Causal and localizable quantum operations

David Beckman,[1,*] Daniel Gottesman,[2,3,†] , M. A. Nielsen,[4,1,‡] and John Preskill[1,§]

[1]*Institute for Quantum Information, California Institute of Technology, Pasadena, California 91125*
[2]*Microsoft Corporation, One Microsoft Way, Redmond, Washington 98052*
[3]*Computer Science Division, EECS, University of California, Berkeley, California 94720*
[4]*Center for Quantum Computer Technology, University of Queensland, Queensland 4072, Australia*

We examine constraints on quantum operations imposed by relativistic causality. A bipartite superoperator is said to be *localizable* if it can be implemented by two parties (Alice and Bob) who share entanglement but do not communicate; it is *causal* if the superoperator does not convey information from Alice to Bob or from Bob to Alice. We characterize the general structure of causal complete-measurement superoperators, and exhibit examples that are causal but not localizable. We construct another class of causal bipartite superoperators that are not localizable by invoking bounds on the strength of correlations among the parts of a quantum system. A bipartite superoperator is said to be *semilocalizable* if it can be implemented with one-way quantum communication from Alice to Bob, and it is *semicausal* if it conveys no information from Bob to Alice. We show that all semicausal complete-measurement superoperators are semilocalizable, and we establish a general criterion for semicausality. In the multipartite case, we observe that a measurement superoperator that projects onto the eigenspaces of a stabilizer code is localizable.

I. INTRODUCTION

What are the *observables* of a relativistic quantum theory? Standard wisdom holds that any self-adjoint operator that can be defined on a spacelike slice through space-time is measurable in principle. But in fact, for most such operators the measurement is forbidden by relativistic causality, and hence impossible.

More generally, it is often stated that the possible ways that the state of a quantum system can change are described by *quantum operations*—completely positive trace-nonincreasing linear maps of density operators to density operators [1,2]. But in a relativistic quantum theory, typical operations would allow superluminal signaling, and are therefore unphysical.

Relativistic quantum-field theory allows us to identify an algebra of observables that is compatible with the causal structure of space-time [3]. Despite this marvelous achievement, puzzles and open questions remain. Our objective in this paper is to understand better the restrictions on operations that are imposed by special relativity. Mostly, we will consider a simplified version of the problem in which the physical system is divided into two separated parts: part A, which is controlled by a party that we will call Alice, and part B, which is controlled by Bob. Initially, Alice and Bob share a joint quantum state whose density operator ρ_{AB} is not known, and they wish to transform the state to $\mathcal{E}(\rho_{AB})$, where $\mathcal{E}$ is a specified operation.

If Alice and Bob were able to communicate by sending quantum information back and forth, then they would be able to apply *any* operation $\mathcal{E}_{AB}$ to their state. We want to determine what operations they can implement if *no* communication (quantum or classical) is permitted. In a relativistic setting, these are the operations that can be realized if Alice's action and Bob's action are spacelike-separated events. We will, though, permit Alice and Bob to make use of a shared entangled *ancilla* state that might have been prepared earlier and distributed to them.

While Alice and Bob are permitted to perform measurements, Bob cannot know the outcome of Alice's measurement, and Alice cannot know the outcome of Bob's. Therefore, we will largely restrict our attention to *trace-preserving* quantum operations, also known as *superoperators*, where no postselection of the quantum state based on the measurement outcome is allowed. We say that a bipartite superoperator is *localizable* if it can be implemented by Alice and Bob acting locally on the shared state and the shared ancilla, without any communication from Alice to Bob or Bob to Alice.

Another important concept is that of a *causal operation*. We say that an operation is causal if it does not allow either party to send a signal to the other. More precisely, imagine that Bob applies a local superoperator $\mathcal{B}$ to his half of the state he shares with Alice just before the global operation $\mathcal{E}$ acts on the joint system, and that Alice makes a local measurement on her half just after $\mathcal{E}$ acts. If Alice's measurement can acquire any information about what operation was applied by Bob, then we say that Bob can signal Alice. The operation is causal if no such signaling is possible in either direction.

Entanglement shared by Alice and Bob cannot be used to send a superluminal signal from Alice to Bob or from Bob to Alice. Therefore, any localizable superoperator is surely a causal superoperator. What about the converse? It might seem reasonable to expect that *any* operation, if it respects the principle that information cannot propagate outside the forward light cone, should be physically realizable in prin-

*Email address: beckman@theory.caltech.edu
†Email address: gottesma@eecs.berkeley.edu
‡Email address: nielsen@physics.uq.edu.au
§Email address: preskill@theory.caltech.edu

ciple. However, we will show otherwise by exhibiting some examples of superoperators that are causal but not localizable.

We obtain weaker notions of localizability and causality by considering communication in just one direction. We say that a superoperator is *semilocalizable* if it is possible to implement it with *one-way* quantum communication from Alice to Bob. Such operations are physically realizable if Bob's action takes place in the future light cone of Alice's action. Similarly, we say that an operation is *semicausal* if it does not allow Bob to send a signal to Alice. Obviously, a semilocalizable superoperator is semicausal— communication from Alice to Bob cannot help Bob to send a message to Alice. What about the converse? If one believes that causality is a very special property of operations that is not likely to hold "by accident," then it is natural to formulate the following conjecture, suggested to us by DiVincenzo [4].

Conjecture. Every semicausal superoperator is semilocalizable.

We will prove this conjecture for the special case of complete orthogonal measurement superoperators. Whether it holds in general remains an open question.

The problem of characterizing what measurements are possible was raised by Dirac [5], and interesting examples of impossible measurements were pointed out in Refs. [6–8]. That relativistic causality may restrict the operators that can be measured in a field theory was first emphasized by Landau and Peierls [9] (though their particular concerns were well answered by Bohr and Rosenfeld [10]). More recently, these restrictions have been noted by a variety of authors [11–16]. In [17], we have addressed some particular causality issues that arise in non-Abelian gauge theories.

To apply our notion of localizability to quantum-field theory, we must adopt the convenient fiction that the entangled ancilla is an external probe not itself described by the field theory, and that its local coupling to the fields is completely adjustable. This idealization is highly questionable in a quantum theory of gravity, and even for quantum-field theory on flat space-time it is open to criticism. In particular, field variables in spatially adjacent regions are inevitably entangled [3], so that no strict separation between field and ancilla variables is really possible. On the other hand, if the probe variables are "heavy" with rapidly decaying correlations and the field variables are "light," then our idealization is credible and worthy of study.

Should the conjecture that semicausality implies semilocalizability prove to be true, then we will have a general and powerful criterion for deciding if a superoperator can be executed with one-way communication. Even so, we will lack a fully satisfactory way of characterizing the observables of a relativistic quantum theory, as no communication is possible if an operation is carried out on a spacelike slice. The existence of causal quantum operations that are not localizable establishes a perplexing gap between what is *causal* and what is *local*.

In Sec. II, we formulate precise definitions of causal, semicausal, localizable, and semilocalizable, and we point out a large class of localizable superoperators characterized by local stabilizer groups. We describe the general structure of semicausal and causal complete-measurement superoperators in Sec. III. We show that semicausal complete-measurement superoperators are semilocalizable in Sec. IV, and exhibit some causal complete-measurement superoperators that are not localizable in Sec. V. In Sec. VI, we exploit bounds on the strength of quantum correlations to construct another class of causal superoperators that are not localizable and we note a connection between localizability and quantum communication complexity. We prove in Sec. VII that a semicausal unitary transformation must be a tensor product. Some further criteria for semicausality are developed in Sec. VIII and Sec. IX contains some concluding comments. Proofs of two of our theorems are included as appendices.

II. CAUSALITY AND LOCALIZABILITY

In this section, we formally define the properties of quantum operations that we wish to explore—causality, semicausality, localizability, and semilocalizability—and we discuss some examples that illustrate these concepts.

A. Causality

Any permissible way in which the state of a quantum system can change is described by a quantum operation, a completely positive trace-nonincreasing linear map of density operators to density operators. An important special case is a trace-preserving map, or *superoperator*. A superoperator $\mathcal{E}$ can be interpreted as a generalized measurement with an unknown outcome; its action on a density operator ρ has an operator-sum representation

$$\mathcal{E}(\rho) = \sum_{\mu} M_{\mu} \rho M_{\mu}^{\dagger}, \tag{1}$$

where the operation elements M_{μ} obey the normalization condition

$$\sum_{\mu} M_{\mu}^{\dagger} M_{\mu} = I. \tag{2}$$

An operation is a generalized measurement in which a particular outcome has been selected, but the density operator has not been renormalized. It too can be represented as in Eq. (1), but where the sum over μ is restricted to a subset of a set of operators obeying Eq. (2)—that is, the eigenvalues of $\Sigma_{\mu} M_{\mu}^{\dagger} M_{\mu}$ are no greater than 1. For a general operation, $\mathrm{tr}\,\mathcal{E}(\rho)$ can be interpreted as the probability of the observed outcome.

Every superoperator has a unitary representation. To implement the superoperator $\mathcal{E}_S$ acting on Hilbert space $\mathcal{H}_S$, we can introduce an ancilla with Hilbert space $\mathcal{H}_R$, prepare a pure state $|\psi\rangle \in \mathcal{H}_R$ of the ancilla, perform a unitary transformation U on $\mathcal{H}_S \otimes \mathcal{H}_R$, and then discard the ancilla,

$$\mathcal{E}_S(\rho_S) = \mathrm{tr}_R[U(\rho_S \otimes |\psi\rangle_{R\,R}\langle\psi|)U^{\dagger}]. \tag{3}$$

A general operation has a similar representation, except that after U is applied, an (not necessarily complete) orthogonal
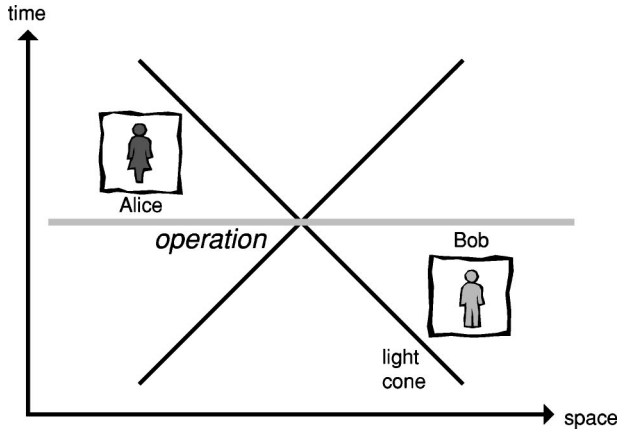
FIG. 1. An operation on a time slice. If the operation allows spacelike-separated Alice and Bob to communicate, then it is not *causal* and hence not physically implementable.

measurement is carried out on the ancilla and a particular result is selected (without renormalizing the density operator).

