Atemporal diagrams for quantum circuits

Robert B. Griffiths, Shengjun Wu,* and Li Yu
Physics Department, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213, USA

Scott M. Cohen
Physics Department, Duquesne University, Pittsburgh, Pennsylvania 15282, USA
(Received 24 August 2005; published 11 May 2006)

A system of diagrams is introduced that allows the representation of various elements of a quantum circuit, including measurements, in a form which makes no reference to time (hence "atemporal"). It can be used to relate quantum dynamical properties to those of entangled states (map-state duality), and suggests useful analogies, such as the inverse of an entangled ket. Diagrams clarify the role of channel kets, transition operators, dynamical operators (matrices), and Kraus rank for noisy quantum channels. Positive (semidefinite) operators are represented by diagrams with a symmetry that aids in understanding their connection with completely positive maps. The diagrams are used to analyze standard teleportation and dense coding, and for a careful study of unambiguous (conclusive) teleportation. A simple diagrammatic argument shows that a Kraus rank of 3 is impossible for a one-qubit channel modeled using a one-qubit environment in a mixed state.

DOI: 10.1103/PhysRevA.73.052309 PACS number(s): 03.67.Mn, 03.65.Ud

I. INTRODUCTION

In some sense the problem of entangled states and the problem of noisy channels, two central issues in quantum information theory, are one and the same problem. Parallels between them have been well known for some time, at least to those who know them. For an extensive and extremely helpful discussion of this duality between quantum maps and quantum states, with copious references to earlier literature, see [1]. The present paper continues an effort, begun with the study of channel kets in [2], to systematize this relationship in a manner which, as far as is practical, avoids making reference to a particular choice of basis, thus using properties of operators rather than their (basis-dependent) matrices. The aim is to make these methods and the corresponding point of view more accessible to those less familiar with their possibilities, and to classify different problems of quantum information theory, whether solved or unsolved, in a systematic way which reveals underlying connections. In [2] a classification scheme was proposed using properties of an entangled system at a single time, for reasons there discussed. The present paper, following this same motivation, is devoted to procedures for reducing quantum circuits, together with "preparation," "measurement," and (at least to some extent) "classical communication," to a form in which time plays no role, using a system of what we call *atemporal diagrams*.

The atemporal form of a particular problem is not necessarily the one which yields the best physical intuition. Entanglement is not an easier concept than transmission of quantum information, and we often find that a good way of extracting physical insight from an atemporal diagram is to think of it as some sort of quantum channel. The diagrams do, however, provide a precise and systematic notation for a variety of things for which Dirac notation, despite its many advantages, is not altogether ideal, and this without introducing a direction of time, something always present in quantum circuit diagrams. Often an atemporal diagram will suggest an analogy between a problem stated in terms of how a system develops in time and another which refers to properties of a quantum system at a single time. While some of these analogies are already known (e.g., a noiseless quantum channel is "like" a fully entangled state), others are not so obvious. In addition, the diagrams can speed up analysis of some situations by allowing one to say, "Well, that is obvious," without engaging in lengthy reasoning or complicated algebra. To be sure, the effectiveness of diagrams of this sort is to some extent subjective. We ourselves have found the scheme presented in this paper to be extremely helpful, and we hope other members of the community will share our enthusiasm or, better yet, come up with a superior, more powerful system.

Section II defines our notation for Hilbert spaces and explains the rules for constructing diagrams, together with a small number of examples. These include diagrams for positive (semidefinite) operators, which exhibit a particular symmetry, "transposers" which change kets into maps and vice versa, and the notion of the inverse of an entangled ket.

Applications begin in Sec. III A with a discussion of how atemporal diagrams, and in particular the notion of the inverse of an entangled ket, can be used to "solve" certain exercises in quantum circuits. The (by now) standard protocols for teleportation and dense coding are the subject of diagrammatic analysis in Sec. III B.

In Sec. IV atemporal diagrams are applied to the problem of unambiguous or conclusive teleportation, in which a partially entangled pure state replaces a fully entangled state as the shared resource, and the agents know whether the process succeeds or fails. Earlier work on this problem, summarized in Sec. IV E, can be unified and generalized in a very systematic way by using the notion of the inverse of an en-

*Present address: Hefei National Laboratory for Physical Science at the Microscale, University of Science and Technology of China, Hefei, Anhui 230026, People's Republic of China.

tangled ket. (Somewhat analogous results on dense coding, which also make use of map-state duality but not atemporal diagrams, are the subject of a different publication [3].)

Noisy quantum channels are the topic of Sec. V. In Sec. V A diagrams corresponding to a channel ket, transition operator, and dynamical operator are used to further illuminate the meanings of objects named in [2]. (The "dynamical operator" corresponds to the "dynamical matrix" of [1].) In particular, certain positivity conditions become self-evident because of the symmetry of the diagrams, and this topic is pursued further in Sec. V B with a discussion of completely positive maps. With the help of diagrams one sees that the Kraus rank of a noisy quantum channel is the same as the (ordinary) rank of a suitable "cross operator." This fact is used in Sec. V C to construct a straightforward, simple proof that the Kraus rank of a noisy one-qubit channel modeled using a one-qubit environment in a mixed state cannot have the value 3. Numerical evidence for this was reported in [4]; however, ours is an analytic proof.

Section VI contains a summary and notes some open questions. Appendix A, using a mathematical result derived in Appendix B, addresses some technical points about the unambiguous operations employed in Sec. IV.

II. DIAGRAMS

A. Hilbert spaces

Our notation treats bras, kets, operators, entangled states, superoperators, and the like in a uniform manner: each is an element of a suitable Hilbert space: a complex linear space on which an inner product is defined. Since we will be dealing with many such spaces, part of the notational problem is making clear precisely which space it is to which a particular entity or "object" belongs. We assume that all the Hilbert spaces of interest to us are finite.

Just as in Dirac notation, where a ket $|\psi\rangle$ and a bra $\langle\psi|$ mean different things, we find it useful to distinguish a Hilbert space $\mathcal{H}$ inhabited by kets from the dual space $\mathcal{H}^\dagger$ inhabited by bras. Lowercase Roman letter subscripts are used to distinguish Hilbert spaces for different systems, and we adopt the abbreviations

$$\mathcal{H}_{ab} = \mathcal{H}_a \otimes \mathcal{H}_b, \quad \hat{\mathcal{H}}_a = \mathcal{H}_a \otimes \mathcal{H}_a^\dagger, \quad \mathcal{L}_{ba} = \mathcal{H}_b \otimes \mathcal{H}_a^\dagger \quad (1)$$

for the *tensor product* of two spaces (sums of dyads $|a\rangle \otimes |b\rangle$), the space of *operators* (sums of dyads $|a\rangle\langle a'|$), and the space of *linear maps* from $\mathcal{H}_a$ to $\mathcal{H}_b$ (sums of dyads $|b\rangle\langle a|$). Omitting the $\otimes$ symbols in (1) and similar expressions will cause no harm. Note that $\mathcal{H}_{ab}=\mathcal{H}_{ba}$, but the order of the subscripts in $\mathcal{L}_{ba}$ is important: it is that of the dyad $|b\rangle\langle a|$, and also that of the subscripts in a matrix representing such a map. Each of these *compound* spaces, and others like them, is itself a Hilbert space, with addition and scalar multiplication defined in the obvious way, and a scalar inner product which for two elements $|a\rangle\otimes|b\rangle\otimes\langle c|$ and $|a'\rangle \otimes |b'\rangle\otimes\langle c'|$ in $\mathcal{H}_{ab}\otimes\mathcal{H}_c^\dagger$, to take a specific example, would be $\langle a|a'\rangle\cdot\langle b|b'\rangle\cdot\langle c'|c\rangle$. The *dimension* of a Hilbert space $\mathcal{H}_x$ will be denoted by d_x, and of course $d_{ab}=d_a\times d_b$.

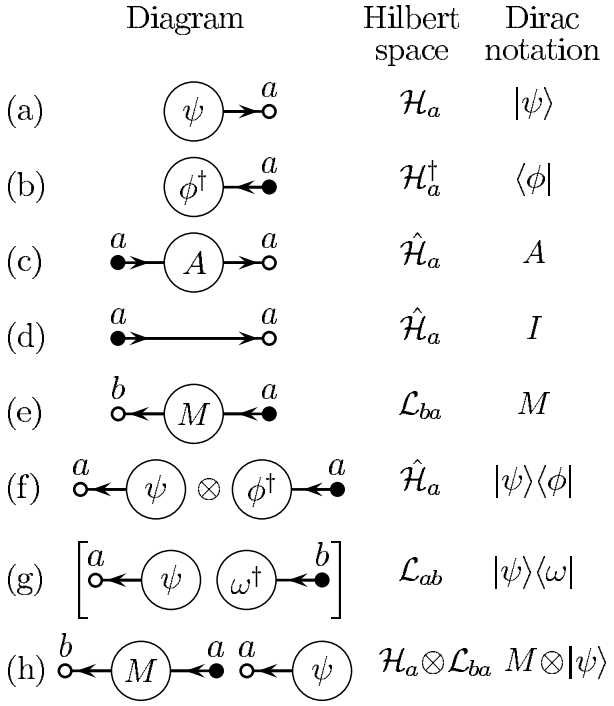

FIG. 1. Examples of simple diagrams (a)–(e) and product diagrams (f)–(h).

B. Examples of diagrams

Each "object" which is an element of some Hilbert space can be denoted by a diagram in which a symbol representing it is placed inside a circle or square or some other shape forming the *center*, connected by a certain number of *legs*, straight or curved lines, to *nodes* which specify the relevant Hilbert space. Several examples are shown Fig. 1. A node is *open* (○) or *closed* (●) depending on whether it refers to the Hilbert space or its dual, and is identified by a letter placed near it. Sometimes we shall refer to these as *active* nodes, in contrast to the inactive nodes present on internal lines of a diagram following contraction; the latter are generally not shown on the diagram (see Sec. II C). In addition to the nodes, there are arrows on the legs pointing outward from the center towards open nodes, and inward from closed nodes towards the center. The identity operator I is indicated by a single line, see (d), without a center; this notation will be justified in Sec. II C.

The Dirac notation, shown in the third column of Fig. 1, resembles that in the diagrams, with an important exception. The bra symbol $\langle\phi|$ corresponds to an object with the label $\phi^\dagger$ rather than ϕ; were the object labeled ϕ, the bra symbol would instead be $\langle\phi^\dagger|$. One can think of the Dirac $\langle|$ as effectively placing a dagger on the symbol it contains.

There is no rule that prescribes the orientations of the legs emerging from the center of an object; these may be chosen as convenient, as illustrated in the opposite orientations shown in parts (c) and (e) of the figure. Of course, if there are several legs, restricting oneself to rigid rotations will make visual identification simpler if the object is used more than once (e.g., the box labeled V in Fig. 11 below). But this

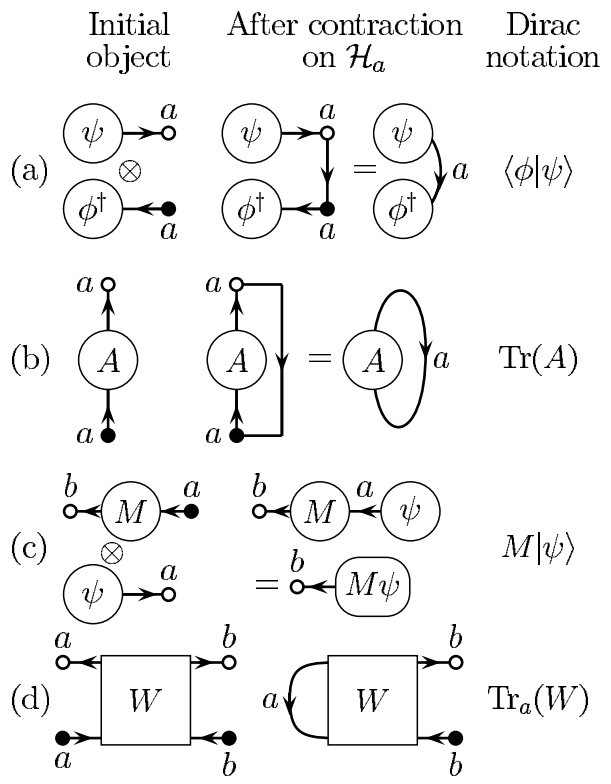

FIG. 2. Examples illustrating contraction.

is not an absolute rule, as the symbol in the center and the labels on the nodes or legs remove any ambiguity.

Objects which are tensor products of entities on distinct Hilbert spaces can be drawn as disconnected pieces, illustrated in parts (f)–(h) of Fig. 1. Sometimes it may be helpful to emphasize that different disconnected pieces belong to the same diagram by placing a $\otimes$ between adjacent pieces (f), or enclosing the pieces inside a large pair of enclosing brackets (g). However, there is no rule that these constructions have to be used, and they can be omitted when ambiguity is unlikely, as in (h). There is also no rule that says that tensor products must be represented as disconnected pieces; the one in (h), for example, could be represented using a single center and three legs terminating in open a and b nodes and a closed a node.

There is some redundancy in using both arrows and nodes, and one or the other could be omitted without changing the significance of a diagram. Retaining the nodes makes it easier to distinguish legs from other things in a complicated diagram. On the other hand, it is convenient to eliminate them when carrying out contractions; see below.

