s}_1=[\sqrt{f},0,\sqrt{1-f}] \quad \text{and} \quad \mathbf{s}_2=[\sqrt{f},0,-\sqrt{1-f}]. \tag{A19}$$

These are in the $(\hat{\sigma}_1\text{-}\hat{\sigma}_3)$ plane, and symmetrically spaced about [1,0,0]. So, to achieve the desired optimum local fidelity copier, the appropriate transformation of the input states is found to be

$$|\psi_i^A\rangle \rightarrow (U_H \otimes U_H)|\psi_i\rangle, \tag{A20a}$$

where $|\psi_i\rangle$ is given by Eq. (A7), and the unitary transformations are

$$U_H=\begin{pmatrix} \cos\xi_H & \sin\xi_H \\ -\sin\xi_H & \cos\xi_H \end{pmatrix} \quad \text{where} \quad \xi_H=\frac{\phi_r(r_H)-\pi}{4}. \tag{A20b}$$

This can be written as Eqs. (18) and (19).

APPENDIX B: DERIVATION OF UNENTANGLED OPTIMAL COPIER

Consider copiers producing product states of the copies. This transformation can be written

$$\hat{\rho}_i^A \rightarrow \hat{\rho}_i^B \otimes \hat{\rho}_i^B \otimes \hat{\rho}_i^x, \tag{B1}$$

where $\hat{\rho}_i^B$ are the copies and $\hat{\rho}_i^x$ is a helper machine state. The only other constraint on the copier is that it must be unitary, which means that traces are preserved. This immediately leads to $\hat{\rho}_i^B$ and $\hat{\rho}_i^x$ being pure because the input states are pure (via $\mathrm{Tr}[\hat{\rho}^2]$). Furthermore,

$$f=(\mathrm{Tr}[\hat{\rho}_1^B\hat{\rho}_2^B])^2\mathrm{Tr}[\hat{\rho}_1^x\hat{\rho}_2^x]=f_{12}^2 f_x, \tag{B2}$$

where f_{12} and f_x are the fidelities between, respectively, the two copy and two machine states produced after input of originals. Thus, since $f_x \leqslant 1$, it follows that $\sqrt{f} \leqslant f_{12} \leqslant 1$.

Let us start with optimizing for one-state information transfer I_1. It is easily shown that for equiprobable input states, I_1 satisfies Eq. (14a) with the distinguishability parameter given by

$$q=\sqrt{1-f_{12}}. \tag{B3}$$

This is most straightforward to show using the Bloch vectors of the copies. Since I_1 is monotonically increasing with q, it will be maximized when q is maximized. This is when $f_{12}=\sqrt{f}$.

Now let us look at I_H. For qubit copy states, this is again given by Eq. (A16), and since the copies are pure, $r=1$, and one finds $q_H=\sqrt{f_{12}}$. With $r=1$, I_H depends only on q_H, and will reach extreme values either when

$$\frac{dI_H}{dq_H}=\frac{1}{2}\log_2\left(\frac{1-q_H}{1+q_H}\right)=0, \tag{B4}$$

or at the endpoints of the q_H range: $q_H=(f^{1/4}\text{or}1)$. One sees that Eq. (B4) is only satisfied for $q_H=f_{12}=f=0$, so for general f, extreme values of I_H are reached at $f_{12}=1$ or $f_{12}=\sqrt{f}$. $f_{12}=1$ leads to $I_H=0$, so the optimal value for f_{12} is again $\sqrt{f}$. Thus the same copiers that are optimal in I_1 are also optimal in I_H.

Lastly, let us look at local fidelity. The fidelity between any two pure states is given by

$$F(\hat{\rho}_1,\hat{\rho}_2)=\tfrac{1}{2}(1+\cos\phi), \tag{B5}$$

in terms of ϕ, the angle between their Bloch vectors. To minimize the average over both possible inputs of this Bloch angle between originals and copies, we choose the Bloch vectors of the copies to lie in the same plane as the Bloch vectors of the originals, and to be symmetric about the same axis. Obviously, in this case, the local fidelity will be maximized if the Bloch angle between the copies is as similar to the Bloch angle between the originals as possible (since the Bloch angle between original and copy is half the difference between these). Since $f_{12}=\sqrt{f} \geqslant f$, this means that we want $f_{12}=\sqrt{f}$ again. Hence, the unentangled optimal copier given in Sec. IV C is optimal in all three indicators considered in this article.