But what quantum operations are physically possible? The general answer is not known, but it is known that many operations of the form Eq. (1) are unphysical because they run afoul of relativistic causality [11–16]. Consider, as in Fig. 1, two parties Alice and Bob who perform spacelike-separated actions. Just prior to the implementation of the operation $\mathcal{E}$, Bob performs a local operation on the degrees of freedom in his vicinity, and just after the implementation of $\mathcal{E}$, Alice performs a local measurement of the degrees of freedom in her vicinity. If Alice is able to acquire any information about what local operation Bob chose to apply, then Bob has successfully sent a superluminal signal to Alice.

If a bipartite operation $\mathcal{E}$ does not enable such superluminal signaling from Bob to Alice, then we will say that $\mathcal{E}$ is *semicausal*. If $\mathcal{E}$ does not allow signaling in either direction, we will say that $\mathcal{E}$ is *causal*. (In our discussions of semicausality, we will normally adopt the convention that Bob is the party attempting to send the signal and Alice is the party attempting to receive it. This somewhat perverse convention is chosen in order to be consistent with the definition of semilocalizable that is introduced below.) An operation that is not semicausal is said to be *acausal*. A causal operation will sometimes be called *fully causal*, when we wish to emphasize the contrast with semicausality.

Alice's part of the bipartite system will be called "system A," with Hilbert space $\mathcal{H}_A$, and Bob's half is "system B" with Hilbert space $\mathcal{H}_B$. We consider a quantum operation $\mathcal{E}$ that acts on states in $\mathcal{H}_A \otimes \mathcal{H}_B$. Alice might have under her control not only A, but also an ancilla system R with Hilbert space $\mathcal{H}_R$, and Bob might control an ancilla system S with Hilbert space $\mathcal{H}_S$. Suppose that Bob wants to attempt to exploit the operation $\mathcal{E}$ to send a signal to Alice. Alice and Bob could share an initial density operator ρ_{RABS} living in $\mathcal{H}_R \otimes \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_S$, and Bob could apply a superoperator $\mathcal{B}_{BS}$ to his half of the state. (Bob is restricted to a superoperator, rather than a trace-decreasing operation, because Alice is unaware of the outcome of any measurement per-

formed by Bob.) Then after the operation $\mathcal{E}$ is applied, we obtain Alice's final density operator by tracing over Bob's system and ancilla

$$\rho_{RA} = \mathrm{tr}_{BS}[\mathcal{E}_{AB}(I_{RA} \otimes \mathcal{B}_{BS}(\rho_{RABS}))]. \quad (4)$$

Here by $\mathcal{E}_{AB}$ we mean the operation $I_R \otimes \mathcal{E}_{AB} \otimes I_S$ that acts trivially on the ancillas. Note that if $\mathcal{E}$ is not trace preserving, then ρ_{RA} might not be normalized, in which case Alice's final state is the normalized density operator $\rho_{RA}/\mathrm{tr}\,\rho_{RA}$. Finally, Alice performs a measurement of this state.

If Alice's state depends at all on the superoperator $\mathcal{B}$ that Bob applies, then the mutual information of Bob's superoperator and Alice's measurement is nonzero. Hence Bob can transmit classical information to Alice over a noisy channel with nonzero capacity; that is, Bob can signal Alice. We arrive then, at the following.

Definition. A bipartite operation $\mathcal{E}$ is semicausal if and only if $(\rho_{RA}/\mathrm{tr}\,\rho_{RA})$ is independent of Bob's superoperator $\mathcal{B}$ for all possible choices of the shared initial state ρ_{RABS}.

(Excluded from consideration is the case $\rho_{RA} = 0$, corresponding to an outcome that occurs with probability zero.)

This criterion for semicausality is rather unwieldy; fortunately it can be simplified. One useful observation is that, while in our definition of semicausality we allowed the initial state ρ_{RABS} shared by Alice and Bob to be entangled, we could without loss of generality restrict their initial state to be a product state.

Suppose that a *superoperator $\mathcal{E}$* is not semicausal. Then there is an initial state ρ_{RABS} shared by Alice and Bob, and a superoperator $\mathcal{B}$ that can be applied by Bob, such that

$$\mathrm{tr}_{BS}[\mathcal{E}_{AB}(\rho_{RABS})] \neq \mathrm{tr}_{BS}[\mathcal{E}_{AB}(I_{RA} \otimes \mathcal{B}_{BS}(\rho_{RABS}))]. \quad (5)$$

Now any bipartite density operator can be expanded as

$$\rho_{RABS} = \sum_\mu \lambda_\mu \rho_\mu \otimes \sigma_\mu, \quad (6)$$

where ρ_μ and σ_μ are density operators of Alice's and Bob's systems (including ancillas), respectively, and the λ_μ's are nonvanishing real numbers. Of course, if the λ_μ's were all positive, then ρ_{RABS} would be a separable state. But if we allow the λ_μ's to be negative, then such an expansion exists for any state.

Since the superoperator $\mathcal{E}$ is linear, we may rewrite Eq. (5) as

$$\sum_\mu \lambda_\mu \mathrm{tr}_{BS}[\mathcal{E}(\rho_\mu \otimes \sigma_\mu)] \neq \sum_\mu \lambda_\mu \mathrm{tr}_{BS}[\mathcal{E}(\rho_\mu \otimes \mathcal{B}(\sigma_\mu))], \quad (7)$$

which can be satisfied only if

$$\mathrm{tr}_{BS}[\mathcal{E}(\rho_\mu \otimes \sigma_\mu)] \neq \mathrm{tr}_{BS}[\mathcal{E}(\rho_\mu \otimes \mathcal{B}(\sigma_\mu))], \quad (8)$$

for at least one μ. Therefore, if Alice and Bob prepare the appropriate product state $\rho_\mu \otimes \sigma_\mu$, $\mathcal{E}$ allows Bob to signal Alice. Furthermore, since each of ρ_μ and σ_μ and $\mathcal{B}(\sigma_\mu)$

can be realized as an ensemble of pure states, there is a signaling protocol in which Alice's and Bob's initial states are pure.

Once we recognize that there is a signaling protocol such that the initial state is a product state, we can see that the ancillas are superfluous. Bob sends his signal by choosing one of the two pure states $|\psi\rangle_{BS}$, $|\psi'\rangle_{BS}$; since tracing over S commutes with $\mathcal{E}$, we can just as well say that Bob starts with a mixed state ρ_B or ρ'_B of system B alone. Furthermore, if Bob can signal Alice by preparing one of ρ_B, ρ'_B, then he must be able to do it by preparing pure states that arise in the ensemble realizations of these density operators.

Finally, if signaling is possible, then Alice can receive the signal by preparing an initial pure state $|\varphi\rangle_{RA}$. Bob's action, together with $\mathcal{E}$, subjects system A to one of two possible operations, resulting in two distinguishable final states. But for these final states of RA to be distinguishable, the two operations must produce different outcomes acting on at least one of the *pure* states of A appearing in the Schmidt decomposition of $|\varphi\rangle_{RA}$. Therefore, Alice could just as well dispense with R, and prepare an initial pure state $|\varphi\rangle_A$ of A alone.

Thus we have proved:

Theorem 1. If the bipartite superoperator $\mathcal{E}$ is not semicausal, then signaling is possible with pure initial states and without ancillas: there are pure states $|\psi\rangle_B, |\psi'\rangle_B \in \mathcal{H}_B$ and $|\varphi\rangle_A \in \mathcal{H}_A$ such that

$$\mathrm{tr}_B[\mathcal{E}((|\varphi\rangle\langle\varphi|)_A \otimes (|\psi\rangle\langle\psi|)_B)]$$
$$\neq \mathrm{tr}_B[\mathcal{E}((|\varphi\rangle\langle\varphi|)_A \otimes (|\psi'\rangle\langle\psi'|)_B)]. \quad (9)$$

We note that semicausal superoperators form a convex set. If each $\mathcal{E}_a$ is a superoperator, then so is the combination

$$\mathcal{E} = \sum_a p_a \mathcal{E}_a, \quad (10)$$

where the p_a's are nonnegative and sum to 1. It follows from the linearity of the $\mathcal{E}_a$'s and the definition of semicausality that $\mathcal{E}$ is semicausal if each $\mathcal{E}_a$ is.

A somewhat less obvious property is that the semicausal superoperators form a semigroup—a composition of semicausal operations is semicausal. This follows from:

Theorem 2. Suppose that $\mathcal{E}$ is a semicausal bipartite superoperator, and that the two bipartite density operators ρ and σ satisfy $\mathrm{tr}_B\rho = \mathrm{tr}_B\sigma$. Then $\mathrm{tr}_B\mathcal{E}(\rho) = \mathrm{tr}_B\mathcal{E}(\sigma)$.

Proof. First we note that it is possible to choose a basis of linearly independent operators acting on a Hilbert space $\mathcal{H}$, such that each element of the basis is a one-dimensional projector. Let $\{P_\mu\}$ denote such a basis for Alice's Hilbert space $\mathcal{H}_A$ and let $\{Q_\mu\}$ denote such a basis for Bob's Hilbert space $\mathcal{H}_B$. Then we may expand the bipartite density operators ρ and σ as

$$\rho = \sum_{\mu\nu} \alpha_{\mu\nu} P_\mu \otimes Q_\nu,$$

$$\sigma = \sum_{\mu\nu} \alpha'_{\mu\nu} P_\mu \otimes Q_\nu. \quad (11)$$

Since $\mathrm{tr}\, Q_\nu = 1$, the property $\mathrm{tr}_B\rho = \mathrm{tr}_B\sigma$ can be rewritten as $\sum_{\mu\nu}\alpha_{\mu\nu}P_\mu = \sum_{\mu\nu}\alpha'_{\mu\nu}P_\mu$, which implies, since the P_μ's are a basis, that

$$\sum_\nu \alpha_{\mu\nu} = \sum_\nu \alpha'_{\mu\nu}, \quad (12)$$

for each μ.