C. Contractions and matrices

A *contraction*, meaning an inner product, trace, or partial trace, can be carried out on any object whose diagram contains both a closed and an open node referring to the same Hilbert space, i.e., to the Hilbert space and its dual. It is performed by drawing a line connecting these two nodes, with an arrow pointing from the open to the closed node. There are several examples in Fig. 2. As the contracted object makes no reference to the Hilbert space of the contracted nodes, they can be omitted, and the three arrows pointing in the same direction replaced by just one. It is useful to place a single letter next to the remaining arrow on the contracting line to indicate on which space the contraction has been carried out. (The original nodes can be put back in, if desired, but they should then not be confused with the *active* nodes, which always terminate a single line.) This process of simplification is indicated explicitly in Fig. 2(a) and (b), but only the end result in (c) and (d). (One could also leave the two nodes joined by the contracting line and omit the arrows, but this is a bulkier notation.) Using the abbreviation just mentioned justifies the use of a single line to indicate the identity operator, Fig. 1(d), as can be seen by constructing the diagram for $I|\psi\rangle=|\psi\rangle$. Of course, both arrows and labels are sometimes omitted when the context determines what they should be. Figure 2(c) illustrates still another abbreviation which is often, but not always, convenient. Two centers connected by a single contraction line can be reduced to a single center containing the symbol corresponding to the head of the arrow to the left of that corresponding to the tail. The order of these symbols is important, and (usually) agrees with standard notation.

To obtain the *matrix* associated with some object, it is necessary to specify a basis or bases. We limit ourselves to orthonormal bases, although the same diagrammatic approach will work for more general choices. Examples are shown in Fig. 3, together with the corresponding Dirac notation for the matrix. Note that column vectors, row vectors, and objects labeled by three or more indices are included in the scheme, and the elements of a rectangular matrix may turn out to be operators or matrices, as in (d). The Dirac notation for items (c) and (d), while appropriate, can sometimes mislead; e.g., $\langle a_j|Y|c_l\rangle$ denotes a ket, not a complex number.

D. Adjoints and positive operators

Given any object O on a tensor product of Hilbert spaces, its adjoint $O^\dagger$ is defined in the usual way using an antilinear map in which all bras become kets and vice versa. For example, the dagger in

$$\left[\sum c_{jkl}\Big(|a_j\rangle\otimes\langle a_k|\otimes|b_l\rangle\Big)\right]^\dagger=\sum c_{jkl}^*\Big(\langle a_j|\otimes|a_k\rangle\otimes\langle b_l|\Big) \tag{2}$$

maps an object from $\hat{\mathcal{H}}_a\mathcal{H}_b$ to $\hat{\mathcal{H}}_a\mathcal{H}_b^\dagger$. Normally when using Dirac notation one would omit the first $\otimes$ on both the left and right sides of this equation, by using $|a_j\rangle\langle a_k|$ and $|a_k\rangle\langle a_j|$, respectively. The rule for diagrams is illustrated in Fig. 4(a), where the right and left sides are adjoints of each other: change every open node to a closed node and vice versa while keeping the node labels the same, reverse the direction of every arrow, and place the adjoint superscript † on the symbol appearing in each center, unless there is one already there, in which case remove it. Note the usual rules, that when † is applied twice to any symbol one gets the symbol back again, and if a symbol is a product, as $M\psi$ in the ab-

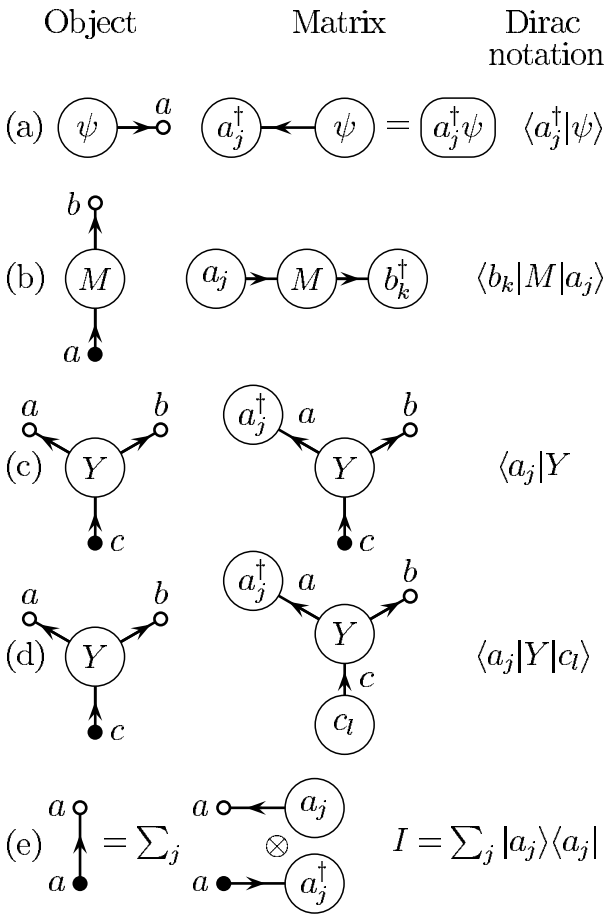

FIG. 3. Parts (a)–(d) are examples of matrix elements of different objects, together with the corresponding Dirac notation; (e) is the diagrammatic form of I as a sum of dyads.

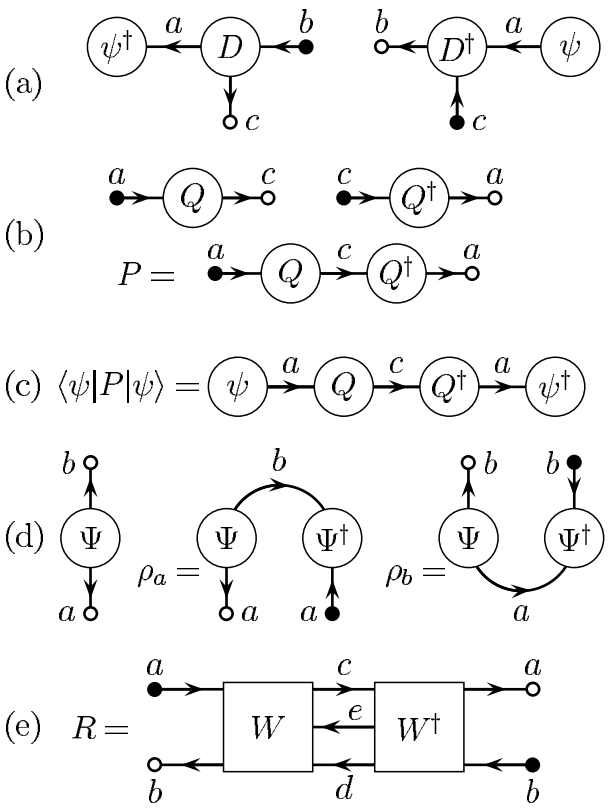

FIG. 4. (a) The two objects are adjoints of each other. (b) A positive operator can be formed by contraction of an object with its adjoint. (c) Diagram symmetry illustrates positivity criterion for P. (d) Reduced density operators for a bipartite ket. (e) Positivity of R follows from symmetry of diagram.

breviated form in Fig. 2(c), one should reverse the order: $(M\psi)^\dagger=\psi^\dagger M^\dagger$. In Fig. 4(a), the adjoint diagram has also been reflected about a vertical axis. While this is often a convenient thing to do, especially when one intends to carry out a contraction to produce a positive operator (see below), it is *not* essential: the significance of any diagram does not depend upon its orientation or the order of the legs (assuming they are properly labeled).

By a *positive operator* on a Hilbert space $\mathcal{H}$ we mean a Hermitian operator in $\hat{\mathcal{H}}$ whose eigenvalues are all nonnegative. The longer term "positive semidefinite" is more precise. [It is important to distinguish a positive operator in the sense just defined from a *positive* or *completely positive superoperator*, referring to a map from a space of operators, say, $\hat{\mathcal{H}}_a$ to a different (or possibly the same) space of operators $\hat{\mathcal{H}}_b$; see Sec. V.] A positive operator P has a positive square root, and hence it can be always be written in the form

$$P=Q^\dagger Q \tag{3}$$

by setting $Q=Q^\dagger$ equal to this square root. Conversely, any operator P of the form (3) is necessarily positive, even in cases in which Q is a map of $\mathcal{H}$ to another Hilbert space $\mathcal{H}'$, in which case $Q^\dagger$ maps from $\mathcal{H}'$ back to $\mathcal{H}$.

The diagram representing the right side of (3) is formed by a contraction of the diagram for Q with that of its adjoint, Fig. 4(b), and exhibits a characteristic symmetry which is often useful for visually identifying positive operators: reflection across a particular plane changes closed to open nodes and vice versa, reverses the directions of all arrows, and applies a superscript † to all symbols. Part (c) of the figure shows how the well-known characterization of a positive operator by the condition that $\langle\psi|P|\psi\rangle$ be positive finds a natural expression in diagrammatic terms when P is of the form (3). Diagrams corresponding to the reduced density operators

$$\rho_a=\mathrm{Tr}_b(|\Psi\rangle\langle\Psi|),\quad \rho_b=\mathrm{Tr}_a(|\Psi\rangle\langle\Psi|) \tag{4}$$

of an entangled ket $|\Psi\rangle\in\mathcal{H}_{ab}$ are shown in (d). Finally, a more complicated example is given in (e), where the structure of the diagram shows that R is a positive operator on $\mathcal{L}_{ba}$, i.e., as a map of the linear space of $\mathcal{H}_a$-to-$\mathcal{H}_b$ maps onto itself. (Think of R as "operating" through contraction of the left pair of nodes with something in $\mathcal{L}_{ba}$ in order to produce something else in this space, corresponding to the right pair of nodes.) The symmetry which shows it to be positive is reflection about a line bisecting the c, d, and e lines; note that not all of the arrows need go in the same direction. (One may

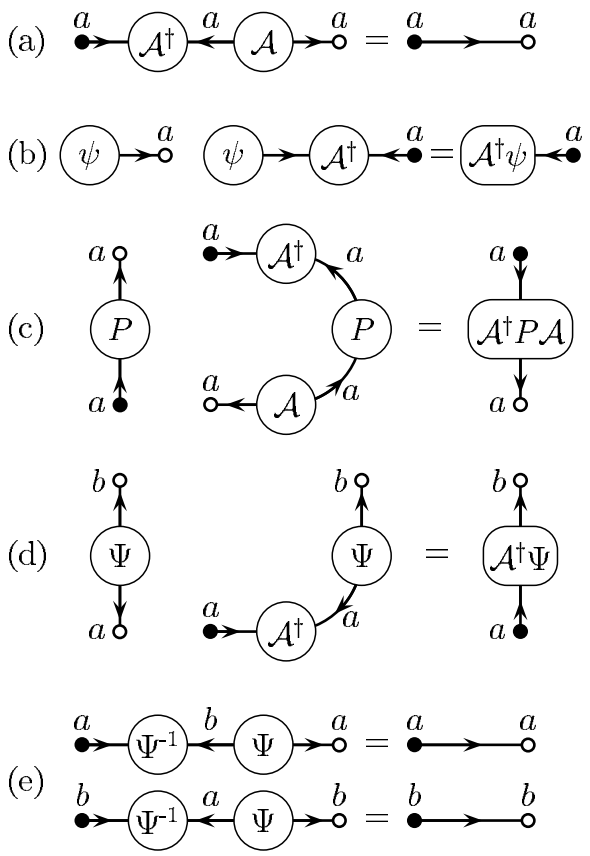

FIG. 5. (a) The adjoint of a transposer is its inverse. (b)–(d) Transposers change open to closed nodes, or vice versa. (e) Inverse of an entangled ket.

also think of R as a "top-to-bottom" map of $\hat{\mathcal{H}}_a$ to $\hat{\mathcal{H}}_b$. As such it is not a positive operator, since the domain and range are different, though it is, as will be discussed in Sec. V, a completely positive superoperator.)

E. Transpose, rank, inverse

Let $\{|a_j\rangle\}$ be an orthonormal basis for the space $\mathcal{H}_a$, and on the tensor product $\mathcal{H}_a \otimes \mathcal{H}_a$ of $\mathcal{H}_a$ with itself define the *transposer* $\mathcal{A}$ and its adjoint $\mathcal{A}^\dagger$ as a fully entangled ket and bra,

$$|\mathcal{A}\rangle = \sum_j |a_j\rangle \otimes |a_j\rangle, \quad \langle \mathcal{A}| = \sum_j \langle a_j| \otimes \langle a_j|, \tag{5}$$

with normalization

$$\langle \mathcal{A}|\mathcal{A}\rangle = d_a. \tag{6}$$

[Readers who prefer to distinguish the two copies of $\mathcal{H}_a$ may think of $\mathcal{A}$ as defined on $\mathcal{H}_a \otimes \mathcal{H}_{a'}$, change the second a_j to a'_j in each of the expressions in (5), and make appropriate modifications in Fig. 5.]