Choosing Bloch vector parameters such that Eq. (B1) holds, $f_{12}=\sqrt{f}$, and local fidelity is optimized, easily leads to the copier given in Eq. (25). It is simplest to use Bloch vectors for this calculation.

APPENDIX C: SOME FIDELITY-OPTIMIZED COPIERS

This section gives a brief summary of the fidelity-optimized copiers that are compared to the information-optimized ones in Sec. V. Expressions are given in terms of f, the square overlap between the two input states. Much more detail is given in the literature.

1. The copier that optimizes the global fidelity

The quantum copying machine that optimizes the global fidelity between the combined state of both copies and a state consisting of unentangled perfect copies has been found by Bruß *et al.* [8] The copies produced are (with the help of a little algebra)

$$\hat{\rho}_1^B=\frac{1}{2}\begin{pmatrix} 1+\sqrt{\frac{1-f}{1+f}} & \frac{f+\sqrt{f}}{1+f} \\ \frac{f+\sqrt{f}}{1+f} & 1-\sqrt{\frac{1-f}{1+f}} \end{pmatrix}, \tag{C1a}$$

$$\hat{\rho}_2^B=\frac{1}{2}\begin{pmatrix} 1-\sqrt{\frac{1-f}{1+f}} & \frac{f+\sqrt{f}}{1+f} \\ \frac{f+\sqrt{f}}{1+f} & 1+\sqrt{\frac{1-f}{1+f}} \end{pmatrix}. \tag{C1b}$$

The local fidelity is [from Eq. (47) in Ref. [8]]

$$F(\hat{\rho}_i^A,\hat{\rho}^B,i)=\frac{1}{2}\left(1+\frac{(1-f)\sqrt{1+f}+f(1+\sqrt{f})}{1+f}\right), \tag{C2}$$

and the one-state copied information is given by Eq. (14a) with distinguishability parameter

$$q=\sqrt{\frac{1-f}{1+f}}. \tag{C3}$$

The ultimate copied information is given by the expression (A16), where r, the magnitude of the Bloch vectors of the copies, is in this case

$$r=\frac{\sqrt{1+f(1+2\sqrt{f})}}{1+f}, \tag{C4a}$$

and the parameter q_H is

$$q_H=\frac{f+\sqrt{f}}{1+f}. \tag{C4b}$$

2. The copier that optimizes the local fidelity

As in Appendix C 1, Bruß *et al.* have found the copier that optimizes the local fidelity between a copy and the originals [8,24]. From Eqs. (C1)–(C6), and (C12) and subsequent discussion in Ref. [8], the copies are in the states

$$\hat{\rho}_1^B=\frac{\sec 2\phi}{2}\begin{pmatrix} \cos 2\phi+\sqrt{1-f} & (1+\sqrt{f})\sin 2\phi \\ (1+\sqrt{f})\sin 2\phi & \cos 2\phi-\sqrt{1-f} \end{pmatrix}, \tag{C5a}$$

$$\hat{\rho}_2^B=\frac{\sec 2\phi}{2}\begin{pmatrix} \cos 2\phi-\sqrt{1-f} & (1+\sqrt{f})\sin 2\phi \\ (1+\sqrt{f})\sin 2\phi & \cos 2\phi+\sqrt{1-f} \end{pmatrix}, \tag{C5b}$$

where the angle ϕ is defined by

$$\sin 2\phi=\frac{\sqrt{f}-1+\sqrt{1-2\sqrt{f}+9f}}{4\sqrt{f}}. \tag{C5c}$$

The local fidelity is [rearranging Eq. (C11) of Ref. [8]]

$$F(\hat{\rho}_i^A,\hat{\rho}_i^B)=\tfrac{1}{2}\{1+\cos 2\phi[1-f+\sqrt{f}(1+\sqrt{f})\sin 2\phi]\}. \tag{C6}$$

After some algebra, one finds that

$$q=\sqrt{1-f}\cos 2\phi, \tag{C7a}$$

$$r=\cos 2\phi\sqrt{1-f+(1+\sqrt{f})^2\sin^2 2\phi}, \tag{C7b}$$

$$q_H=\sin 2\phi\cos 2\phi(1+\sqrt{f}), \tag{C7c}$$

which can be used in expressions (14a) and (A16), respectively, to find I_1 and I_H.