Applying the superoperator $\mathcal{E}$, we find that

$$\mathcal{E}(\rho) = \sum_{\mu\nu} \alpha_{\mu\nu} \mathcal{E}(P_\mu \otimes Q_\nu),$$

$$\mathcal{E}(\sigma) = \sum_{\mu\nu} \alpha'_{\mu\nu} \mathcal{E}(P_\mu \otimes Q_\nu). \quad (13)$$

Furthermore, since each Q_ν is a pure state, a unitary transformation applied by Bob can transform any one of the Q_ν's to any other; therefore the semicausality of $\mathcal{E}$ implies that the operator $\mathrm{tr}_B\mathcal{E}(P_\mu \otimes Q_\nu)$ is independent of ν. Denoting this operator by R_μ, we have

$$\mathrm{tr}_B\mathcal{E}(\rho) = \sum_{\mu\nu} \alpha_{\mu\nu} R_\mu,$$

$$\mathrm{tr}_B\mathcal{E}(\sigma) = \sum_{\mu\nu} \alpha'_{\mu\nu} R_\mu; \quad (14)$$

Equation (12) then implies that $\mathrm{tr}_B\mathcal{E}(\rho) = \mathrm{tr}_B\mathcal{E}(\sigma)$. This completes the proof of Theorem 2.

The semigroup property of semicausal superoperators is a simple corollary as follows.

Corollary If $\mathcal{E}_1$ and $\mathcal{E}_2$ are semicausal bipartite superoperators, then their composition $\mathcal{E}_2 \bigcirc \mathcal{E}_1$ is also semicausal.

Proof. Suppose that the bipartite density operator ρ can be transformed to σ by a superoperator applied by Bob. Then the semicausality of $\mathcal{E}_1$ implies that $\mathrm{tr}_B\mathcal{E}_1(\rho) = \mathrm{tr}_B\mathcal{E}_1(\sigma)$, and Theorem 2 applied to $\mathcal{E}_2$ implies that $\mathrm{tr}_B\mathcal{E}_2(\mathcal{E}_1(\rho)) = \mathrm{tr}_B\mathcal{E}_2(\mathcal{E}_1(\sigma))$; therefore $\mathcal{E}_2 \bigcirc \mathcal{E}_1$ is semicausal.

B. Localizability

Physics is local. If a physical system has many parts that are remote from one another, then the evolution of the system is governed by local "parties" that act on the different parts of the system separately. In particular, since communication outside the light cone is impossible, the operations that can be applied to a physical system at a fixed time are those that require no communication among the local parties. We call an operation of this type *localizable*.

Although they are not permitted to communicate, the parties are free to exploit any resources that might have been

prepared in advance. In particular, they are permitted to have a shared ancilla that might be in an entangled quantum state, and to consume their shared entanglement in the course of executing their operation.

In the case of a system with two parts, one controlled by Alice, the other by Bob, these considerations motivate the following definition.

Definition. A bipartite superoperator $\mathcal{E}$ is *localizable* if and only if

$$\mathcal{E}(\rho_{AB}) = \mathrm{tr}_{RS}[\mathcal{A}_{RA} \otimes \mathcal{B}_{BS}(\rho_{AB} \otimes \rho_{RS})] \qquad (15)$$

for some shared ancilla state ρ_{RS} and local superoperators $\mathcal{A}_{RA}$, $\mathcal{B}_{BS}$.

In fact, by extending the ancilla, the state ρ_{RS} can be "purified" and the local superoperators can be replaced by unitary transformations; thus without loss of generality we may use instead the following definition.

Definition. A bipartite superoperator $\mathcal{E}$ is *localizable* if and only if

$$\mathcal{E}(\rho_{AB}) = \mathrm{tr}_{RS}[U_{RA} \otimes V_{BS}(\rho_{AB} \otimes \rho_{RS}) U_{RA}^\dagger \otimes V_{BS}^\dagger] \quad (16)$$

for some shared ancilla pure state ρ_{RS} and local unitary transformations U_{RA} and V_{BS}.

Localizable superoperators form a convex set. To see this, we note that with shared entanglement, Alice and Bob can simulate shared randomness (a weaker resource). For example, suppose they share an ancilla prepared in the state

$$|\Phi\rangle_{RS} = \sum_a \sqrt{p_a} |a\rangle_R \otimes |a\rangle_S, \qquad (17)$$

where $\{|a\rangle_R\}$ is an orthonormal basis for Alice's Hilbert space $\mathcal{H}_R$, $\{|a\rangle_S\}$ is an orthonormal basis for Bob's Hilbert space $\mathcal{H}_S$, and the p_a's are non-negative real numbers that sum to 1. Then if Alice and Bob both perform measurements that project onto these bases, each obtains the outcome $|a\rangle$ with probability p_a. Now let $\{\mathcal{E}_a\}$ be a set of localizable operations. Alice and Bob can consult their shared randomness, and then carry out a local protocol that applies the operation $\mathcal{E}_a$ with probability p_a, thus achieving a local implementation of the convex sum $\Sigma_a p_a \mathcal{E}_a$.

Of course, a superoperator is surely localizable if it is a tensor product of superoperators applied by Alice and by Bob, $\mathcal{E} = \mathcal{E}_A \otimes \mathcal{E}_B$. By convexity, any superoperator of the form

$$\mathcal{E} = \sum_a p_a \mathcal{E}_{A,a} \otimes \mathcal{E}_{B,a} \qquad (18)$$

is also localizable. There are some less obvious examples of localizable operations, as we will soon see.

We are also interested in bipartite operations that can be implemented with communication in just one direction. We call such operations *semilocalizable*. In our discussions of semilocalizability, we will normally adopt the convention that Alice is permitted to send quantum information to Bob, but Bob cannot send anything to Alice. A semilocalizable

operation is one that can be performed in principle if Bob's action is in the forward light cone of Alice's action.

We could, equivalently, provide Alice and Bob with prior shared entanglement, and restrict them to classical communication. These two notions of semilocalizability are equivalent because prior entanglement and classical communication from Alice to Bob enable Alice to teleport quantum information to Bob. Conversely, if Alice can send qubits to Bob, she can establish shared entanglement with him, and she can send him classical messages.

An operation that is not semilocalizable is said to be *unlocalizable*. A localizable operation will sometimes be called *fully localizable*, when we wish to emphasize the contrast with semilocalizability.

If Alice can send quantum information to Bob, then Alice and Bob both have access to the same ancilla: Alice performs a local operation on the ancilla and her half of the shared state, she sends the ancilla to Bob, and then Bob performs a local operation on the ancilla and his half of the state. Thus we arrive at the following.

Definition. A bipartite operation $\mathcal{E}$ is *semilocalizable* if and only if

$$\mathcal{E}(\rho_{AB}) = \mathrm{tr}_R[(\mathcal{B}_{BR} \bigcirc \mathcal{A}_{RA})(\rho_{AB} \otimes \rho_R)] \qquad (19)$$

for some ancilla state ρ_R, where $\mathcal{A}_{RA}$ is an operation and $\mathcal{B}_{BR}$ is a superoperator.

Note that the product $\mathcal{B}_{BR} \bigcirc \mathcal{A}_{RA}$ is a composition of operations (with Alice's operation acting first), not a tensor product; the operations do not commute because they act on the same ancilla. We have allowed Alice to apply an operation that is not necessarily trace preserving, since Alice can perform a measurement whose outcome is known to both Alice and Bob, but Bob is restricted to a superoperator because the outcome of a measurement that he performs is not known by Alice. If the operation $\mathcal{E}$ is a superoperator, then so must be $\mathcal{A}_{RA}$, and in fact we can take $\mathcal{B}_{BR}$ and $\mathcal{A}_{RA}$ to be unitary transformations without loss of generality.

An obvious consequence of this definition is that semilocalizable (or localizable) superoperators form a semigroup: $\mathcal{E}_2 \bigcirc \mathcal{E}_1$ is semilocalizable if $\mathcal{E}_1$ and $\mathcal{E}_2$ are both semilocalizable.

C. Orthogonal measurement superoperators

One of our goals is to characterize the *observables* of a relativistic quantum theory: what self-adjoint operators are really measurable?

When we speak of a "measurement" of an observable whose support is on a time slice, we need not require that the measurement outcome be instantaneously known by anyone. We might imagine instead that many parties distributed over the slice perform simultaneous local operations. Later the data collected by the parties can be assembled and processed at a central location to determine the measurement result.

Then we may say that the operation performed on the slice is a measurement with an unknown outcome. If $\{E_a\}$ is the set of orthogonal projectors onto the eigenspaces of the observable, the effect of this operation on a density operator ρ is

$$\rho \rightarrow \mathcal{E}(\rho) = \sum_a E_a \rho E_a. \qquad (20)$$

We will call an (trace-preserving) operation of this form an *orthogonal measurement superoperator*. In the special case where each projector E_a is one dimensional, it is a *complete orthogonal measurement superoperator*, or just a complete-measurement superoperator. The causality and localizability properties of complete-measurement superoperators will be extensively discussed in the next few sections.

First, let us clarify the concept of semicausality by pointing out an example, noted by Sorkin [14], of an incomplete-measurement superoperator that is not semicausal. It is a two-outcome incomplete Bell measurement performed on a pair of qubits. The orthogonal projectors corresponding to the two outcomes are

$$E_1 = |\phi^+\rangle\langle\phi^+|,$$

$$E_2 = I - |\phi^+\rangle\langle\phi^+|, \qquad (21)$$

where $|\phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$. Suppose that the initial pure state shared by Alice and Bob is $|01\rangle_{AB}$. This state is orthogonal to $|\phi^+\rangle$, so that outcome 2 occurs with probability one, and the state is unmodified by the superoperator. Afterwards Alice still has a density operator $\rho_A = |0\rangle\langle0|$.