Figure 5(a) shows that $\mathcal{A}^\dagger$ is the inverse of $\mathcal{A}$, which makes $|\mathcal{A}\rangle$ the analog of a unitary map. Parts (b)–(d) show how a transposer can be used to convert open to closed nodes and vice versa. In part (b) the transposer applied to a ket ψ yields a bra $\mathcal{A}^\dagger\psi$ whose row vector in the basis $\{|a_j\rangle\}$ is the transpose, in the usual sense of that term (without complex conjugation), of the column vector corresponding to the ket in this same basis. The transpose of an operator P, part (c), requires the use of both $\mathcal{A}$ and $\mathcal{A}^\dagger$, and the result can be denoted $\mathcal{A}^\dagger P \mathcal{A}$, following the rule given in Sec. II C. The conventional notation P^T is more compact, but carries less information, since it does not show the basis dependence of the resulting operator.

Part (d) of the figure shows how a transposer can change an entangled ket $\Psi \in \mathcal{H}_{ab}$ into a map $\mathcal{A}^\dagger\Psi \in \mathcal{L}_{ba}$. Although such a change is not ordinarily called a "transpose," the term (or, if one prefers, "generalized transpose") seems appropriate. In Dirac notation it takes a particularly simple form if, along with the orthonormal basis $\{|a_j\rangle\}$ defining $\mathcal{A}$, one introduces a basis $\{|b_k\rangle\}$ for $\mathcal{H}_b$ and expands both Ψ and $M = \mathcal{A}^\dagger\Psi$ in these bases:

$$|\Psi\rangle = \sum_{jk} \mu_{jk} |a_j\rangle \otimes |b_k\rangle, \quad M = \langle \mathcal{A}|\Psi\rangle = \sum_{jk} \mu_{jk} |b_k\rangle\langle a_j|. \tag{7}$$

Note that the matrix μ_{jk} is the same in both expressions; all the transpose does is change $|a_j\rangle$ to $\langle a_j|$. Also note that in the case $d_a = d_b$, if $|\Psi\rangle$ is a fully entangled ket normalized so that $\langle\Psi|\Psi\rangle = d_a$, M is a unitary map, and vice versa.

The *rank* $\mathrm{Rn}(M)$ of the map M in (7) can be defined as the rank of the matrix μ_{jk} (the dimension of the space spanned by its columns). This definition depends only on M and not on the choice of bases, because the rank of a matrix is left unchanged upon left and right multiplication by nonsingular square matrices: see p. 13 in [5]. The rank of $|\Psi\rangle$ can be defined in exactly the same way, and is often called the *Schmidt rank*, which we sometimes denote by σ, because it is the number of positive Schmidt coefficients in the diagonal expansion

$$|\Psi\rangle = \sum_j \lambda_j |\bar{a}_j\rangle \otimes |\bar{b}_j\rangle, \quad \lambda_j \geqslant 0, \tag{8}$$

obtained using appropriate orthonormal bases $\{|\bar{a}_j\rangle\}$ and $\{|\bar{b}_j\rangle\}$. These Schmidt coefficients are none other than the singular values of the map M.

Continuing the list of useful parallels between maps and entangled kets, if the Schmidt rank σ of $|\Psi\rangle$ is equal to $d_a = d_b$, one can define its *inverse* using the formula

$$\langle (\Psi^{-1})^\dagger| = \sum_{jk} \nu_{kj} \langle a_j| \otimes \langle b_k| = \sum_j \lambda_j^{-1} \langle \bar{a}_j| \otimes \langle \bar{b}_j|, \tag{9}$$

where ν is the inverse of the μ matrix in (7): $\Sigma_k \mu_{jk}\nu_{kl} = \delta_{jl}$. As indicated in the diagram in Fig. 5(e), Ψ^{-1} is both the left and right inverses of Ψ. [The dagger on the left side of (9) is absent in the diagram; see the remarks accompanying Fig. 1.] Of course, Ψ is the inverse of Ψ^{-1}. Although the first expression in (9) defines the inverse using a particular basis, it is easy to show that the resulting bra does not depend on the choice of basis. In addition, it is worth noting that the inverse of $\Psi^\dagger$ is the same as the adjoint of Ψ^{-1}, so the order of minus one and dagger in $\Psi^{-1\dagger}$ does not matter. When the rank of $|\Psi\rangle$ is smaller than d_a, or if $d_a \neq d_b$, one can define a "gen-

eralized inverse" by only retaining the terms with $\lambda_j > 0$ on the right side of (9). See the discussion, for maps, in [5], p. 421. In this paper we do not need such a generalized inverse, so will not discuss it further.

In the case of an object with three or more nodes, such as Y in Fig. 3, the discussion of rank becomes more complicated, because Y can be turned into a map from one space to another in several different ways. The most natural interpretation of Y, given that the c node is closed and the a and b nodes are open, is as a map $Y_{ab;c}$ from $\mathcal{H}_c$ to $\mathcal{H}_{ab}$, and as such it has a rank which can be written as $\mathrm{Rn}(Y_{ab;c})$. But one can just as well regard Y as a map $Y_{bc;a}$ from $\mathcal{H}_a^\dagger$ to $\mathcal{H}_b\mathcal{H}_c^\dagger = \mathcal{L}_{bc}$, and as such it will have a rank $\mathrm{Rn}(Y_{bc;a})$ different (in general) from $\mathrm{Rn}(Y_{ab;c})$. Or, equivalently, $\mathrm{Rn}(Y_{bc;a})$ is the rank of $Y' = \mathcal{A}^\dagger Y \mathcal{C}$, obtained by applying transposers to the a and c nodes of Y, when it is regarded as a map $Y'_{bc;a}$ from $\mathcal{H}_a$ to $\mathcal{H}_{bc}$. One way to think about these different possibilities is to write Y in the Dirac form

$$Y = \sum_{j,k,l} \Big(\langle a_j, b_k | Y | c_l \rangle \Big) |a_j\rangle \otimes |b_k\rangle \otimes \langle c_l|. \tag{10}$$

using orthonormal bases $\{|a_j\rangle\}$, $\{|b_k\rangle\}$, $\{|c_l\rangle\}$. If one regards $\langle a_j, b_k|Y|c_l\rangle$ as a matrix with columns labeled by l and rows by the double label (j,k), its rank will be $\mathrm{Rn}(Y_{ab;c})$, whereas to obtain $\mathrm{Rn}(Y_{bc;a})$ consider it a matrix with columns labeled by j and rows by the double label (k,l). There is no reason to expect these two ranks to be equal, and of course there is a third possibility $\mathrm{Rn}(Y_{ac;b})$, which could be different from the other two. Note that our subscript notation simply divides the collection of nodes into two sets, with the object understood as a map from the space labeled by the second set to that labeled by the first. Thus $Y_{ab;c}$ and $Y_{ba;c}$ denote the same map. In addition, since transposing a matrix does not change its rank, $\mathrm{Rn}(Y_{ab;c})$ is the same as $\mathrm{Rn}(Y_{c;ab})$, even though the maps are distinct. Consequently there are only three possible ranks associated with the object Y, and it makes no difference if one changes open nodes to closed nodes or vice versa using appropriate transposers. Atemporal diagrams help one to see that a single object can be associated with several different ranks, but they do not in and of themselves place any constraints on these different possibilities.

F. The virtue of being careless

The system of diagrams introduced above is precise in the sense that the (tensor product) Hilbert space identified with an object is specified by the types, open or closed, of nodes in its diagram, along with the labels attached to them, with the space $\mathcal{H}^\dagger$ of bras carefully distinguished from the space $\mathcal{H}$ of kets, as in Dirac notation. One can convert an open to a closed node or vice versa by introducing a transposer, Sec. II E, but in this case the symbol for the transposer becomes part of the diagram, as in Figs. 5(b)–5(d), and indicates precisely which basis was employed for the transposition, a concept that is not basis independent.

While such precision is often valuable, there are circumstances, especially when one is carrying out a first, informal analysis of a problem, when it is useful to *ignore* the difference between open and closed nodes, or the direction of arrows on contractions lines, and treat two diagrams that differ only in these respects as "essentially the same." To put the matter in a different way, there are analogies between two diagrams which are "identical" in this loose sense, analogies which are worth exploring, because they might produce helpful insights.

For example, if Ψ in Fig. 5(d) is a fully entangled ket on two Hilbert spaces of the same dimension, the corresponding map $\mathcal{A}^\dagger\Psi$ is a unitary map, up to multiplication by a constant. Thus "fully entangled" and "unitary" are "essentially the same." This helps one understand, at an intuitive level, the essence of teleportation as discussed below in Sec. III, using the diagrams in Fig. 7 below, in which (c) and (d) both represent unitary channels. Another example is Schmidt rank [see the discussion in connection with (8)], which is the ordinary rank of a map which is "essentially the same" as an entangled ket. The Kraus rank discussed below in Sec. V A can in a similar way be identified with the ordinary rank of the "cross operator" V in Fig. 10(b) below, which maps the f node part into the a, b node part, without worrying about whether these nodes are open or closed; see the discussion of rank in Sec. II E. A fourth example, not illustrated in the present paper, is that of letting time flow backwards in a quantum circuit.

In some cases one can easily identify the formal equivalence that makes such diagrams "essentially the same." Thus a fully entangled state has equal Schmidt coefficients, and these correspond to the equal singular values that characterize a scalar multiple of a unitary operator. However, experience suggests that it is generally more fruitful to think about the analogy, and ask whether it provides some useful insight, than it is to work out the formal correspondence.

III. INTERCHANGE: TELEPORTATION AND DENSE CODING

A. Interchange

As a first example of how a quantum circuit can be represented by an atemporal diagram, consider the circuit equation in Fig. 6(a): Given an entangled state $|\Psi\rangle$ on $\mathcal{H}_a \otimes \mathcal{H}_b$, with $d_b = d_a = d$ and an operator A on $\mathcal{H}_a$, find, if possible, an operator B on $\mathcal{H}_b$ such that

$$(I_a \otimes B)|\Psi\rangle = (A \otimes I_b)|\Psi\rangle. \tag{11}$$

Part (b) of the figure shows the equation reduced to atemporal diagrams, and part (c) shows a strategy for solving it if $|\Psi\rangle$ is of rank d, and thus has an inverse: insert the identity in the form $\Psi^{-1}\Psi$ between the a node and A on the right side of (b), and it is evident that the final set of three objects in (c) represent the solution, as written in diagrammatic form in (d). It is not very easy to interpret the Dirac equivalent

$$B = \langle (\Psi^{-1})^\dagger | A | \Psi \rangle \tag{12}$$

apart from saying it means the same thing as (d) in the figure.

When the rank of $|\Psi\rangle$ is less than d, this strategy will not work, and in general (11) has no solution. However, a weaker version, Fig. 6(e), where the problem is to find B and

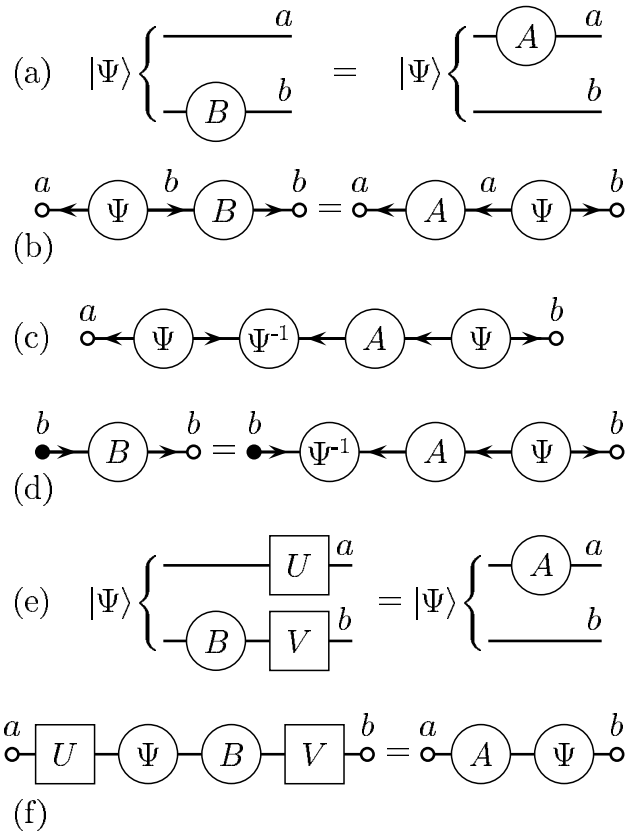

FIG. 6. The circuit equation (a), or its atemporal equivalent (b), for the unknown operator B can be solved as shown in (c) and (d) if Ψ has an inverse. Solving (e) or (f) for B and the unitaries U and V is discussed in the text.

two unitary operators U and V such that, for a given $|\Psi\rangle$ and A,

$$(U \otimes V)(I_a \otimes B)|\Psi\rangle = (A \otimes I_b)|\Psi\rangle, \tag{13}$$

can always be solved, even when one imposes the additional condition that B be unitarily equivalent to A, in the sense that there is a unitary map S from $\mathcal{H}_a$ to $\mathcal{H}_b$ such that $B = SAS^\dagger$. The strategy (see Appendix A of [6]) is to choose orthonormal bases in which $|\Psi\rangle$ has Schmidt form (8), and define

$$B = \sum_{jk} \langle \bar{a}_j|A|\bar{a}_k\rangle \cdot |\bar{b}_j\rangle\langle\bar{b}_k|, \tag{14}$$

which is obviously unitarily equivalent to A in the sense just defined: $S = \Sigma_j |\bar{b}_j\rangle\langle\bar{a}_j|$. Given this definition for B, it is evident that the two entangled kets

$$|\Psi_A\rangle = (A \otimes I_b)|\Psi\rangle, \quad |\Psi_B\rangle = (I_a \otimes B)|\Psi\rangle \tag{15}$$

are mapped into each other by interchanging the two spaces $\mathcal{H}_a$ and $\mathcal{H}_b$, a symmetry which guarantees that they possess the same set of coefficients when each is expanded in Schmidt form, relative to appropriate Schmidt bases. The "local unitaries" U and V are then those that map the Schmidt bases for $|\Psi_B\rangle$ to those of $|\Psi_A\rangle$.