3. The UQCM

The universal quantum copying machine [3] copies any two-dimensional input states with an equal, optimal, local fidelity of 5/6. This copier is unique among those mentioned in this article, in that it uses a machine helper state which becomes entangled with both copies after the process is complete. Given the input states (10) used in this article, the UQCM will create the copies

$$\hat{\rho}_1^B=\frac{1}{6}\begin{pmatrix} 3+\sqrt{1-f} & 2\sqrt{f} \\ 2\sqrt{f} & 3-2\sqrt{1-f} \end{pmatrix}, \tag{C8a}$$

$$\hat{\rho}_1^B=\frac{1}{6}\begin{pmatrix} 3-\sqrt{1-f} & 2\sqrt{f} \\ 2\sqrt{f} & 3+2\sqrt{1-f} \end{pmatrix}. \tag{C8b}$$

To calculate I_1 and I_H, use

$$q=\tfrac{2}{3}\sqrt{1-f}, \tag{C9a}$$

$$r=\tfrac{2}{3}, \tag{C9b}$$

$$q_H=\tfrac{2}{3}\sqrt{f} \tag{C9c}$$

in expressions (14a) and (A16).

[1] W. K. Wootters and W. H. Zurek, Nature (London) **299**, 802 (1982).
[2] D. Dieks, Phys. Lett. **92A**, 271 (1982).
[3] V. Bužek and M. Hillery, Phys. Rev. A **54**, 1844 (1996).
[4] R. Jozsa, J. Mod. Opt. **41**, 2315 (1994).
[5] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher, Phys. Rev. Lett. **76**, 2818 (1996).
[6] N. G. B. Huttner, Phys. Lett. A **228**, 13 (1997).
[7] N. Gisin and S. Massar, Phys. Rev. Lett. **79**, 2153 (1997).
[8] D. Bruß , D. DiVincenzo, A. Ekert, C. A. Fuchs, C. Macchiavello, and J. A. Smolin, Phys. Rev. A **57**, 2368 (1998).
[9] N. Gisin, Phys. Lett. A **242**, 1 (1998).
[10] Lu-Ming Duan and Guang-Can Guo, Phys. Lett. A **243**, 261 (1998).
[11] N. J. Cerf, Acta Phys. Slov. **48**, 115 (1998).
[12] V. Bužek, M. Hillery, and B. Rednik, Acta Phys. Slov. **48**, 177 (1998).
[13] R. F. Werner, Phys. Rev. A **58**, 1827 (1998).
[14] V. Bužek, M. Hillery, and P. L. Knight, Fortschr. Phys. **46**, 521 (1998).
[15] M. Murao, D. Jonathan, M. B. Plenio, and V. Vedral, Phys. Rev. A **59**, 156 (1999).

[16] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948); **27**, 623 (1948).

[17] A measurement scheme can in general be represented by a set of positive operators $\{\hat{A}_i\}$ which satisfy $\Sigma_i \hat{A}_i = \hat{I}$, and the probability of obtaining measurement result x_i from a state $\hat{\rho}$ is $\mathrm{Tr}[\hat{A}_i \hat{\rho}]$ [19].

[18] A. S. Holevo, Probl. Inf. Transm. **9**, 110 (1973).

[19] C. M. Caves and P. D. Drummond, Rev. Mod. Phys. **66**, 481 (1994).

[20] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W. K. Wootters, Phys. Rev. A **54**, 1869 (1996).

[21] A. S. Holevo, IEEE Trans. Inf. Theory **44**, 269 (1998).

[22] Lu-Ming Duang and Guang-Can Guo, Phys. Rev. Lett. **80**, 4999 (1998).

[23] If $r_m = 1$, use the largest real root, and if $r_m = 0$, ϕ_m is of no physical significance and thus can take any value, all of which lead to an information content of zero, anyway.

[24] C. A. Fuchs, Fortschr. Phys. **46**, 535 (1998).

[25] As well as a ''dummy'' input state of Hilbert space dimension 2, which is always in the same state initially. This is needed to preserve unitarity of the transformation.