But what if, before the superoperator acts, Bob performs a unitary that rotates the state to $|00\rangle_{AB}$? Since this state is an equally weighted superposition of $|\phi^+\rangle$ and $|\phi^-\rangle = (|00\rangle - |11\rangle)/\sqrt{2}$, the two outcomes occur equiprobably, and in either case the final state is maximally entangled, so that Alice's density operator afterwards is $\rho_A = I/2$, where I denotes the identity. Alice can make a measurement that has a good chance of distinguishing the density operators $|0\rangle\langle0|$ and $I/2$, so that she can decipher a message sent by Bob. By a similar method, Alice can send a signal to Bob. The measurement superoperator is acausal.

On the other hand, some orthogonal measurement superoperators are causal. For example, measurement of a tensor product observable $A \otimes B$ is obviously causal—Alice and Bob can induce decoherence in the basis of eigenstates of a tensor product through only local actions. But there are other examples of causal measurement superoperators that are a bit less obvious. One is complete Bell measurement, i.e., decoherence in the Bell basis

$$|\phi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle),$$

$$|\psi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle). \qquad (22)$$

No matter what Bob does, the shared state after Bell measurement is maximally entangled, so that Alice always has $\rho_B = I/2$, and she cannot extract any information about Bob's activities.

Though Bell measurement is a causal operation, it is not something that Alice and Bob can achieve locally without

additional resources. But the Bell measurement superoperator is localizable—Alice and Bob can implement it if they share an entangled ancilla. In fact, shared randomness, a weaker resource than entanglement, is sufficient for this purpose [18]. Suppose that Alice and Bob share a pair of qubits, and also share a string of random bits. At a particular time, they both consult two bits of the random string; depending on whether they read 00, 01, 10, or 11, they both apply the unitary operator I, X, Z, or Y, where I is the identity, and $\{X, Y, Z\}$ are the 2×2 Pauli matrices

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (23)$$

Together, then, Alice and Bob apply the superoperator

$$\mathcal{E}(\rho) = \frac{1}{4}[(I \otimes I)\rho(I \otimes I) + (X \otimes X)\rho(X \otimes X)$$

$$+ (Y \otimes Y)\rho(Y \otimes Y) + (Z \otimes Z)\rho(Z \otimes Z)]. \quad (24)$$

The four Bell states are simultaneous eigenstates of $X \otimes X$ and $Z \otimes Z$ (and therefore also $Y \otimes Y = -(X \otimes X)(Z \otimes Z)$) with eigenvalues ± 1: $Z \otimes Z$ specifies a parity bit that distinguishes ϕ from ψ and $X \otimes X$ specifies a phase bit that distinguishes $+$ from $-$. Hence we easily verify that $\mathcal{E}$ preserves each of the four Bell basis states and annihilates all the terms in ρ that are off the diagonal in the Bell basis.

The Bell measurement superoperator can be viewed as a refinement, or "completion," of the acausal incomplete-measurement superoperator of Eq. (21)—that is, Bell measurement is obtained by resolving the three-dimensional projector E_2 of Eq. (21) into a sum of three mutually orthogonal one-dimensional projectors. Thus, this example illustrates that a completion of an acausal measurement superoperator can sometimes be causal. On the other hand, there are other ways of refining the superoperator of Eq. (21) that yield acausal complete-measurement superoperators. For example, the two-qubit superoperator with projectors

$$E_1 = |\phi^+\rangle\langle\phi^+|,$$

$$E_2 = |\phi^-\rangle\langle\phi^-|,$$

$$E_3 = |01\rangle\langle01|,$$

$$E_4 = |10\rangle\langle10|, \qquad (25)$$

is easily seen to be acausal by applying the criterion of Theorem 3 below. In fact, it is a general feature that if an orthogonal measurement superoperator $\mathcal{E}$ allows Bob to signal Alice, then there exists a completion of $\mathcal{E}$ that also allows signaling, with the same signal states $|\psi\rangle_B$ and $|\psi'\rangle_B$. This result is proved in Ref. [16].

Since the Bell measurement superoperator can be implemented with shared randomness, one may wonder whether shared randomness is sufficient for the implementation of arbitrary localizable superoperators. But it is easy to think of localizable superoperators for which shared randomness does not suffice—shared entanglement is necessary. For example,
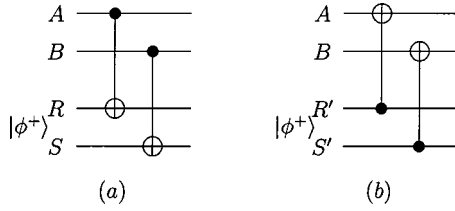
FIG. 2. Local implementation of the Bell measurement superoperator using shared entanglement. In (a), the two controlled-NOT gates imprint the parity bit of AB onto the ancilla RS. In (b), the two controlled-NOT gates imprint the phase bit of AB onto the ancilla $R'S'$. Tracing over the ancillas in the Bell basis, we find that the AB pair decoheres in the Bell basis.

Alice and Bob can locally perform a two-qubit operation in which they throw their qubits away, and replace them with a $|\phi^+\rangle$ drawn from their shared ancilla. This operation can turn a product state into an entangled state, which would be impossible with local operations and shared randomness alone. We will discuss another example of a localizable superoperator that cannot be implemented with shared randomness in Sec. VI.

We note in passing that if Alice and Bob have ancilla pairs prepared in the state $|\phi^+\rangle_{RS}$ (where R denotes Alice's ancilla qubit, and S denotes Bob's), then they can implement the Bell measurement superoperator by executing the quantum circuits shown in Fig. 2 and discarding their ancilla qubits. The circuit in Fig. 2(a) flips the ancilla pair from $|\phi^+\rangle$ to $|\psi^+\rangle$ if the parity bit of the AB state is $Z_A \otimes Z_B = -1$, and the circuit in Fig. 2(b) flips the ancilla pair from $|\phi^+\rangle$ to $|\phi^-\rangle$ if the phase bit of the AB state is $X_A \otimes X_B = -1$. Thus, the values of the parity and phase bits of the AB pair, and only this information, become imprinted on the ancilla pairs. Tracing over the ancilla pairs in the Bell basis, then, induces decoherence in the AB Bell basis.

In fact, if Alice and Bob share entanglement and perform the circuits of Fig. 2, they can execute Bell measurement, including postselection, on the AB pair by measuring their ancilla qubits and broadcasting their results classically. After executing circuit Fig. 2(a), Alice measures Z_R and Bob measures Z_S. Then the parity bit of the AB pair is the parity of these measurement outcomes. After executing circuit Fig. 2(b), Alice measures $X_{R'}$ and Bob measures $X_{S'}$. Then the phase bit of the AB pair is the parity of the measurement outcomes. This example is instructive, as it reminds us again that entanglement is a more powerful resource than shared randomness. If Alice and Bob were limited to shared randomness and classical communication, they would be unable to create entanglement, and so would lack the capability of doing Bell measurement with postselection on their shared qubit pair.

If Alice and Bob did not have entanglement to start with, they would still be able to perform Bell measurement with postselection on their shared pair if they could send ancilla qubits to a central laboratory for later quantum processing, as illustrated in Fig. 3. Here Alice entangles her qubit A first with ancilla qubit R and then with R', while Bob entangles his qubit B first with ancilla qubit S and then with S'. The ancilla qubits are collected, and a Bell measurement is per-
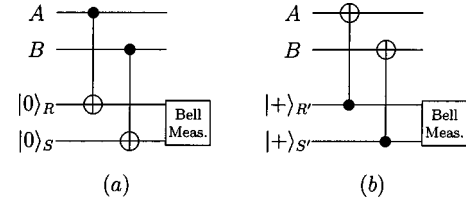

FIG. 3. Bell measurement through entanglement swapping. Alice performs local controlled-NOT gates on her qubit A and the ancilla qubits RR'. Bob performs local controlled-NOT gates on his qubit B and the ancilla qubits SS'. Later, the ancilla qubits are collected, and Bell measurement is performed on the pairs RS and $R'S'$. The Bell measurements on the ancilla realize Bell measurement on AB, by "swapping" entanglement of AR, AR', BS, BS' for entanglement of AB, RS, and $R'S'$.

formed on RS and $R'S'$. The RS measurement yields the correct parity bit of the AB pair, and a random phase bit; if the measured RS phase bit is -1, then the phase bit of the AB state is flipped. The $R'S'$ measurement yields the "correct" phase bit (possibly flipped by the RS measurement), and a random parity bit; if the measured $R'S'$ parity bit is -1, then the parity bit of the AB state is flipped. Taken together, the two Bell measurements on the ancilla qubits act as a projective measurement of $\mathcal{H}_{AB}$ onto the Bell basis, followed by one of the transformations $I \otimes I$, $X \otimes X$, $Y \otimes Y$, or $Z \otimes Z$; which of the four transformations has been applied, as well as the identity of the AB Bell state resulting from the projection, can be inferred from outcomes of the RS and $R'S'$ measurements. If the AB state is initially a product state, Bell measurement of the ancilla qubits establishes entanglement of A with B by "swapping" AR and BS entanglement for AB and RS entanglement [19].

This example is also instructive. It reminds us that decoherence induced on a time slice can sometimes be reversed later through the operation of a "quantum eraser" [20]. If we were to trace out the ancilla qubits right after applying the controlled-NOT operations of Fig. 3 (before the Bell measurement), then Fig 3(a) would induce decoherence, not in an entangled basis, but rather in the product basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$; similarly Fig. 3(b) would induce decoherence in the product basis $\{|++\rangle, |+-\rangle, |-+\rangle, |--\rangle\}$, where

$$|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle). \tag{26}$$

As one would expect, without entanglement or shared randomness, Alice and Bob are unable to implement decoherence in the Bell basis with their local operations alone. But a later measurement *including postselection* can "undo" the decoherence in the product basis and establish decoherence in the Bell basis instead.