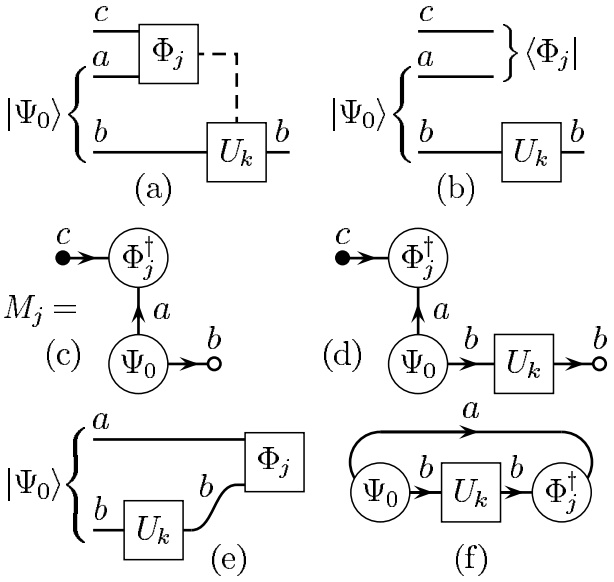

FIG. 7. Circuit diagrams (a) and (b) for teleportation, along with atemporal diagrams (c) representing M_j, and (d) the effective quantum channel from $\mathcal{H}_c$ to $\mathcal{H}_b$. The dense coding quantum circuit (e) leads to the atemporal diagram (f).

The corresponding atemporal diagram in Fig. 6(f), while not used in deriving this result, immediately suggests an extension, especially with arrows omitted (the reader can easily supply them): Interchanging two objects A and Ψ is possible at a cost of unitary equivalence and introducing additional unitaries. The objects could be two operators, or an entangled ket and an entangled bra.

B. Teleportation and dense coding

The standard teleportation arrangement shown in Fig. 7(a) employs three Hilbert spaces $\mathcal{H}_a$, $\mathcal{H}_b$, and $\mathcal{H}_c$, of identical dimension d. An initial fully entangled state $|\Psi\rangle$ on $\mathcal{H}_{ab}$, is used to teleport a state $|c\rangle$ of $\mathcal{H}_c$, by means of a measurement on $\mathcal{H}_{ac}$ in a basis $\{|\Phi_j\rangle\}$ of fully entangled states. The outcome j is used to select one of a set $\{U_k\}$ of d^2 unitary operators to apply to $\mathcal{H}_b$ in order to yield a final state equal to $V|c\rangle$, where V is some unitary map, fixed and independent of j, from $\mathcal{H}_c$ to $\mathcal{H}_a$. One usually thinks of V as the identity operator, assuming that an appropriate correspondence has been set up between these two Hilbert spaces; for a careful discussion see Sec. V of [7].

In constructing an atemporal diagram to represent teleportation in the form just discussed, we first ignore the "classical communication" step indicated by the dashed curve in part (a) of the figure—we shall return to this aspect later—and interpret the measurement outcome as indicating the prior state of the quantum system, as indicated in the circuit in part (b). (Although not yet found in elementary textbooks, this is a perfectly consistent way of viewing measurement outcomes; see Chap. 17 of [8].) If we expand the initial state

$$|\Omega\rangle = |c\rangle \otimes |\Psi\rangle = \sum_j |\Phi_j\rangle \otimes M_j|c\rangle, \tag{16}$$

in the basis $\{|\Phi_j\rangle\}$, the Kraus operator M_j can be formally expressed in Dirac notation as

$$M_j = \langle \Phi_j | \Psi \rangle, \tag{17}$$

or perhaps with greater clarity using the atemporal diagram of Fig. 7(c). When the final unitary U_k is added to the diagram, we arrive at part (d), which provides an atemporal representation of the quantum channel starting at the initial c and ending at the final b of parts (a) or (b) of the figure.

In interpreting the atemporal diagram in Fig. 7(c) it is helpful to adopt the normalization

$$\langle \Psi | \Psi \rangle = \langle \Phi_j | \Phi_j \rangle = d, \tag{18}$$

in which case M_j will be a unitary operator. That is, there is a unitary map from the c input to the b line preceding U_k in part (a) of the figure, a map which depends on the measurement outcome j. Note that one does not have to think of node b in (c) as referring to a time following that at which the measurement on $\mathcal{H}_{ac}$ occurs; the channel, in the sense of statistical correlations within an appropriate framework of histories, is present both at earlier and later times. One advantage of an *atemporal* diagram is that one reaches this conclusion without becoming entangled in misleading notions like "wave function collapse."

Once the unitary nature of (c) has been understood, the role of the final unitary in (d) becomes obvious: in order to obtain a perfect quantum channel corresponding to the identity map, the final unitary should undo the action of the preceding unitary, so that in the case of outcome j, one should employ

$$U_j = M_j^\dagger. \tag{19}$$

This means, of course, using a different unitary for each measurement outcome j, and that is why it is essential in actual teleportation protocols that its value be transmitted from the point where the measurement occurs to the point where the unitary correction will occur. While this "classical communication" plays a critical role in the protocol, it does not play an essential role in understanding what is going on in quantum mechanical terms. The virtue of the atemporal diagram is that it focuses attention on the latter, which is to say the quantum correlations whose correct calculation requires the consistent use of quantum theory. Once these correlations are taken care of, other aspects of the situation can be understood through straightforward application of ideas from classical physics.

The quantum circuit for dense coding, Fig. 7(e), has been drawn in a way to emphasize the similarity with teleportation. Initially there is a fully entangled state $|\Psi\rangle$ on $\mathcal{H}_{ab}$, and one of an appropriate collection of d^2 unitaries is applied to the $\mathcal{H}_b$ part, following which a measurement is carried out in a fully entangled basis $\{|\Phi_j\rangle\}$. [The line connecting U_k to Φ_j is drawn curved in (e) as a reminder that one typically thinks of this process as transmitting information from one physical location to another, but this is obviously not essential for understanding the quantum physics of the situation.] Upon assuming, as before, that the measurement outcome represents the prior quantum state, one arrives at the atemporal diagram in (f). [The reader may want to insert the obvious analog of part (b) of the figure as an intermediate step.]

This closed loop, as it contains no nodes, is a complex number whose absolute square can be interpreted as a probability up to a suitable normalization constant. If, as before, the normalization (18) is adopted, so that $\Phi_j^\dagger \Psi$ is a unitary operator, dividing by d^2 will yield the conditional probability of measurement outcome j given the use of the unitary U_k. The conventional choice which makes the loop equal to 0 for $j \neq k$ is obtained using (19). Since this is an atemporal diagram, one could just as well have drawn (f) with $\Phi_j^\dagger$ to the left of Ψ, whence it is evident that (f) is obtained from (d) by "closing the loop." Whereas the close connection of teleportation with dense coding has been pointed out in the past (see in particular [9]), the diagrammatic approach provides a particularly simple way of seeing the relationship.

IV. UNAMBIGUOUS (CONCLUSIVE) TELEPORTATION

A. Introduction

If the state $|\Psi\rangle$ in Fig. 7 is an entangled state of Schmidt rank $d=d_a=d_b=d_c$ but not fully entangled, ordinary teleportation of some arbitrary state is not possible, but a modified protocol, which we call *unambiguous* teleportation, will succeed with a probability less than 1. In one version, Alice carries out a POVM measurement on the initial state $|\Omega\rangle$ in (16) using a collection of positive operators $\{G_j, j=0,1,\ldots\}$ on $\mathcal{H}_{ca}$ which sum to the identity. If the outcome corresponds to G_j for some $j \geqslant 1$, the *conclusive* case, Bob applies an appropriate unitary correction U_j, and the initial state $|c\rangle$ on $\mathcal{H}_c$ has been successfully teleported to $\mathcal{H}_b$. However, if it corresponds to G_0, the *inconclusive* case, the attempted teleportation has failed in the sense that in general there is no way Bob can apply a correction which will ensure the correct output. In a different version, Alice measures $|\Omega\rangle$ in a fully entangled basis of $\mathcal{H}_{ca}$, as in the standard protocol, but Bob upon learning the value of j attempts to carry out a nonunitary correction L_j, which, if successful, completes the teleportation. The two can be combined: some general POVM by Alice and a nonunitary correction by Bob when he learns the outcome, as discussed below in Sec. IV C

The adjective "unambiguous" refers to the fact that at least one of the parties carrying out the process, who can always inform the other, is aware of whether the protocol succeeds (conclusive case) or fails (inconclusive case). While the term "conclusive teleportation" is often employed, we think "unambiguous" has a somewhat clearer meaning. Also, its use in this context agrees with the closely related idea of unambiguous discrimination—see, for example, [10]—and it is more precise than the term "conclusive" as used at present by Mor (the originator of "conclusive teleportation") and his associates: see Sec. IV E. (Other names are sometimes employed, e.g., probabilistic teleportation.) For any protocol of this type, let q_j be the probability that Alice obtains outcome j and r_j for $j \geqslant 1$ the probability that Bob's correction will succeed—if the latter is a unitary operation, $r_j=1$. Then the overall probability of success is

$$p_s = \sum_{j \geqslant 1} p_j = \sum_{j \geqslant 1} q_j r_j, \tag{20}$$

with $p_j=q_j r_j$ the probability that Alice's outcome is j and Bob's correction is successful.

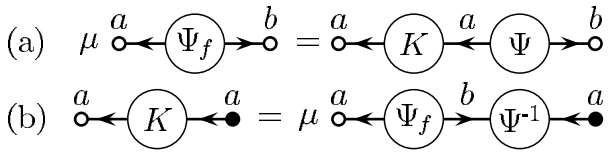

FIG. 8. Definition of operator K using Ψ^{-1}.

What might at first seem like a completely different strategy is for Alice to carry out *unambiguous entanglement concentration* on $|\Psi\rangle$ to obtain a fully entangled state $|\Psi_f\rangle$. If the process is successful, teleportation can then be carried out using the standard protocol, employing $|\Psi_f\rangle$ as the entangled resource. The optimum probability for successful unambiguous entanglement concentration is $p_c = d\lambda_m^2$, Sec. IV B, where λ_m is the minimum Schmidt coefficient in (8), and consequently it is always possible to carry out unambiguous teleportation using a two-step process with this probability of overall success. On the other hand, p_c is also an upper bound on the probability of success of *any* unambiguous teleportation protocol for, as pointed out in p. 90 of [11], Alice can generate a fully entangled pair in her laboratory and teleport half of it to Bob, thus effecting unambiguous entanglement concentration by means of unambiguous teleportation. Of course, the unambiguous entanglement concentration could equally well be carried out by Bob, with the same probability of success.

Based on this idea one can construct various protocols of the type mentioned earlier that have the maximum probability of success allowed by the partially entangled state used as a resource. Some of these are considered in Sec. IV D following an analysis in Sec. IV C of the general case of an arbitrary POVM by Alice leading to an arbitrary (in general nonunitary) correction by Bob. We have found this analysis helpful in, among other things, understanding why it is that in the case of unambiguous teleportation, just as in regular teleportation, Alice and Bob learn nothing about the nature of the teleported state, despite the fact that in a general protocol of the sort described in Sec. IV C, Alice *can* acquire statistical information about the initial state.

While our results could no doubt be obtained without using atemporal diagrams, we think they add both motivation and clarity to arguments which are more difficult to understand when expressed in purely algebraic form. Our approach serves to provide a unified perspective on a number of different results in the literature, as discussed in Sec. IV E.

B. Entanglement concentration

The essence of ordinary teleportation resides in the observation that the operator M_j in (17), with diagram in Fig. 7(c), is for every j a multiple of a unitary operator. This is no longer the case if $|\Psi\rangle$ is not fully entangled, but if it has a Schmidt rank of d, and thus an inverse in the sense discussed in Sec. II E, unitarity can be restored by inserting a suitable operator K in the a link in part (c) of Fig. 7, provided $K|\Psi\rangle$ is a constant μ times a fully entangled state $|\Psi_f\rangle$, as indicated in Fig. 8(a). The explicit form of K can be obtained using the inverse Ψ^{-1} of Ψ, as shown in Fig. 8(b). In particular, it is convenient to assume that

$$|\Psi_f\rangle = (1/\sqrt{d})\sum_j |\bar{a}_j\rangle \otimes |\bar{b}_j\rangle, \tag{21}$$

using the same bases as in (8), in which case

$$K = (\mu/\sqrt{d})\sum_j (1/\lambda_j)|\bar{a}_j\rangle\langle\bar{a}_j|. \tag{22}$$

Unambiguous entanglement concentration means that Alice carries out the operation K in an unambiguous manner, using an apparatus which clearly indicates success or failure; e.g., think of a green or red light going on. Further details and some subtleties are considered in Appendix A. The maximum probability of success (the green light going on), which is also achievable, is given by

$$p_c = \langle\Psi|K^\dagger K|\Psi\rangle = d\lambda_m^2 \tag{23}$$

when $\mu = \sqrt{d}\lambda_m$ is chosen so that the maximum eigenvalue of $K^\dagger K$ equal to 1; here λ_m is the minimum of the coefficients in the Schmidt expansion (8) for $|\Psi\rangle$. (Note that in the special case in which $|\Psi\rangle$ is fully entangled, $\lambda_m = 1/\sqrt{d}$ and the probability of success is $p_c = 1$, as expected.) Needless to say, the same thing can be achieved by placing Ψ^{-1} in the b rather than in the a link of Fig. 7(c); i.e., Bob rather than Alice can carry out unambiguous entanglement concentration, with exactly the same probability of success.