D. Local stabilizers

The observation that shared randomness is sufficient to induce decoherence in the Bell basis can be substantially generalized. Consider a superoperator $\mathcal{E}$ that acts on a density operator ρ as

$$\mathcal{E}(\rho) = \frac{1}{|G|} \sum_{g \in G} U(g) \rho U(g)^{\dagger}, \tag{27}$$

where the $U(g)$'s provide a (not necessarily irreducible) unitary representation of the group G, and $|G|$ denotes the order of G. The Hilbert space $\mathcal{H}$ in which ρ resides can be decomposed into spaces that transform irreducibly under the group G. Let us choose an orthonormal basis

$$\{|R,a,i\rangle\}; \tag{28}$$

here R labels the irreducible representations of G, a labels the sectors of $\mathcal{H}$ that transform as the irreducible representation R (a particular irreducible representation might occur multiple times), $i = 1, 2, \ldots, n_R$ labels states of a basis for the vector space on which R acts, and n_R is the dimension of R. Expressed in this basis, the representation $U(g)$ is

$$U(g) = \sum_{R,a,i,j} |R,a,i\rangle D^{(R)}(g)_{ij} \langle R,a,j|, \tag{29}$$

where $D^{(R)}(g)_{ij}$ is a matrix element of the irreducible representation R. These matrix elements obey the orthogonality relations

$$\frac{1}{|G|} \sum_{g \in G} D^{(R)}(g)_{ij} D^{(R')}(g)_{kl}^{*} = \frac{1}{n_R} \delta^{RR'} \delta_{ik} \delta_{jl}. \tag{30}$$

Substituting Eq. (29) into Eq. (27) and applying the orthogonality relations, we find

$$\langle R,a,i|\mathcal{E}(\rho)|R',b,j\rangle = \delta^{RR'} \frac{1}{n_R} \delta_{ij} \sum_{k=1}^{n_R} \langle R,a,k|\rho|R,b,k\rangle. \tag{31}$$

Thus we see that the superoperator $\mathcal{E}$ destroys the coherence of a superposition of states that transform as distinct irreducible representations of G. Within a given irreducible representation, it randomizes the state, replacing the density operator by a multiple of the identity. Some off-diagonal terms in the density operator can survive, if a given irreducible representation occurs in $\mathcal{H}$ more than once.

Now suppose that the Hilbert space $\mathcal{H}$ is shared by n parties; it has a tensor product decomposition

$$\mathcal{H} = \otimes_{\alpha=1}^{n} \mathcal{H}_{\alpha}. \tag{32}$$

And suppose that each element $U(g)$ of the representation of G is a tensor product

$$U(g) = \otimes_{\alpha=1}^{n} U(g)_{\alpha}. \tag{33}$$

Then the n parties can perform the operation $\mathcal{E}$ by consulting their shared randomness—if they are instructed to apply the group element $g \in G$, the party α applies $U(g)_{\alpha}$ to her portion of the state.

In the Bell measurement case discussed above, the four-dimensional Hilbert space of two qubits transforms as the representation

$$\{I \otimes I, X \otimes X, -Y \otimes Y, Z \otimes Z\} \tag{34}$$

of the group $G = Z_2 \times Z_2$. The group G is Abelian, and the four Bell states transform as distinct one-dimensional irreducible representations of G. Therefore, the superoperator $\mathcal{E}$ induces decoherence in the Bell basis.

The same ideas apply to any *stabilizer code* [21,22,1,2]. Consider, for example, an Abelian group G acting on n qubits, generated by $n-k$ operators, where each generator is a tensor product of single-qubit Pauli operators. If each qubit is entrusted to a distinct party, then by consulting shared randomness, the n parties can apply a random element of the group G to their density operator. The superoperator they implement acts trivially on each 2^k-dimensional code space—an eigenspace of the generators with specified eigenvalues. But it destroys the coherence of a superposition of eigenspaces with different eigenvalues. In the notation of Eq. (31), the index R labels the stabilizer eigenvalues, and the indices a,b label the basis states in a code space with a specified value of R. Because the group is Abelian, all of the irreducible representations are one dimensional, and Eq. (31) becomes

$$\mathcal{E}(\rho) = \frac{1}{|G|} \sum_{g \in G} U(g) \rho U(g)^{\dagger}$$

$$= \sum_{R,a,b} |R,a\rangle \langle R,a|\rho|R,b\rangle \langle R,b|$$

$$= \sum_{R} E_R \rho E_R, \tag{35}$$

where E_R projects onto the subspace with specified stabilizer eigenvalues.

The observation that Bell *measurement* (including postselection) can be achieved with shared entanglement, local operations, and broadcasting of classical data can be generalized to any CSS stabilizer code; i.e., any code of the class constructed by Calderbank and Shor [23], and Steane [24]. An n-qubit stabilizer code is of the CSS type if the stabilizer generators can be chosen so that each generator is either a tensor product of Z's and I's or a tensor product of X's and I's. Imagine that each of the n qubits is in the custody of a separate party. Steane [25] has observed that the measurement of the stabilizer generators can always be achieved by carrying out these steps: (1) preparation of a suitable entangled ancilla that is distributed to the n parties, (2) local quantum gates applied by each party, acting on her qubit and her part of the ancilla, (3) local measurements by each party, and (4) classical post processing of the measurement outcomes. In the case of Bell measurement on a pair of qubits, the stabilizer generators are $X \otimes X$, $Z \otimes Z$, and the entangled ancilla state is $|\phi^+\rangle$.

An example of a superoperator associated with a non-Abelian group is the "twirling" operation that transforms a two-qubit state into a Werner state [18]. In that case, the group is A_4, the order-12 subgroup of the rotation group that preserves a tetrahedron. Under A_4, the state $|\psi^-\rangle$ transforms trivially, while the other three Bell states $|\phi^+\rangle, |\psi^+\rangle, |\phi^-\rangle$

transform as a three-dimensional irreducible representation. Two parties Alice and Bob consult their shared randomness and apply a random element of A_4; according to Eq. (31), this operation transforms any initial density operator into a state of the Werner form

$$\rho(F) = F|\psi^-\rangle\langle\psi^-| + \frac{1-F}{3}(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-|$$

$$+ |\psi^+\rangle\langle\psi^+|), \qquad (36)$$

while preserving the fidelity $F = \langle\psi^-|\rho|\psi^-\rangle$. Note that unlike the acausal operation defined by Eq. (21), this localizable operation transforms any initial state into an incoherent mixture of Bell states; hence Alice's final density operator is always $I/2$, and Alice is unable to receive a signal from Bob.

III. CAUSAL AND SEMICAUSAL COMPLETE-MEASUREMENT SUPEROPERATORS

In the next three sections, we will investigate the causality and localizability properties of measurement superoperators that project onto a complete orthonormal basis. We will show that semicausal operations of this class are semilocalizable, and that fully causal operations of this class are not necessarily fully localizable.

Suppose that the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ have dimensions N_A and N_B, respectively, and let $\{E^a = |a\rangle\langle a|, a = 1,2,\ldots,N_A N_B\}$ denote a complete set of orthogonal one-dimensional projectors on $\mathcal{H}_A \otimes \mathcal{H}_B$. By tracing over $\mathcal{H}_B$, we obtain from these projectors $N_A N_B$ positive operators acting on $\mathcal{H}_A$, each with unit trace, defined by

$$\sigma_A^a = \text{tr}_B(E_{AB}^a). \qquad (37)$$

Since the E_{AB}^a's are complete, these operators satisfy

$$\sum_a \sigma_A^a = \text{tr}_B I_{AB} = N_B I_A; \qquad (38)$$

that is, $\{N_B^{-1}\sigma_A^a\}$ is a positive operator-valued measure (POVM) on $\mathcal{H}_A$ with $N_A N_B$ outcomes.

The semicausal complete orthogonal measurement superoperators (those that do not allow Bob to signal Alice) can be simply characterized by a property of the σ_A^a's, thanks to the following theorem.

Theorem 3. A complete orthogonal measurement superoperator is semicausal if and only if it has the following property. For each pair of operators $\{\sigma_A^a, \sigma_A^b\}$, either $\sigma_A^a = \sigma_A^b$ or $\sigma_A^a \sigma_A^b = 0 = \sigma_A^b \sigma_A^a$.

That is, any pair of σ_A^a's must be either identical or orthogonal, if and only if the superoperator is semicausal.

Theorem 3 is proved in Appendix A, but one can readily see that the result is plausible. If $E^a = |a\rangle\langle a|$, then, in order to signal Alice, Bob by acting locally needs to induce a transition from the state $|a\rangle$ to the state $|b\rangle$ for some a and b; furthermore, Alice must be able to detect the difference between $|a\rangle$ and $|b\rangle$. But if σ_A^a and σ_A^b are orthogonal, then Bob is unable to induce the transition, and if $\sigma_A^a = \sigma_A^b$, then
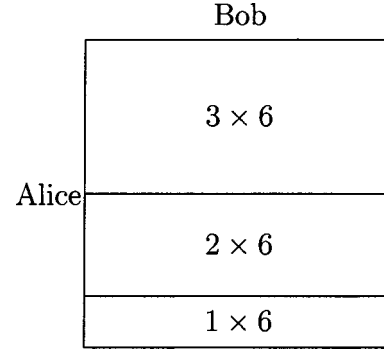

FIG. 4. A semicausal complete orthogonal measurement in 6×6 dimensions. Alice's Hilbert space is decomposed into three mutually orthogonal subspaces, of dimensions 3, 2, and 1. The measurement projects onto an orthonormal basis, where each element of the basis is a maximally entangled state of one of Alice's three subspaces with Bob's space.

Alice cannot tell the difference. On the other hand, if $\sigma_A^a \sigma_A^b = 0$ and $\sigma_A^a \neq \sigma_A^b$, then Bob can induce the transition and Alice can distinguish the states; hence a signaling protocol can be devised.

Applying Theorem 3, we can see that all semicausal complete measurements have a simple structure. Suppose that $\{E^a\}$ is a semicausal complete measurement. For any one of the σ_A^a, let $\mathcal{H}_A^a$ denote the subspace of $\mathcal{H}_A$ on which σ_A^a has its support (σ_A^a is strictly positive on this subspace and vanishes on the orthogonal subspace). According to Theorem 3, the support $\mathcal{H}_A^b$ of each σ_A^b either coincides with $\mathcal{H}_A^a$ or is orthogonal to $\mathcal{H}_A^a$, and in the former case we have $\sigma_A^a = \sigma_A^b$.