C. General teleportation protocol

Consider a protocol in which Alice carries out a POVM $\{G_j\}$, and when Bob is informed of outcome j he carries out an unambiguous operation L_j with a probability of success that may be less than 1. We shall assume that for every $j \geqslant 1$, G_j is proportional to a projector onto a pure state. There is no harm in making this assumption for a protocol which achieves the maximum possible probability of success, since any positive operator in a POVM can always be refined without reducing the amount of information available to Alice, and thus to Bob. Hence we write

$$G_j = |\Gamma_j\rangle\langle\Gamma_j| \text{ for } j \geqslant 1, \quad G_0 = I - \sum_{j\geqslant 1} G_j. \tag{24}$$

If Bob's operation L_j is successful, the net result will be a unitary map V from $\mathcal{H}_c$ to $\mathcal{H}_a$ as indicated in Fig. 9(a), up to a constant of proportionality designated by $\sqrt{p_j}$. Note that V must be independent of j apart from a possible (uninteresting) phase. Part (b) of Fig. 9 is the diagrammatic solution to the equation for $\langle\Gamma_j|$ in (a), and from it one can construct the diagram for G_j in (c). The probability that Alice will obtain outcome j *and* Bob will succeed in carrying out the corresponding operation L_j is given by

$$\langle\Omega|L_j^\dagger G_j L_j|\Omega\rangle = \langle\Omega|L_j^\dagger|\Gamma_j\rangle\langle\Gamma_j|L_j|\Omega\rangle = \||\langle\Gamma_j|L_j|\Omega\rangle\|^2 = p_j, \tag{25}$$

assuming an initial normalized state (16). Here $\langle\Gamma_j|L_j|\Omega\rangle \in \mathcal{H}_b$ is the ket indicated in Fig. 9(d), and the right side of this figure justifies the final equality in (25), since $|c\rangle$ is normalized and V is unitary.

For the POVM to be physically realizable it must be the case that

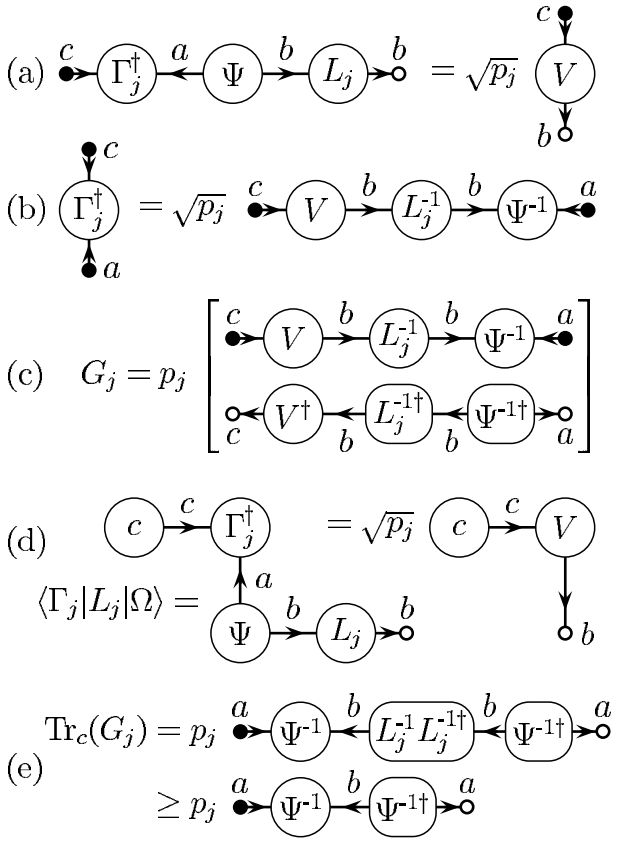

FIG. 9. Diagrams associated with unambiguous teleportation using a POVM followed by a nonunitary correction.

$$\sum_{j\geqslant 1} G_j \leqslant I_c \otimes I_a, \tag{26}$$

which implies that

$$d \cdot I_a \geqslant \sum_j \mathrm{Tr}_c(G_j) \geqslant \left(\sum_{j\geqslant 1} p_j\right) \Psi^{-1}\Psi^{-1\dagger}, \tag{27}$$

where the second inequality, Fig. 9(e), comes about from the fact that the largest eigenvalue of the positive operator $L_j^\dagger L_j$ cannot exceed 1 (Appendix A), so the smallest eigenvalue of its inverse $L_j^{-1}L_j^{-1\dagger}$ is not less than 1, and therefore $L_j^{-1}L_j^{-1\dagger} \geqslant I_b$. As the largest eigenvalue of the operator $\Psi^{-1}\Psi^{-1\dagger}$ is $1/\lambda_m^2$, (27) implies that

$$p_s = \sum_{j\geqslant 1} p_j \leqslant d\lambda_m^2 = p_c, \tag{28}$$

in agreement with the argument given in Sec. IV A.

D. Optimum strategies

Since optimal unambiguous entanglement concentration followed by a standard teleportation protocol achieves the optimal probability of success, an obvious way to produce an optimal strategy of the form described in Sec. IV C is to produce Alice's POVM by combining the operator K for entanglement concentration with the fully entangled basis $\{|\Phi_j\rangle\}$ appropriate for standard teleportation, setting

$$|\Gamma_j\rangle = (I_c \otimes K^\dagger)|\Phi_j\rangle. \tag{29}$$

in (24). Using the fact that $\Sigma_j |\Phi_j\rangle\langle\Phi_j|$ is the identity on $\mathcal{H}_{ac}$, one can check that

$$\sum_{j\geqslant 1} G_j = \sum_{j\geqslant 1} |\Gamma_j\rangle\langle\Gamma_j| = I_c \otimes K^\dagger K \leqslant I_c \otimes I_a, \tag{30}$$

as the largest eigenvalue of $K^\dagger K$ is 1, so the POVM, with G_0 given by (24), is physically realizable. Bob, on the other hand, carries out unitary corrections which always succeed, as in standard teleportation. By combining (29), (22) with $\mu = \sqrt{d}\lambda_m$, (8) and (25), with L_j and $L_j^\dagger$ omitted from the last, one finds that $p_j = \lambda_m^2/d$, independent of the initial $|c\rangle$ and independent of j, so (28) is an equality, as expected.

Obviously, Bob rather than Alice could carry out the entanglement concentration which takes $|\Psi\rangle$ to $|\Psi_f\rangle$ using an operator L which is the analog of K in (22), and, if successful, utilize the outcome of Alice's measurements in the fully entangled basis $\{|\Phi_j\rangle\}$ to apply a unitary correction U_j in case j to his half of $|\Psi_f\rangle$. His two steps can be combined into one by defining an operation

$$L_j = U_j L. \tag{31}$$

Thus in this protocol Alice's POVM is the standard projective measurement in a fully entangled basis, whereas Bob's correction L_j will typically not be unitary, and will sometimes fail. Because the shared resource is only partially entangled, the probability that Alice will obtain outcome j depends in general on the state $|c\rangle$ she is trying to teleport, contrary to what one might have expected from the general rule (see, for example, [7]) that teleportation is only possible when neither Alice nor Bob learns anything about the state being teleported. However, the *combined* probability p_j that Alice obtains j *and* Bob succeeds in carrying out L_j is λ_m^2/d independent of $|c\rangle$, in agreement with the general rule. It is also independent of j, so (28) is again an equality, as expected.

One can design other optimum procedures by imagining the entanglement concentration task shared by Alice and Bob; e.g., each could do half of it based upon

$$|\Psi^{-1/2}\rangle = \sum_j \lambda_j^{-1/2} |\bar{a}_j\rangle \otimes |\bar{b}_j\rangle. \tag{32}$$

After that they carry out measurements and unitary corrections as in the standard protocol. Upon combining Alice's concentration and measurement steps in a single POVM, and Bob's concentration and unitary correction steps in nonunitary operators as in (31), one arrives at a combined protocol with, once again, the optimal probability of success. There may well be other possibilities, but it seems clear that in one way or another they will have to employ the inverse, (9), of the partially entangled ket $|\Psi\rangle$, and can be understood in terms of how they achieve this.

E. Literature

Unambiguous entanglement concentration is our term for the "Procrustean method" introduced by Bennett *et al.* [12]. It is often referred to as distillation or entanglement concen-

tration of a single copy, in contrast to schemes (again see [12]), where operations are carried out on multiple independent systems in the same entangled state. That the probability of success in (23) is the maximum possible using local operations and classical communication was shown by Vidal [13] and by Lo and Popescu [6].

The notion of unambiguous teleportation was introduced by Mor in 1996 in unpublished work [14] for $d=2$, using a protocol in which Alice carries out a POVM and Bob a unitary correction. For a complete account including further developments [15], see [11]. During this development the term *conclusive teleportation* acquired a broader sense, and in [11] it is "perfect conclusive teleportation" that corresponds to our "unambiguous teleportation." Among other things, these authors obtained the optimum probability of success, (28) with $d=2$, and pointed out the close connection between unambiguous teleportation and entanglement concentration, as noted above in Sec. IV A. Their thinking of a POVM carried out on an entangled state as producing some sort of a mysterious choice-at-a-distance (telePOVM), a perspective derived from Hughston *et al.* [16], employs vivid imagery that is not actually necessary for a sober discussion of teleportation in terms of local quantum mechanics when conditional probabilities are used in a consistent way; see [17], and Chaps. 23 and 24 of [8].

Nonunitary corrections by Bob in the $d=2$ case, as well as modifications of Alice's measurement scheme relative to the standard teleportation protocol, were considered by Li *et al.* [18] and by Bandyopadhyay [19]. Both articles contain explicit protocols which achieve the optimum probability of success. By contrast, Agrawal and Pati in [20,21] (the two papers are quite similar) restricted Alice to projective measurements and Bob to unitary corrections, so their protocol does not have the optimum probability of success. However, they made the interesting observation that the measurement basis states used by Alice must have the same entanglement as that of the shared resource. From our perspective this arises from the fact that a $d=2$ bipartite ket and its inverse have the same entanglement, since they have the same Schmidt coefficients when normalized in the same way. However, this is no longer true in higher dimensions $d \geqslant 3$, where the entanglement of the inverse is in general not the same as that of the original ket. (The term "entanglement matching" occurs in [18], but its significance is not clear.)

There have been several studies of unambiguous teleportation in the case of general $d=d_a=d_b=d_c$, the one we have been considering. Son *et al.* [22] and Roa *et al.* [23] both employ protocols in which Alice uses a POVM and Bob applies a unitary correction. In both cases the probability of success is optimal in the sense that (28) is an equality. However, their arguments that they have optimal procedures are not easy to follow, and in [23] the aim seems to be that of maximizing the *average* fidelity (over all pure input states) of the imperfect teleportation scheme, rather than maximizing the probability p_s (in our notation) of conclusive teleportation with unit fidelity.

The approach of Kurucz and colleagues [24,25] is similar to ours in focusing attention on the inverse of the partially entangled resource state. They construct the inverse by first using the entangled state $|\Psi\rangle$ (in our notation) to produce an antilinear map from $\mathcal{H}_a$ to $\mathcal{H}_b$, which then has an inverse of the usual sort. They use this to place conditions on Alice's measurement, again converted to an antilinear map, so that it will result in a unitary channel from $\mathcal{H}_c$ to $\mathcal{H}_b$, with Bob carrying out a final unitary correction. They consider general d, but do not address the question of a measurement strategy that achieves the maximum overall probability of success, unlike the work cited in the previous paragraph. The use of antilinear maps, which were also discussed in connection with teleportation by Uhlmann [26,27], has a certain elegance in that these maps are basis independent, unlike their linear counterparts, such as the one on the right side of our Fig. 5(d), for which a basis needs to be specified. In our own approach the inverse of a bipartite entangled ket is also (when it exists) a basis-independent concept, so it is not clear to us that antilinear maps possess a significant conceptual advantage, but we are happy to leave this to the reader's judgment.

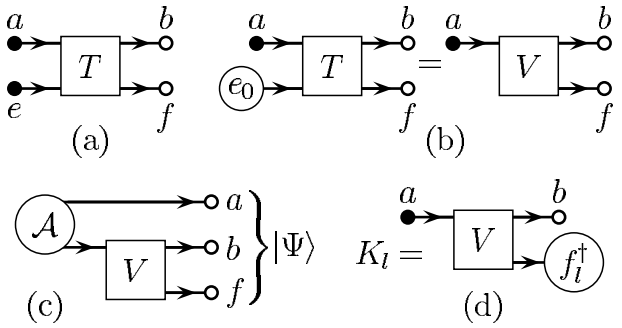

FIG. 10. Model for a noisy channel based on (a) the unitary map T which induces the isometry V in (b) if the environment is initially in a pure state $|e_0\rangle$. Transposing the isometry yields the channel ket $|\Psi\rangle$ in (c). A Kraus operator (d) obtained using (33).

Inverses may also be taken in a manner that makes explicit use of particular bases, as in the work of Li *et al.* [28]. They allow a very general measurement on Alice's part and also a nonunitary correction by Bob, for a system of arbitrary dimensionality, as in Sec. IV C above. However, they only consider a single measurement by Alice, and thus do not address the issue of optimizing unambiguous teleportation.

We have deliberately omitted referring to the substantial literature on some closely related topics: teleportation using a mixed state as a resource, teleportation between systems of different dimensionality, and the average fidelity in cases where the resource is partially entangled. It may be that atemporal diagrams are of some use in this broader context, but that remains to be seen.

V. NOISY QUANTUM CHANNELS

A. Channel operators

A standard model of a noisy quantum channel is shown in Fig. 10. It begins with a unitary time-development operator T in (a), which maps the channel input $\mathcal{H}_a$ and environment $\mathcal{H}_e$ to the channel output $\mathcal{H}_b$ and environment $\mathcal{H}_f$ at a later time. Unitarity implies that $\mathcal{H}_{ae}$ and $\mathcal{H}_{bf}$ have equal dimension, but the dimensions d_a and d_b of $\mathcal{H}_a$ and $\mathcal{H}_b$ might be different.

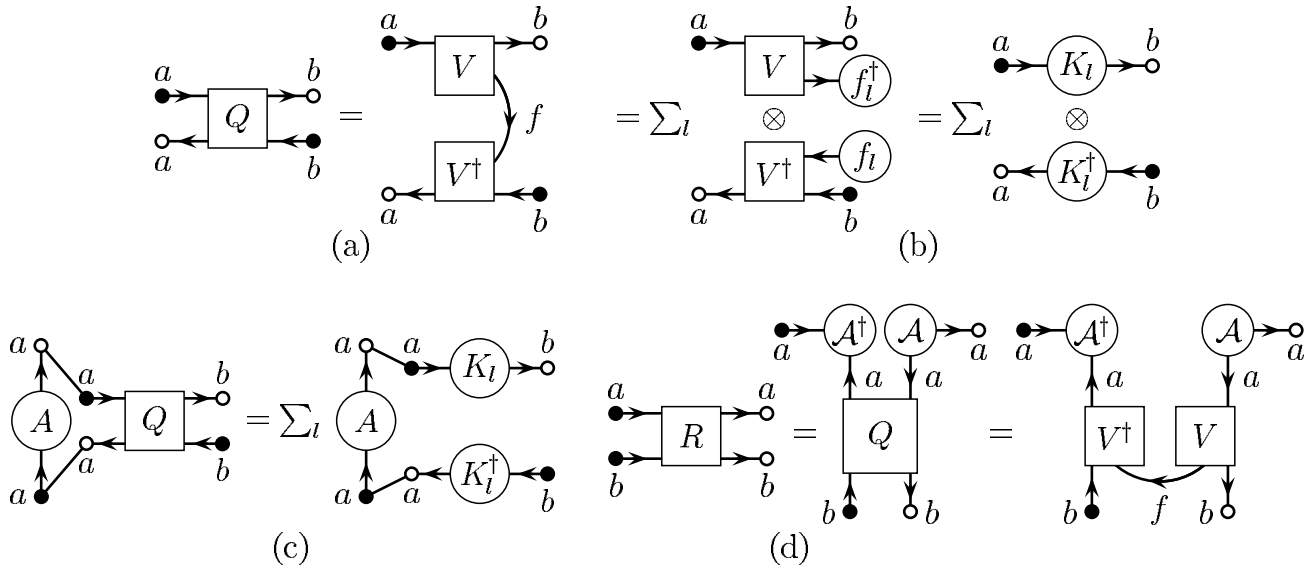

FIG. 11. (a) Definition of the transition operator Q. (b) Equivalent expressions for Q in terms of Kraus operators. (c) Contraction with Q yields the superoperator map $\mathcal{V}(A)$; see Eq. (35). (d) The dynamical operator R is the partial transpose of Q

(The notation used here follows that of [2], with minor modifications, to facilitate comparison.) The diagram can be thought of as a quantum circuit to which nodes and arrows have been added, thus converting it to the corresponding atemporal diagram. If one assumes the environment is initially in a pure state $|e_0\rangle$, the appropriate diagram in part (b) of the figure has three active nodes and represents the isometry $V=T|e_0\rangle$ mapping $\mathcal{H}_a$ to $\mathcal{H}_{bf}$. Part (c) of the figure shows the channel ket $|\Psi\rangle \in \mathcal{H}_{abf}$ obtained by applying a transposer $\mathcal{A}$ to the a node of V. (In [2] a fully entangled state $|\phi\rangle$ was used instead of $|\mathcal{A}\rangle$, giving the same result apart from normalization.) Part (d) is the diagrammatic form for a Kraus operator

$$K_l = \langle f_l|V, \tag{33}$$

an element of $\mathcal{L}_{ba}$, where $\{|f_l\rangle\}$ is some orthonormal basis for $\mathcal{H}_f$. (One can, if one wants, think of K_l as associated with a measurement on $\mathcal{H}_f$.)

The collection of Kraus operators $\{K_l\}$ depends, of course, on the choice of basis $\{|f_l\rangle\}$ used to define them. However, the subspace of $\mathcal{L}_{ba}$ spanned by their linear combinations is independent of this choice, and we shall refer to its dimension as the *Kraus rank* κ of the noisy channel. This subspace is the range of the "cross operator" obtained when V is regarded as a map from $\mathcal{H}_f^\dagger$ to $\mathcal{L}_{ba}$, from the bottom to the top of Fig. 10(b), written $V_{ba;f}$ in the notation employed in Sec. II E. Consequently, the Kraus rank is the (ordinary) rank of this cross operator,

$$\kappa = \mathrm{Rn}(V_{ba;f}) \leqslant \min\{d_a d_b, d_f\}, \tag{34}$$

where the inequality reflects the fact that the rank of a matrix cannot exceed the number of rows or the number of columns.

A channel is usually discussed in terms of the superoperator $\mathcal{V}$ that maps $\hat{\mathcal{H}}_a$ to $\hat{\mathcal{H}}_b$; in particular, it maps density operators at the channel entrance to density operators at the channel output. Two ways of writing it are

$$\mathcal{V}(A) = \mathrm{Tr}_a[(A \otimes I)Q] = \sum_l K_l A K_l^\dagger, \tag{35}$$

where $Q \in \hat{\mathcal{H}}_a \otimes \hat{\mathcal{H}}_b$ is the *transition operator*, in the notation of [2], and the Kraus operators were defined in (33) above. The object Q corresponds to an atemporal diagram with four legs obtained by contracting the tensor product of V with its adjoint $V^\dagger$, Fig. 11(a). By replacing the f contraction line, thought of as I_f, with $\Sigma_l |f_l\rangle\langle f_l|$, see Fig. 3(e), one obtains Q in the form shown in Fig. 11(b) as a sum of tensor products of Kraus operators. Thus parts (a) and (b) of the figure correspond to writing Q as

$$Q = \mathrm{Tr}_f(V \otimes V^\dagger) = \sum_l K_l \otimes K_l^\dagger. \tag{36}$$

The meaning of these equations is probably clearest if they are used in conjunction with the atemporal diagrams in part (a) and (b) of Fig. 11. The way in which Q generates the superoperator $\mathcal{V}$ in (35), as a map from the left pair of nodes to the right pair of nodes, is indicated in part (c) of this figure; note how contraction with the operator A yields a diagram with the open and closed b nodes signifying an operator on $\mathcal{H}_b$.

The *dynamical operator*

$$R = \mathcal{A}^\dagger Q \mathcal{A} \tag{37}$$

is the partial transpose of Q in the basis $\{|a_j\rangle\}$, (5), and a map of $\mathcal{H}_{ab}$ onto itself. The corresponding diagram is shown in Fig. 11(d), where R acts from left to right and Q has been rotated by 90° relative to its orientation in (a). The left-to-right reflection symmetry of the final diagram shows that R is a positive (semidefinite) operator, of the general form $WW^\dagger$, where $W=V\mathcal{A}$, regarded as a map from $\mathcal{H}_f^\dagger$ to $\mathcal{H}_{ab}$, may be thought of as a transposed form of the $V_{ba;f}$ cross operator introduced previously. Thus the rank of W is equal to the Kraus rank κ, and since the rank of $WW^\dagger$ is the same as that

of W (see, e.g., p. 13 of [5]), the Kraus rank can also be defined, as in [2], to be the (ordinary) rank of the dynamical operator R. This in turn is the rank of Q when it is regarded as a map from $\mathcal{L}_{ba}$ to $\mathcal{L}_{ba}$, i.e., from bottom to top in Fig. 11(a).

B. Complete positivity

As is well known, the superoperator $\mathcal{V}$ for a quantum channel is a *completely positive* map from $\hat{\mathcal{H}}_a$ to $\hat{\mathcal{H}}_b$ in the following sense. Let $\hat{\mathcal{H}}_s$ be the space of operators on a Hilbert space $\mathcal{H}_s$ distinct from any of those we have been considering, and define the superoperator $\mathcal{W}$ mapping $\hat{\mathcal{H}}=\hat{\mathcal{H}}_a \otimes \hat{\mathcal{H}}_s$ to $\hat{\mathcal{H}}'=\hat{\mathcal{H}}_b \otimes \hat{\mathcal{H}}_s$ through

$$\mathcal{W}(A \otimes S) = \mathcal{V}(A) \otimes S, \tag{38}$$

for any S in $\hat{\mathcal{H}}_s$; i.e., $\mathcal{W}$ is the tensor product of $\mathcal{V}$ with an identity map on $\hat{\mathcal{H}}_s$. Then $\mathcal{V}$ is defined to be a *completely positive map* provided any superoperator $\mathcal{W}$ constructed in this manner is a *positive map*, in the sense that whenever $P \in \hat{\mathcal{H}}$ is a positive (semidefinite) operator, $\mathcal{W}(P) \in \hat{\mathcal{H}}'$ is also a positive (semidefinite) operator.

This definition is neither simple nor constructive: it provides no direct way to check whether some superoperator is or is not completely positive. The diagram in Fig. 11(a) suggests a simpler characterization. If one thinks of Q as acting from "bottom to top," which means as a map of $\mathcal{L}_{ba}$ to itself, the atemporal diagram has the characteristic symmetry of a positive (definite) operator; see Sec. II D and the examples in Fig. 4(b) and 4(e). In fact, the positivity of this "cross operator" of the transition operator Q is a necessary and sufficient condition that $\mathcal{V}$ be a completely positive map, as first pointed out by Choi. (See theorem 2 of [29]; the reader who finds the argument difficult to follow may wish to consult [30].) A diagrammatic proof of this result is given in Fig. 12.

Figure 12(a) shows that if A is a positive operator on $\mathcal{H}_a$, $\mathcal{V}(A)$ is a positive operator on $\mathcal{H}_b$. Here A has been written in the form $N^\dagger N$, see Fig. 4(b), and one can see that the operator on $\mathcal{H}_b$ which results from contraction with Q is positive, because of the (top-to-bottom) reflection symmetry of the diagram. The transition operator corresponding to the superoperator $\mathcal{W}$ in (38) is shown in Fig. 12(b), where the two horizontal lines representing the identity superoperator on $\hat{\mathcal{H}}_s$ are drawn in a way which preserves the top-to-bottom reflection symmetry. Thus when $\mathcal{W}$ is applied to an arbitrary positive operator, shown in (c) as $X^\dagger X$, the result is an operator which is positive. Note that any positive operator on $\mathcal{L}_{ba}$ can be expressed in the form of a diagram similar to that used for Q in Fig. 11(a), with V some object that need not be an isometry. Thus what we have shown is that if the "cross operator" of the transition operator defining any superoperator $\mathcal{V}$ is positive (semidefinite), the superoperator $\mathcal{V}$ itself is completely positive, according to the standard definition based on (38).