Thus each σ_A^b with support on $\mathcal{H}_A^a$ must equal σ_A^a, and furthermore, the sum of the operators with support on this subspace, according to Eq. (38), must be $N_B I_A^a$, where I_A^a denotes the projector onto $\mathcal{H}_A^a$. Therefore each σ_A^a with support on $\mathcal{H}_A^a$ is proportional to I_A^a, and since we also know that $\text{tr}\,\sigma_A^a = 1$, must be

$$\sigma_A^a = \frac{1}{d_A^a} I_A^a, \qquad (39)$$

where d_A^a is the dimension of $\mathcal{H}_A^a$. We also conclude that the number of σ_A^b's with support on $\mathcal{H}_A^a$ is $N_B d_A^a$.

We see that the state $|a\rangle$ is a maximally entangled state of the form

$$|a\rangle = \frac{1}{\sqrt{d_A^a}} \sum_{i=1}^{d_A^a} |i\rangle_A \otimes |i'\rangle_B, \qquad (40)$$

where here $\{|i\rangle_A\}$ denotes an orthonormal basis for $\mathcal{H}_A^a$, and the $|i'\rangle_B$'s are mutually orthogonal states of $\mathcal{H}_B$. The general structure of a semicausal complete-measurement operation, then, is as illustrated in Fig. 4. Alice's Hilbert space can be decomposed into mutually orthogonal subspaces

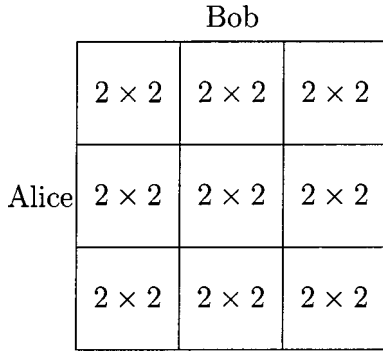$$\mathcal{H}_A = \oplus_\alpha \mathcal{H}_A^\alpha, \qquad (41)$$

Bob

FIG. 5. A causal complete orthogonal measurement in 6×6 dimensions. Alice's six-dimensional Hilbert space is partitioned into three mutually orthogonal subspaces, each of dimension 2, and Bob's Hilbert space is similarly partitioned. The measurement projects onto an orthonormal basis, where each element of the basis is a maximally entangled state of one of Alice's subspaces with one of Bob's subspaces.

where $\mathcal{H}_A^\alpha$ has dimension d_A^α. Of the $N_A N_B$ states comprising the basis $\{|a\rangle\}$, $N_B d_A^\alpha$, all maximally entangled, have support on $\mathcal{H}_A^\alpha$.

A *fully causal* complete measurement is more highly constrained. Since the measurement is semicausal in both directions, both the σ_A^a's and the σ_B^a's obey the conditions specified in Theorem 3. If we choose one particular $\mathcal{H}_A^\alpha$, there are $N_B d_A^\alpha$ elements of the basis with support on this space, and associated with these are $N_B d_A^\alpha \sigma_B^a$'s, all of rank d_A^α, and any two of which must either coincide or be orthogonal. Therefore, the σ_B^a's partition $\mathcal{H}_B$ into mutually orthogonal subspaces, all of dimension d_A^α; it follows that d_A^α must divide N_B, that the number of orthogonal subspaces is N_B / d_A^α, and that $(d_A^\alpha)^2 \sigma_B^a$'s have support on each space.

Applying the same argument again, but with $\mathcal{H}_A$ and $\mathcal{H}_B$ interchanged, we see that the σ_A^a's also partition $\mathcal{H}_A$ into mutually orthogonal subspaces, all of dimension d_A^α. We conclude that a causal complete measurement has the structure illustrated in Fig. 5. Alice's space can be decomposed into r_A subspaces $\mathcal{H}_A^\alpha$, each of dimension d (so that $N_A = r_A d$), and Bob's space can be decomposed into r_B subspaces $\mathcal{H}_B^\beta$, each of dimension d (with $N_B = r_B d$). The measurement projects onto an orthonormal basis, where each element of the basis is a maximally entangled state of some $\mathcal{H}_A^\alpha$ with some $\mathcal{H}_B^\beta$. There are $r_A r_B$ ways to choose α and β, and there are d^2 maximally entangled states for each pair of subspaces.

The extreme cases are $d=1$, for which we have a product basis $\{|\alpha, \beta\rangle_{AB}\}$, and $d = N_A = N_B$, for which the measurement is a projection onto a maximally entangled basis of $\mathcal{H}_{AB}$.

Comparing Fig. 4 and 5 makes it clear that a semicausal measurement need not be fully causal. Indeed, this feature is quite obvious, since transmission of information from Alice to Bob can allow Alice to signal Bob but does not enable Bob to signal Alice. To make this point more explicit, consider the 2×2 example illustrated in Fig. 6. The measurement projects onto the orthonormal basis
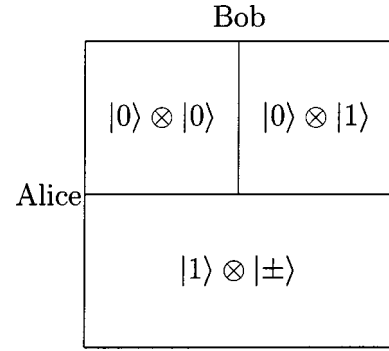
Bob

FIG. 6. A semicausal complete orthogonal measurement in 2×2 dimensions. The orthonormal basis shown partitions Alice's space into mutually orthogonal one-dimensional subspaces; hence Bob cannot signal Alice. But since Bob's space is not so partitioned, Alice can signal Bob.

$$|0\rangle_A \otimes |0\rangle_B, \quad |0\rangle_A \otimes |1\rangle_B,$$
$$|1\rangle_A \otimes |+\rangle_B, \quad |1\rangle_A \otimes |-\rangle_B, \tag{42}$$

where

$$|\pm\rangle = \frac{1}{\sqrt{2}} (|0\rangle \pm |1\rangle). \tag{43}$$

Tracing over Bob's system we obtain Alice's projectors

$$\sigma_A^{00} = \sigma_A^{01} = |0\rangle\langle 0|,$$
$$\sigma_A^{1+} = \sigma_A^{1-} = |1\rangle\langle 1|, \tag{44}$$

which satisfy the criterion of Theorem 3; hence Bob cannot signal Alice. Tracing over Alice's system we obtain Bob's projectors

$$\sigma_B^{00} = |0\rangle\langle 0|, \quad \sigma_B^{01} = |1\rangle\langle 1|,$$
$$\sigma_B^{1+} = |+\rangle\langle +|, \quad \sigma_B^{1-} = |-\rangle\langle -|, \tag{45}$$

which violate the criterion of Theorem 3; hence Alice can signal Bob.

A particular protocol that allows Alice to signal Bob works as follows. Bob prepares his qubit in the state $|0\rangle$ and Alice prepares hers in one of the states $|0\rangle$, $|1\rangle$. After the operation $\mathcal{E}$ is applied, Bob's density operator is

$$\rho_B^0 = |0\rangle\langle 0|, \tag{46}$$

if Alice prepared $|0\rangle$, and his density operator is

$$\rho_B^1 = I/2, \tag{47}$$

if Alice prepared $|1\rangle$. Since $\rho_B^0 \neq \rho_B^1$, we conclude that $\mathcal{E}$ is not causal.

Note that the measurement that projects onto the basis Eq. (42) is obviously semilocalizable; in fact it can be executed with one-way *classical* communication from Alice to Bob. Alice measures in the basis $\{|0\rangle_A, |1\rangle_A\}$, and sends her mea-

surement outcome to Bob. Then Bob measures in the basis $\{|0\rangle_B, |1\rangle_B\}$ if Alice's outcome was $|0\rangle_A$, and in the basis $\{|+\rangle_B, |-\rangle_B\}$ if Alice's outcome was $|1\rangle_A$.

IV. SEMICAUSAL COMPLETE-MEASUREMENT SUPEROPERATORS ARE SEMILOCALIZABLE

We now have learned enough about the structure of semi-causal complete measurements to see that any semicausal complete-measurement superoperator is semilocalizable. Suppose that Alice and Bob share a quantum state, and wish to perform a measurement that projects onto the basis $\{|a\rangle\}$, where each $|a\rangle$ is a maximally entangled state of a subspace of $\mathcal{H}_A$ with a subspace of $\mathcal{H}_B$. Alice can perform a partial measurement that identifies the subspace $\mathcal{H}_A^\alpha$ of $\mathcal{H}_A$, and then send her half of the state to Bob, who can finish the measurement and identify $|a\rangle$. To complete the procedure, they simply replace their original state by a state identical to $|a\rangle$ that can be drawn from their shared ancilla. Since Bob can convert any maximally entangled state of $\mathcal{H}_A^\alpha$ and $\mathcal{H}_B$ into $|a\rangle$ by performing a unitary transformation on $\mathcal{H}_B$, he and Alice can replace the original state by $|a\rangle$ without any further communication. From these observations we obtain the following.

Theorem 4. A semicausal complete orthogonal measurement superoperator is semilocalizable.

Proof. Suppose that Alice and Bob share the state ρ_{AB} in their joint Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. In addition, an ancillary Hilbert space $\mathcal{H}_R \otimes \mathcal{H}_S$, isomorphic to $\mathcal{H}_A \otimes \mathcal{H}_B$, is initially under Alice's control. To implement the semicausal operation

$$\mathcal{E}_{AB}(\rho_{AB}) = \sum_a E_{AB}^a (\rho_{AB}) E_{AB}^a \qquad (48)$$

with one-way quantum communication from Alice to Bob, they proceed as follows: First, Alice performs a partial measurement that projects ρ_{AB} onto her mutually orthogonal subspaces $\mathcal{H}_A^\alpha$, obtaining the outcome

$$\rho_{AB}^\alpha = \frac{E_A^\alpha(\rho_{AB}) E_A^\alpha}{\mathrm{tr}(E_A^\alpha \rho_{AB})} \qquad (49)$$

with probability

$$p_\alpha = \mathrm{tr}(E_A^\alpha \rho_{AB}), \qquad (50)$$

where E_A^α is the projector onto $\mathcal{H}_A^\alpha$. If Alice's measurement outcome is α, she prepares an ancilla state $|\Phi\rangle_{RS} \in \mathcal{H}_R \otimes \mathcal{H}_S$, and a dimension $d_A^\alpha \times d_A^\alpha$ maximally entangled state of $\mathcal{H}_R^\alpha$ with $\mathcal{H}_S$, where $\mathcal{H}_R^\alpha$ is isomorphic to $\mathcal{H}_A^\alpha$. Next, Alice swaps the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_R$, obtaining ρ_{RB}^α and $|\Phi\rangle_{AS}$. She sends ρ_{RB}^α to Bob, along with the S half of the entangled state $|\Phi\rangle_{AS}$.