The converse argument, from $\mathcal{W}$ in the form (38) as a positive map to the positivity of Q as an operator on $\mathcal{L}_{ba}$, begins in Fig. 12(d). We have chosen $\mathcal{H}_s$ to be a copy of $\mathcal{H}_a$

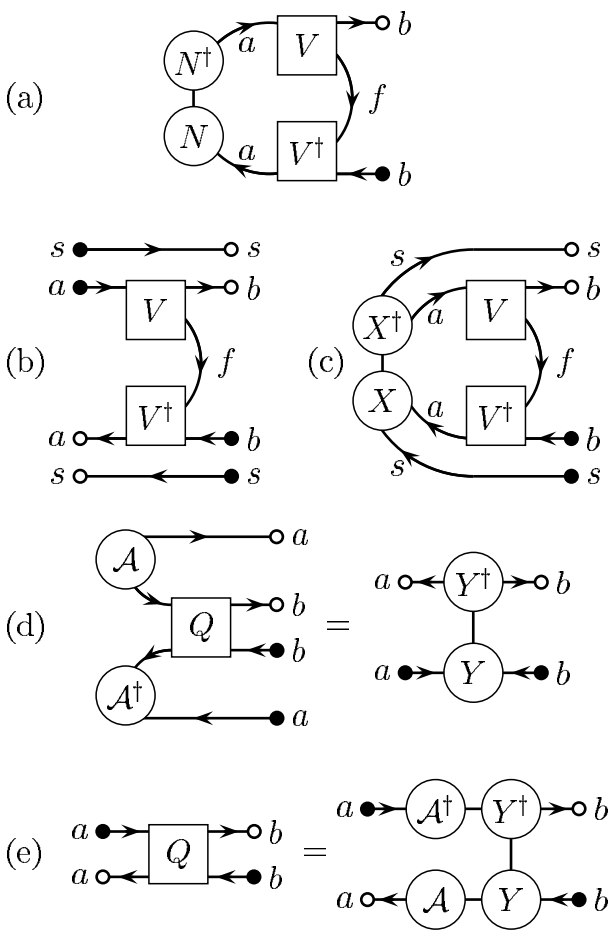

FIG. 12. Argument that complete positivity of the superoperator $\mathcal{V}$ corresponding to the transition operator Q is equivalent to positivity of the "cross operator" mapping $\mathcal{L}_{ba}$ to itself.

and allowed $\mathcal{W}$ to act on a particular operator $\mathcal{A} \otimes \mathcal{A}^\dagger$ formed from the transposer defined in Sec. II E and its adjoint, which by symmetry is a positive (semidefinite) operator on $\mathcal{H}_a \otimes \mathcal{H}_a$. Since by assumption $\mathcal{W}$ maps positive operators to positive operators, the result must be a positive operator on $\mathcal{H}_{ab}$, and thus of the form $Y^\dagger Y$ shown on the right side of (d). The equation for Q in (d) is solved in (e), and the symmetry of the diagram on the right side (read from bottom to top) implies that the cross operator of Q acting on $\mathcal{L}_{ba}$ is positive semidefinite. This completes the proof of equivalence.

In constructing the proof we have, incidentally, proven that the complete positivity of $\mathcal{V}$ is equivalent to the positivity of the *dynamical operator* $R=\mathcal{A}^\dagger Q \mathcal{A}$, the partial transpose of the transition operator Q with respect to the orthonormal basis of $\mathcal{H}_a$ giving rise to $\mathcal{A}$. While the preceding results are not new, the diagrammatic analysis makes the derivation particularly simple by showing that everything hinges on properties of a single "object" Q viewed in various different ways.

C. Channel based on mixed-state environment

For simplicity the following discussion is limited to the case of a noisy channel in which the input and output have

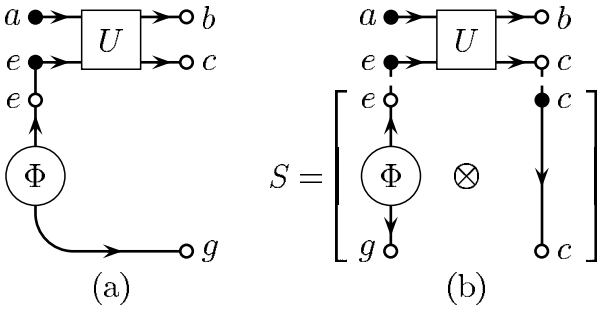

FIG. 13. Channel based on mixed-state environment obtained by tracing $|\Phi\rangle\langle\Phi|$ over $\mathcal{H}_g$. In (b) the atemporal diagram in (a) is redrawn in a way which helps interpreting it as a cross operator.

the same dimension $d_b=d_a$. In order to model this channel using a unitary T applied to the system together with an environment initially in a pure state, as in Fig. 10(b), the dimension $d_e=d_f$ of the environment must be at least as great as the Kraus rank κ; see the remarks in Sec. V A. But κ cannot exceed $d_a d_b=d_a^2$, consistent with the well-known fact that using an environment of dimension $d_e=d_a^2$ and a suitable T one can model any channel of this sort.

However, if the environment is initially in a mixed state ρ_e, d_e^2 rather than d_e becomes the upper bound on κ [see below following (40)], and thus one might have supposed that any noisy channel could be modeled using an environment of dimension $d_e=d_a$, rather than $d_e=d_a^2$, by allowing it to be initially in a mixed state. In fact this is not the case [31], but good criteria for distinguishing channels which can and cannot be modeled using an environment of dimension $d_e<d_a^2$ in a mixed state are not known at present. In the case $d_e=d_a=2$, numerical evidence [4] indicates that κ for such a channel can only take on the values 1, 2, and 4; 3 is excluded. We shall present a simple proof of this result based on a diagrammatic analysis which, while not resolving the general mixed-environment channel problem, does focus attention on what could be a helpful tool: properties of the cross operator corresponding to a unitary operator on a bipartite system.

Figure 13(a) shows the situation we wish to consider as an atemporal diagram drawn in a way which makes it resemble a quantum circuit. The initial mixed state ρ_e of the environment has been "purified" in the usual fashion by introducing a hypothetical reference system $\mathcal{H}_g$ with dimension $d_g=d_e$, so that ρ_e is the partial trace over $\mathcal{H}_g$ of an entangled state $|\Phi\rangle$ on $\mathcal{H}_{eg}$. Thus we are back to the case considered in Fig. 10: an (enlarged) environment in an initial pure state, with a time development operator $T=U\otimes I_g$, where U is a unitary operator acting on $\mathcal{H}_{ae}$, while the final environment Hilbert space $\mathcal{H}_f$ in that figure is now $\mathcal{H}_{cg}$, with $d_c=d_e$.

The Kraus rank κ of this channel is the rank of the cross operator $V_{ba;f}$, (34), and in studying what values it can take it is helpful to represent it in the form, see Fig. 13(b),

$$V_{ba;f}=V_{ba;cg}=U_{ba;ce}S_{ce;cg} \tag{39}$$

of a product of the cross operator (from bottom to top in the figure) $U_{ba;ce}$ for the unitary U and a second (from bottom to top) map $S_{ce;cg}$ corresponding to the object S defined in the figure. In fact, $S_{ce;cg}$ is a tensor product of the map generated by $|\Phi\rangle$, times the identity on $\mathcal{H}_c$, and thus its rank is $d_c=d_e$ times the Schmidt rank σ of $|\Phi\rangle$. Since the rank of the matrix product of two matrices cannot exceed that of either one, (39) leads to the conclusion that

$$\kappa\leqslant\min\{\mathrm{Rn}(U_{ba;ce}),d_e\sigma\} \tag{40}$$

in the notation of Sec. II E, where $\mathrm{Rn}(\cdots)$ stands for rank. If $\sigma=1$, meaning $|\Phi\rangle$ is a product state and the environment $\mathcal{H}_e$ is initially in a pure state, (40) tells us that κ cannot be larger than d_e. The other extreme is $\sigma=d_e=d_g$, so that d_e^2 replaces d_e as a bound, and since $S_{ce;cg}$ is in this case nonsingular, $\kappa=\mathrm{Rn}(U_{ba;ce})$, as $\mathrm{Rn}(U_{ba;ce})$ cannot exceed d_e^2.

In the particular case in which $d_e=d_a=2$, a one-qubit channel modeled using a one-qubit mixed-state environment, there is a convenient parametrization of the two-qubit unitary U which allows one to show that the rank of its cross operator can only take the values 1, 2, and 4; 3 is excluded. See Sec. IV of [32], where this result is stated, in the notation employed there, as proposition 2. This means that κ cannot be 3, because either $\sigma=1$, in which case $\kappa\leqslant 2$ by (40), or else $\sigma=2$, which means the operator $S_{ce;cg}$ is nonsingular and the Kraus rank is equal to that of the cross operator for U. As noted above, the absence of $\kappa=3$ was conjectured on the basis of numerical work in [4]; so far as we know, ours is the first analytic proof of this result.

In the case of a unitaries acting on two qutrits, $d_e=d_a=3$, it has been shown [33] that there are no restrictions on the rank of the cross operator, which can take any value between 1 and 9, contrary to a conjecture in [32]. Studies by one of us [34] confirm that $d_e=d_a=2$ is in this respect somewhat exceptional.

VI. SUMMARY AND OPEN QUESTIONS

The system of diagrams defined in Sec. II has been carefully constructed to make it easy to convert quantum circuits into atemporal diagrams, as illustrated in the various examples in Figs. 6, 7, and 10. Nonetheless, it should be emphasized that the arrows in atemporal diagrams have an abstract significance not connected with the direction of time, and thus one need not be concerned as to whether complicated diagrams, such as those in Fig. 11 with arrows pointing in various different directions, can be given a temporal interpretation. One of the principal advantages of map-state duality resides precisely in the possibility of removing temporal references. It is, for example, useful when interpreting the diagram in Fig. 7(c) to remember that one can think of M_j as a quantum channel from $\mathcal{H}_c$ to $\mathcal{H}_b$ without requiring that $\mathcal{H}_b$ be at a time *later* than the one at which Alice's measurement takes place.

To be sure, the reader may well object that the atemporal feature of such diagrams, however useful it may be for solving formal problems, serves to empty them of any physical or intuitive content. In response, note that such diagrams are best thought of as pre-probabilities in the terminology of [8]—useful for computing probabilities once appropriate frameworks of quantum histories have been introduced- and not physical reality (i.e., actual quantum histories). Wave

functions obtained by unitary time development and used to calculate Born probabilities are also pre-probabilities (Sec. 9.4 of [8]), and thus atemporal diagrams are no less "real" than a variety of other tools used by quantum physicists. This response, while formally correct, conceals an interesting question: when can one find a framework of quantum histories in which a given atemporal diagram makes sense (i.e., assigns physically meaningful probabilities) in terms of a narrative of Alice preparing something, Bob measuring something, and the like? Alice and Bob live in a world which is irreversible in the thermodynamic sense, which makes preparation very different from measurement, whereas microscopic quantum theory is reversible. Connecting the two is not an altogether trivial task, and while physicists typically assign it to philosophers, there may be some aspects which physicists themselves will have to disentangle in order to attain clear ideas about quantum information.

We have given two applications of atemporal diagrams yielding results which represent a generalization or clarification of work in the previous literature. The first concerns unambiguous teleportation, for which our discussion in Sec. IV, in particular Sec. IV C, includes in a unified scheme all the various examples and results in the previous literature for the situation in which the three relevant Hilbert spaces have the same dimension d. By using the notion of the inverse of the shared entangled ket it is easy to produce a variety of optimal protocols and understand how they are related to each other. (The corresponding problem of unambiguous dense coding, the topic of a different paper by some of us [3], turns out to be much more difficult to analyze.) Of course, when teleportation is discussed outside the "unambiguous" framework by considering a mixed state as a resource, or allowing for transmission of states with fidelity less than 1, the problems become more difficult, and it is not known whether or not atemporal diagrams will aid in their solution.

The second application, showing that the Kraus rank of a one qubit noisy channel with one-qubit mixed-state environment cannot have Kraus rank 3, Sec. V C, while of somewhat limited interest in itself—good numerical evidence for it was available some time ago [4]—raises interesting questions about the relationship of a noisy channel (i.e., the corresponding superoperator) to the "cross operator" of the isometry V, or unitary T, used to model it, Fig. 10(b). This is related to, but not the same as the question of how much entanglement can be produced by a unitary, or some more general operation, acting on a bipartite system; see the extensive discussion in [32]. Since a unitary acting on two qubits seems to be exceptional [33,34] the Kraus rank does not look like it will be useful for understanding what is special about other noisy quantum channels based on a mixed-state environment. However, other properties of the cross operator may be relevant. A better understanding of the connection between a quantum channel and the cross operator of the isometry, or the properties of the dynamical operator (matrix), remains an open question whose significance has, we believe, been somewhat sharpened by our diagrammatic approach, even though it did not originate there.

ACKNOWLEDGMENTS

We thank Y. Sun for comments on unambiguous discrimination, V. Gheorghiu for a suggestion used in Appendix B, and K. Życzkowski for comments on the bibliography. The research described here received support from the National Science Foundation through Grant Nos. PHY-0139974 and PHY-0456951.

APPENDIX A: PROBABILITY OF AN UNAMBIGUOUS OPERATION

Let K be a linear map from $\mathcal{H}_a$ to $\mathcal{H}_b$, where these are any two Hilbert spaces. We are interested in the probability that K can be carried out physically as an *unambiguous operation* in the sense that one could, at least in principle, construct an apparatus such that for any input state $|\alpha\rangle \in \mathcal{H}_a$, the output is guaranteed to be $K|\alpha\rangle \in \mathcal{H}_b$, provided an *auxiliary system* at the end of the process has a property S indicating that the operation has been successful.

To be more precise, suppose there is a unitary transformation T that maps $\mathcal{H}_a \otimes \mathcal{H}_x$ to $\mathcal{H}_b \otimes \mathcal{H}_y$, where $\mathcal{H}_x$ and $\mathcal{H}_y$ are the Hilbert spaces of the initial and final auxiliary systems (which could be the same if $d_a = d_b$). Assume the initial state is $|\alpha\rangle \otimes |\xi_0\rangle \in \mathcal{H}_a \otimes \mathcal{H}_x$, with $|\xi_0\rangle$ fixed, and that success corresponds to a subspace of $\mathcal{H}_y$ with projector $S \in \mathcal{H}_y$ chosen in such a way that

$$SV|\alpha\rangle = K|\alpha\rangle \otimes |\eta_\alpha\rangle, \tag{A1}$$

where the isometry $V: \mathcal{H}_a \to \mathcal{H}_{by}$ is defined by

$$V|\alpha\rangle = T(|\alpha\rangle \otimes |\xi_0\rangle), \tag{A2}$$

and $|\eta_\alpha\rangle \in \mathcal{H}_y$ might depend on the initial $|\alpha\rangle$, as suggested by the subscript. Note that K, T, $|\xi_0\rangle$, V, and S are all regarded as fixed quantities characteristic of the apparatus, whereas the initial $|\alpha\rangle$ is thought of as variable, or unknown, and the fixed apparatus must perform in such a way that (A1) is true for *any* initial $|\alpha\rangle$.

None of the kets in (A1) need be normalized, and K could be replaced by cK, c any nonzero constant, without altering the following discussion (e.g., the constant could be absorbed in $|\eta_\alpha\rangle$). That is, the operation should be thought of as mapping rays (one-dimensional subspaces) in $\mathcal{H}_a$ to rays in $\mathcal{H}_b$. The key point is that in the case of success the outcome, (A1), must be a *product* state; if it is entangled, the desired result will arise with, at best, some probability less than 1. Note also that (A1) is assumed to hold for *all* $|\alpha\rangle$ in $\mathcal{H}_a$. If producing (the ray corresponding to) $K|\alpha\rangle$ for a *single* $|\alpha\rangle$ is all that is required, it is trivial to construct an apparatus to do this with certainty. One can interpret success as the positive outcome of an ideal measurement carried out on $\mathcal{H}_y$ to distinguish S from $I-S$; e.g., a green light flashes in case S and a red light for $I-S$. Since we have allowed an arbitrary auxiliary system, there is no need to talk about POVM's; projective measurements are sufficient. And if one uses a framework in which idealized measurements indicate pre-existing properties (Chap. 17 of [8]), there is no need to even talk about measurements.

The linearity of the operators in (A1) and the fact that it holds for all $|\alpha\rangle$ in $\mathcal{H}_a$ means—set $E=K$ and $F=SV$ in Ap-

pendix B—that there is a *fixed* ket $|\bar{\eta}\rangle \in \mathcal{H}_y$, independent of $|\alpha\rangle$, such that

$$SV|\alpha\rangle = K|\alpha\rangle \otimes |\bar{\eta}\rangle = \bar{K}|\alpha\rangle \otimes |\eta_0\rangle, \tag{A3}$$

where the second equality introduces the normalized quantities

$$|\eta_0\rangle = |\bar{\eta}\rangle/\|\bar{\eta}\|, \quad \bar{K} = \|\bar{\eta}\|K, \tag{A4}$$

with $\|\bar{\eta}\| = \sqrt{\langle\bar{\eta}|\bar{\eta}\rangle}$ the usual norm.

The probability of success given an initial normalized state $|\alpha\rangle$, $\|\alpha\|=1$, is given by the Born rule

$$\Pr(S|\alpha) = \langle\alpha|V^\dagger SV|\alpha\rangle = \langle\alpha|\bar{K}^\dagger\bar{K}|\alpha\rangle \leqslant 1, \tag{A5}$$

where the final inequality, necessarily true of a probability, can be checked by writing $I_a = V^\dagger V$ as a sum of the two positive operators $V^\dagger SV$ and $V^\dagger(I-S)V$. As this inequality holds for any (normalized) $|\alpha\rangle$, the maximum eigenvalue of the positive operator $\bar{K}^\dagger\bar{K}$ cannot exceed 1. Thus the probability of success is bounded by

$$\Pr(S|\alpha) \leqslant \langle\alpha|\hat{K}^\dagger\hat{K}|\alpha\rangle, \tag{A6}$$

where $\hat{K}$ is defined as cK when $c>0$ chosen so that the maximum eigenvalue of $\hat{K}^\dagger\hat{K}$ is precisely 1.

In fact, the upper bound is achievable using the following strategy. Since $I-\hat{K}^\dagger\hat{K}$ is a positive operator, the same is true of its positive square root $\hat{L}$, and we can define an isometry

$$V|\alpha\rangle = \hat{K}|\alpha\rangle \otimes |\eta_0\rangle + \hat{L}|\alpha\rangle \otimes |\eta_1\rangle, \tag{A7}$$

using any two orthonormal states $|\eta_0\rangle$ and $|\eta_1\rangle$ in $\mathcal{H}_y$. With $S = |\eta_0\rangle\langle\eta_0|$, (A3) holds with $\bar{K}=\hat{K}$, making (A6) an equality. (Extending V to a unitary T is a straightforward exercise.)

APPENDIX B: VECTOR CONSTANT OF PROPORTIONALITY

Let $\mathcal{V}_a$, $\mathcal{V}_b$, and $\mathcal{V}_c$ be any three linear spaces—we are interested in Hilbert spaces, but the inner product plays no role in the following argument—and suppose that $E:\mathcal{V}_a \to \mathcal{V}_b$ and $F:\mathcal{V}_a \to \mathcal{V}_b \otimes \mathcal{V}_c$ are two linear maps such that

$$F|\alpha\rangle = (E|\alpha\rangle) \otimes |\gamma_\alpha\rangle \tag{B1}$$

for every $|\alpha\rangle$ in $\mathcal{V}_a$, where $|\gamma_\alpha\rangle$ might depend upon $|\alpha\rangle$, as indicated by the subscript.

Theorem. Equation (B1) implies that there is a fixed vector $|\bar{\gamma}\rangle \in \mathcal{V}_c$ such that

$$F|\alpha\rangle = (E|\alpha\rangle) \otimes |\bar{\gamma}\rangle. \tag{B2}$$

Even in the case in which $\mathcal{V}_c$ is one dimensional, the collection of scalars, the result is not trivial, and it is useful to state it as:

Corollary. If G and E are two linear maps from $\mathcal{V}_a$ to $\mathcal{V}_b$ such that

$$G|\alpha\rangle = c(\alpha)E|\alpha\rangle \tag{B3}$$

for all $|\alpha\rangle$, where $c(\alpha)$ is a scalar that may depend on $|\alpha\rangle$, then there is a fixed scalar $\bar{c}$ such that $G=\bar{c}E$.

In constructing the proof, note that if $E|\alpha\rangle$, and hence $F|\alpha\rangle$, is zero, $|\gamma_\alpha\rangle$ can be anything. Thus our the task is to show that for all nonzero $E|\alpha\rangle$ the corresponding $|\gamma_\alpha\rangle$ does not, in fact, depend on $|\alpha\rangle$. Then if we set $|\bar{\gamma}\rangle$ equal to this constant vector, (B2) will hold in all cases, including those in which $E|\alpha\rangle$ vanishes. Let $|\alpha_0\rangle$ and $|\alpha_1\rangle$ be any two distinct elements of $\mathcal{V}_a$ for which

$$|\beta_0\rangle = E|\alpha_0\rangle \neq 0, \quad |\beta_1\rangle = E|\alpha_1\rangle \neq 0, \tag{B4}$$

and let $|\gamma_0\rangle$ and $|\gamma_1\rangle$ be the corresponding elements of $\mathcal{V}_c$ in (B1). In showing that $|\gamma_1\rangle = |\gamma_0\rangle$ we shall consider three cases that exhaust all the possibilities.

Case 1. The vectors $|\alpha_0\rangle$ and $|\alpha_1\rangle$ are linearly dependent, say, $|\alpha_1\rangle = c|\alpha_0\rangle$. Then by linearity $|\beta_1\rangle = c|\beta_0\rangle$, $F|\alpha_1\rangle = cF|\alpha_0\rangle$, and from (B1) it follows at once that $|\gamma_0\rangle = |\gamma_1\rangle$.

Case 2. The vectors $|\alpha_0\rangle$ and $|\alpha_1\rangle$ are linearly independent, but $|\beta_0\rangle$ and $|\beta_1\rangle$ are linearly dependent, so that

$$|\beta_1\rangle = c|\beta_0\rangle \tag{B5}$$

for some $c \neq 0$. This means that $E(|\alpha_1\rangle - c|\alpha_0\rangle) = 0$ and therefore $F(|\alpha_1\rangle - c|\alpha_0\rangle) = 0$. Rewrite the second equality as

$$F|\alpha_1\rangle = |\beta_1\rangle \otimes |\gamma_1\rangle = cF|\alpha_0\rangle = c|\beta_0\rangle \otimes |\gamma_0\rangle. \tag{B6}$$

Using (B5) and $|\beta_1\rangle \neq 0$, see (B4), we conclude that $|\gamma_1\rangle = |\gamma_0\rangle$.

Case 3. The vectors $|\beta_0\rangle$ and $|\beta_1\rangle$ are linearly independent. Then use (B1), with $|\gamma_\alpha\rangle = |\gamma_2\rangle$ when $|\alpha\rangle = |\alpha_0\rangle + |\alpha_1\rangle$, and linearity, to obtain

$$\begin{aligned} F(|\alpha_0\rangle + |\alpha_1\rangle) &= (|\beta_0\rangle + |\beta_1\rangle) \otimes |\gamma_2\rangle \\ &= |\beta_0\rangle \otimes |\gamma_0\rangle + |\beta_1\rangle \otimes |\gamma_1\rangle. \end{aligned} \tag{B7}$$

Rewrite the second equality as

$$|\beta_0\rangle \otimes (|\gamma_2\rangle - |\gamma_0\rangle) = |\beta_1\rangle \otimes (|\gamma_1\rangle - |\gamma_2\rangle) \tag{B8}$$

to make it obvious that, since the $|\beta_j\rangle$ are linearly independent, $|\gamma_2\rangle = |\gamma_0\rangle$ and $|\gamma_1\rangle = |\gamma_2\rangle$, so $|\gamma_1\rangle = |\gamma_0\rangle$.

[1] K. Życzkowski and I. Bengtsson, Open Syst. Inf. Dyn. **11**, 3 (2004), eprint quant-ph/0401119.

[2] R. B. Griffiths, Phys. Rev. A **71**, 042337 (2005).

[3] S. Wu *et al.*, Phys. Rev. A **73**, 042311 (2006).

[4] B. M. Terhal, I. L. Chuang, D. P. DiVincenzo, M. Grassl, and J. A. Smolin, Phys. Rev. A **60**, 881 (1999).

[5] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, England, 1985).

[6] H.-K. Lo and S. Popescu, Phys. Rev. A **63**, 022301 (2001).

[7] M. A. Nielsen and C. M. Caves, Phys. Rev. A **55**, 2547

(1997).
[8] R. B. Griffiths, *Consistent Quantum Theory* (Cambridge University Press, Cambridge, England, 2002).
[9] R. F. Werner, J. Phys. A **34**, 7081 (2001).
[10] A. Chefles and S. M. Barnett, Phys. Lett. A **250**, 223 (1998).
[11] G. Brassard, P. Horodecki, and T. Mor, IBM J. Res. Dev. **48**, 87 (2004).
[12] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A **53**, 2046 (1996).
[13] G. Vidal, Phys. Rev. Lett. **83**, 1046 (1999).
[14] T. Mor, e-print quant-ph/9608005.
[15] T. Mor and P. Horodecki, e-print quant-ph/9906039.
[16] L. P. Hughston, R. Jozsa, and W. K. Wootters, Phys. Lett. A **183**, 14 (1993).
[17] R. B. Griffiths, Phys. Rev. A **66**, 012311 (2002).
[18] W.-L. Li, C.-F. Li, and G.-C. Guo, Phys. Rev. A **61**, 034301 (2000).
[19] S. Bandyopadhyay, Phys. Rev. A **62**, 012308 (2000).
[20] P. Agrawal and A. K. Pati, Phys. Lett. A **305**, 12 (2002).
[21] A. K. Pati and P. Agrawal, J. Opt. B: Quantum Semiclassical Opt. **6**, S844 (2004).
[22] W. Son, J. Lee, M. S. Kim, and Y.-J. Park, Phys. Rev. A **64**, 064304 (2001).
[23] L. Roa, A. Delgado, and I. Fuentes-Guridi, Phys. Rev. A **68**, 022310 (2003).
[24] Z. Kurucz, M. Koniorczyk, and J. Janszky, Fortschr. Phys. **49**, 1019 (2001).
[25] Z. Kurucz, M. Koniorczyk, P. Adam, and J. Janszky, J. Opt. B: Quantum Semiclassical Opt. **5**, S627 (2003).
[26] A. Uhlmann, e-print quant-ph/0301116.
[27] A. Uhlmann, e-print quant-ph/0407244.
[28] C. Li, H.-S. Song, and Y.-X. Luo, Phys. Lett. A **297**, 121 (2002).
[29] M.-D. Choi, Linear Alg. Appl. **10**, 285 (1975).
[30] D. W. Leung, J. Math. Phys. **44**, 528 (2003).
[31] C. Zalka and E. Rieffel, J. Math. Phys. **43**, 4376 (2002).
[32] M. A. Nielsen, C. M. Dawson, J. L. Dodd, A. Gilchrist, D. Mortimer, T. J. Osborne, M. J. Bremner, A. W. Harrow, and A. Hines, Phys. Rev. A **67**, 052301 (2003).
[33] J. E. Tyson, J. Phys. A **36**, 10101 (2003).
[34] L. Yu (unpublished).