Upon receipt, Bob swaps the Hilbert spaces $\mathcal{H}_B$ and $\mathcal{H}_S$, so that Alice and Bob now share $|\Phi\rangle_{AB}$, while ρ_{RS}^α is entirely in Bob's hands. On the state ρ_{RS}^α, Bob performs an orthogonal measurement with projectors E_{RS}^a that are isomorphic to the E_{AB}^a's, obtaining the outcome $|a\rangle_{RS}$ with probability $p_{a|\alpha}$, where

$$p_{a,\alpha} = p_{a|\alpha} p_\alpha = \mathrm{tr}(E_{RS}^a E_R^\alpha \rho_{RS} E_R^\alpha E_{RS}^a). \qquad (51)$$

Since Bob's measurement is just a completion of Alice's partial measurement, we have

$$E_{RS}^a E_R^\alpha = E_R^\alpha E_{RS}^a = \delta_{a,\alpha} E_{RS}^a, \qquad (52)$$

where $\delta_{a,\alpha}$ is 1 if σ_R^a has support on $\mathcal{H}_R^\alpha$ and zero otherwise. Therefore, the probability that Bob obtains outcome a can be expressed as

$$p_a = \sum_\alpha p_{a,\alpha} = \mathrm{tr}_{RS}(E_{RS}^a \rho_{RS}) = \mathrm{tr}_{AB}(E_{AB}^a \rho_{AB}). \qquad (53)$$

(Bob's measurement commutes with Alice's, so it is just as though Bob measured first, and Alice has been provided with incomplete information about what Bob found.)

Now, since the state $|a\rangle_{RS}$ prepared by Bob's measurement is a $d_A^\alpha \times d_A^\alpha$ maximally entangled state of $\mathcal{H}_R$ with $\mathcal{H}_S$, Bob can apply a suitable unitary transformation to his half of the state $|\Phi\rangle_{AB}$ that he now shares with Alice, rotating it to the state $|a\rangle_{AB}$. Thus, Alice and Bob have converted their initial state ρ_{AB} to $|a\rangle_{AB}$ with probability $p_a = \langle a|\rho_{AB}|a\rangle$. Finally, Bob discards the ancilla RS, and Alice and Bob discard the record of their measurement outcomes. We have described a protocol with one-way quantum communication that executes the semicausal complete-measurement super-operator $\mathcal{E}_{AB}$. This proves Theorem 4.

Note that if we dispense with the last step, in which Alice and Bob discard their records, then we see that not just the measurement superoperator, but also the measurement operation with postselection, is semilocalizable: with one-way quantum communication from Alice to Bob, the state is projected onto the basis, and the measurement outcome is known by Bob (though not by Alice).

V. CAUSAL COMPLETE-MEASUREMENT SUPEROPERATORS NEED NOT BE LOCALIZABLE

Now that the general structure of causal complete measurements is known, we can address whether causal complete measurements are localizable. In fact, we will be able to construct examples of causal measurements that are provably *not* localizable. To accomplish this task, we will identify a property satisfied by localizable operations, and exhibit causal measurements that do not possess this property.

We say that a (not necessarily normalized) pure state $|\psi\rangle$ is an *eigenstate* of a superoperator $\mathcal{E}$ if

$$\mathcal{E}(|\psi\rangle\langle\psi|) = |\psi\rangle\langle\psi|. \qquad (54)$$

The key property of localizable superoperators that we will exploit is given here below.

Theorem 5 If $\mathcal{E}$ is a localizable superoperator on $\mathcal{H}_A \otimes \mathcal{H}_B$, and $|\psi\rangle$, $A \otimes I|\psi\rangle$, and $I \otimes B|\psi\rangle$ are all eigenstates

	Bob	
	01	23
Alice 01	$|\phi^{\pm}\rangle, |\psi^{\pm}\rangle$	$|\phi^{\pm}\rangle, |\psi^{\pm}\rangle$
23	$|\phi^{\pm}\rangle, |\psi^{\pm}\rangle$	$U_B|\phi^{\pm}\rangle, U_B|\psi^{\pm}\rangle$

FIG. 7. A causal complete orthogonal measurement that is not localizable, in 4×4 dimensions. The Hilbert space is divided into four 2×2 quadrants, and the elements of the orthonormal basis are maximally entangled Bell states in each quadrant. Because the Bell basis in the bottom-right quadrant has been twisted by applying the unitary transformation $I\otimes U_B$, the corresponding measurement superoperator cannot be implemented without communication between Alice and Bob.

of $\mathcal{E}$ (where A and B are invertible operators), then $A\otimes B|\psi\rangle$ is also an eigenstate of $\mathcal{E}$.

The proof of Theorem 5 is in Appendix B. Clearly the claim is plausible. By hypothesis, the eigenstate $|\psi\rangle$ of $\mathcal{E}$ is mapped to a new eigenstate if Alice applies A and Bob does nothing, or if Bob applies B and Alice does nothing. Since $\mathcal{E}$ is localizable, Alice and Bob should be able to decide independently whether to apply A and B, and still obtain an eigenstate of $\mathcal{E}$.

A. A twisted partition

Having identified in Theorem 5 a necessary condition for localizability, we proceed to describe causal measurements for which this condition is violated. An example in 4×4 dimensions is illustrated in Fig. 7. Alice's four-dimensional space is partitioned into two two-dimensional subspaces, $\mathcal{H}_A^{01}$ spanned by $\{|0\rangle_A, |1\rangle_A\}$, and $\mathcal{H}_A^{23}$ spanned by $\{|2\rangle_A, |3\rangle_A\}$. Bob's space is similarly partitioned. Let us first consider the case where the measurement projects onto the standard Bell basis $\{|\phi^{\pm}\rangle, |\psi^{\pm}\rangle\}$ in each of the four subspaces $\mathcal{H}_A^{\alpha}\otimes\mathcal{H}_B^{\beta}$; that is, the orthonormal basis is

$$|\phi_{00}^{\pm}\rangle, |\psi_{00}^{\pm}\rangle, \quad |\phi_{02}^{\pm}\rangle, |\psi_{02}^{\pm}\rangle,$$

$$|\phi_{20}^{\pm}\rangle, |\psi_{20}^{\pm}\rangle, \quad |\phi_{22}^{\pm}\rangle, |\psi_{22}^{\pm}\rangle, \qquad (55)$$

where we use the notation

$$|\phi_{ij}^{\pm}\rangle = \frac{1}{\sqrt{2}}(|i,j\rangle \pm |i+1,j+1\rangle),$$

$$|\psi_{ij}^{\pm}\rangle = \frac{1}{\sqrt{2}}(|i,j+1\rangle \pm |i+1,j\rangle). \qquad (56)$$

The superoperator that induces decoherence in this basis *is* localizable, and in fact it can be implemented with shared randomness—no entanglement is required. Alice and Bob can each perform a partial measurement to identify whether the state occupies the subspace $\mathcal{H}^{01}$ or $\mathcal{H}^{23}$. Then they can proceed to implement decoherence in the 2×2 Bell basis as described in Sec. II C. Finally, Alice and Bob discard the

record of the partial measurement to complete the implementation of the measurement superoperator.

But now consider a "twisted" basis in which the basis elements in the $\mathcal{H}_A^{23}\otimes\mathcal{H}_B^{23}$ quadrant of the Hilbert space are rotated by applying a unitary transformation U_B to Bob's half of the state, becoming

$$I_A\otimes U_B|\phi_{22}^{\pm}\rangle, \quad I_A\otimes U_B|\psi_{22}^{\pm}\rangle, \qquad (57)$$

where U_B maps $\mathcal{H}_B^{23}$ to $\mathcal{H}_B^{23}$. Since this new basis still meets the criterion of Theorem 3 in both $\mathcal{H}_A$ and $\mathcal{H}_B$, the corresponding measurement superoperator $\mathcal{E}$ is still causal. But because $\mathcal{E}$ does not satisfy the criterion of Theorem 5 (except in the case were U_B merely permutes the Bell basis), it is no longer localizable. The eigenstates of $\mathcal{E}$ include, for example,

$$|\phi_{00}^{\pm}\rangle,$$

$$|\phi_{20}^{\pm}\rangle = X^2\otimes I|\phi_{00}^{\pm}\rangle,$$

$$|\phi_{02}^{\pm}\rangle = I\otimes X^2|\phi_{00}^{\pm}\rangle, \qquad (58)$$

where X^2 is the four-dimensional Pauli operator that acts on the basis $\{|0\rangle, |1\rangle, |2\rangle, |3\rangle\}$ according to

$$X^2: |i\rangle \rightarrow |i+2 \,(\mathrm{mod}\ 4)\rangle. \qquad (59)$$

If $\mathcal{E}$ is localizable, Theorem 5 requires that

$$|\phi_{22}^{\pm}\rangle = X^2\otimes X^2|\phi^{\pm}\rangle \qquad (60)$$

also be an eigenstate—i.e., that $|\phi_{22}\rangle$ is also an element of the orthonormal basis. This is not so unless U_B is one of the Pauli matrices, up to a phase. Therefore, $\mathcal{E}$ is not localizable.

The method that worked for the untwisted basis ($U_B = I$) fails for the twisted basis. Alice's partial measurement identifies what row the state occupies, and Bob's measurement identifies the column, but neither one has enough information to determine whether or not the state lies in the bottom-right quadrant where the basis is twisted. Without this information, they cannot complete the protocol successfully.

If Bob did have this information, then the protocol could be completed. Hence, not only is $\mathcal{E}$ semilocalizable (like any causal complete-measurement superoperator); furthermore it can actually be implemented with one-way *classical* communication. Alice performs the partial measurement that projects onto $\mathcal{H}_A^{01}$ or $\mathcal{H}_A^{23}$, and sends her measurement outcome to Bob. She also sends to Bob a copy of a table of random numbers that she has generated. Then Bob, after performing his partial measurement, has enough information to determine whether the state occupies the bottom-right quadrant, where the stabilizer generators are

$$X\otimes U_B X U_B^{-1}, \quad Z\otimes U_B Z U_B^{-1}, \qquad (61)$$

or one of the other three quadrants, where the stabilizer generators are

$$X\otimes X, \quad Z\otimes Z. \qquad (62)$$

(Here X and Z denote 2×2 Pauli operators.) When Alice and Bob consult their shared randomness and Alice is directed to apply $\sigma\in\{I,X,Y,Z\}$ to her half of the state, Bob applies $U_B\sigma U_B^{-1}$ in the former case, and σ in the latter case, to induce decoherence in the proper stabilizer eigenstate basis.

Note that, for this protocol to work, the classical communication must be in the proper direction—it must be Bob, not Alice, who chooses from two alternative operations. This statement sounds surprising at first, as we know that a unitary transformation applied by Bob to any maximally entangled state is equivalent to a suitable unitary transformation applied by Alice. However, unitary transformations applied by Alice and by Bob do *not* have equivalent effects when applied to *all* the elements of a maximally entangled *basis*. Correspondingly, it must be Bob, not Alice, who applies the rotation to transform the stabilizer generators of Eq. (62) to those of Eq. (61).

B. A twisted Bell basis

We saw that the general causal complete-measurement superoperator projects onto a basis that partitions $\mathcal{H}_A\otimes\mathcal{H}_B$ into $d\times d$ subspaces. Of course, if $d=1$, the basis is a product basis and the superoperator is trivially localizable. What about the other limiting case, in which $d=N_A=N_B$, so that the basis is maximally entangled? We will give an example of a 4×4 maximally entangled basis, such that the corresponding measurement superoperator is *not* localizable.

In the $d\times d$ case, any maximally entangled state can be expressed as $U\otimes I|\Phi^+\rangle$, where

$$|\Phi^+\rangle=\frac{1}{\sqrt{d}}\sum_{i=1}^{d}|i\rangle\otimes|i\rangle \qquad (63)$$

and U is unitary. The elements of a $d\times d$ maximally entangled basis, then, can be expressed as

$$|\Phi_a\rangle=U_a\otimes I|\Phi^+\rangle; \qquad (64)$$

the requirement that the states are orthogonal becomes

$$\text{tr}(U_a^\dagger U_b)=d\,\delta_{ab}. \qquad (65)$$

For the standard $d\times d$ Bell basis, these unitary transformations can be chosen as

$$U_{a,b}=X^aZ^b, \quad a,b=0,1,\ldots,d-1, \qquad (66)$$

where X and Z are the d-dimensional Pauli operators that act on a basis $\{|0\rangle,|1\rangle,\ldots,|d-1\rangle\}$ as

$$X|i\rangle=|i+1\pmod{d}\rangle,$$

$$Z|i\rangle=\omega^i|i\rangle, \quad \omega=e^{2\pi i/d}. \qquad (67)$$

Any measurement superoperator that projects onto a maximally entangled basis satisfies the criterion of Theorem 3 (in both directions) and is therefore causal. But if the superoperator is localizable, Theorem 5 requires the unitary transformations $\{U_a\}$ satisfying Eq. (65) to obey further restrictions. Note that if Alice and Bob both adopt the Schmidt basis of

$|\Phi_0\rangle$ as their computational bases, then $|\Phi_0\rangle=|\Phi^+\rangle$ and $U_0=I$. Then a useful characterization of the U_a's is provided by the following theorem.

Theorem 6 Let $\mathcal{U}$ be a set of d^2 $d\times d$ unitary matrices satisfying Eq. (65), and let $\mathcal{E}_\mathcal{U}$ be the measurement superoperator that projects onto the orthonormal basis $\{U_a\otimes I|\Phi^+\rangle, U_a\in\mathcal{U}\}$. Suppose that $\mathcal{E}_\mathcal{U}$ is localizable. Then if I, U, and V are all contained in $\mathcal{U}$, so must be $e^{i\phi}UV$, for some phase $e^{i\phi}$.

That is, $\mathcal{U}$ is a projective group.

Proof. Theorem 6 follows easily from Theorem 5. First note that

$$M\otimes I|\Phi^+\rangle=I\otimes M^T|\Phi^+\rangle, \qquad (68)$$

where M is any operator, and the transpose is taken in the computational basis. Then by hypothesis, all of

$$|\Phi^+\rangle, \quad U\otimes I|\Phi^+\rangle, \quad I\otimes V^T|\Phi^+\rangle, \qquad (69)$$

are eigenstates of $\mathcal{E}_\mathcal{U}$. Theorem 5 then implies that

$$U\otimes V^T|\Phi^+\rangle=UV\otimes I|\Phi^+\rangle \qquad (70)$$

is also an eigenstate, and hence an element of the orthonormal basis, up to a phase. This proves the theorem.

Now to exhibit a causal measurement superoperator that is not localizable, it suffices to construct unitary operators that do not satisfy the projective group property specified in Theorem 6. Consider, in the 4×4 case, the 16 unitary operators

$$\mathcal{U}=\begin{Bmatrix} I & Z & Z^2 & Z^3 \\ X & XZ & XZ^2 & XZ^3 \\ X^2 & X^2Z & X^2Z^2 & X^2Z^3 \\ X^3 & X^3\tilde{Z} & X^3Z^2 & X^3\tilde{Z}Z^2 \end{Bmatrix}, \qquad (71)$$

where $\tilde{Z}$ and Z^2 are the diagonal 4×4 matrices

$$\tilde{Z}=\text{diag}(1,1,-1,-1).$$

$$Z^2=\text{diag}(1,-1,1,-1). \qquad (72)$$

We can readily check that these operators obey the orthogonality condition Eq. (65), as I, $\tilde{Z}$, Z^2, and

$$\tilde{Z}Z^2=\text{diag}(1,-1,-1,1) \qquad (73)$$

are all mutually orthogonal. (They are the characters of the four unitary irreducible representations of the group $Z_2\times Z_2$.) However, due to the mismatch of the fourth row of Eq. (71) with the first three rows, $\mathcal{U}$ does not have the projective group property required by Theorem 6. For example, X and X^2Z are contained in $\mathcal{U}$, but their product X^3Z is not proportional to any element of $\mathcal{U}$. Therefore $\mathcal{E}_\mathcal{U}$ is not localizable.

As with any causal complete-measurement superoperator, $\mathcal{E}_\mathcal{U}$ is semilocalizable. But in contrast with the preceding example, classical communication is not sufficient—quantum communication (or equivalently, classical communication

and shared entanglement) is needed to implement the operation. To prove this we can appeal to a result of Ref. [16]: If a superoperator can be implemented with one-way classical communication, and has a maximally entangled state as an eigenstate, then it is localizable. Since the superoperator that projects onto the twisted Bell basis has a maximally entangled state as an eigenstate, and is *not* localizable, we know that it cannot be done with one-way classical communication. Further examples of twisted Bell bases are presented and discussed in Ref. [16].

VI. QUANTUM CORRELATIONS AND LOCALIZABILITY

In this section, we will use a different method to exhibit another class of causal superoperators that are not localizable. The construction exploits fundamental limitations on the strength of correlations among the parts of a quantum system, limitations embodied by the Cirel'son inequality. An operation that produces correlations that are too strong cannot be implemented without communication among the parts.

A related observation is that correlations arising from quantum entanglement are stronger than can be achieved with shared randomness—this is the content of Bell's theorem. We use this idea to construct examples of superoperators that can be locally implemented with prior quantum entanglement, but cannot be locally implemented with shared randomness.

A. The CHSH and Cirel'son inequalities

Suppose that Alice receives a classical input bit $x \in \{0,1\}$ and is to produce a classical output bit a, while Bob receives input bit y and is to produce output bit b. Their goal is to generate output bits that are related to the input bits according to

$$a \oplus b = x \wedge y, \tag{74}$$

where $\oplus$ denotes the sum modulo 2 (the XOR gate) and $\wedge$ denotes the product (the AND gate).

If Alice and Bob are unable to communicate with one another, so that Alice does not know Bob's input and Bob does not know Alice's, then they will not be able to achieve their goal for all possible values of the input bits. Let a_0, a_1 denote the value of Alice's output if her input is $x = 0,1$ and let b_0, b_1 denote Bob's output if his input is $y = 0,1$. They would like their output bits to satisfy

$$a_0 \oplus b_0 = 0,$$

$$a_0 \oplus b_1 = 0,$$

$$a_1 \oplus b_0 = 0,$$

$$a_1 \oplus b_1 = 1; \tag{75}$$

this is impossible, since by summing the four equations we obtain $0 = 1$.

If Alice and Bob always choose the output $a = b = 0$, then they will achieve their goal with probability 3/4, if all pos-

sible values of the input bits are equally probable. The Clauser-Horne-Shimony-Holt (CHSH) *inequality* says that, even if Alice and Bob share a table of random numbers, no higher success probability is attainable. To make the connection with the CHSH inequality as it is usually formulated [26], define random variables with values ± 1 as

$$A = (-1)^{a_0}, \quad A' = (-1)^{a_1},$$

$$B = (-1)^{b_0}, \quad B' = (-1)^{b_1}. \tag{76}$$

Then the CHSH inequality says that for any joint probability distribution governing $A, A', B, B' \in \{\pm 1\}$, the expectation values satisfy

$$\langle AB \rangle + \langle AB' \rangle + \langle A'B \rangle - \langle A'B' \rangle \leq 2. \tag{77}$$

Furthermore, if we denote by p_{xy} the probability that Eq. (75) is satisfied when the input bits are (x,y), then

$$\langle AB \rangle = 2p_{00} - 1,$$

$$\langle AB' \rangle = 2p_{01} - 1,$$

$$\langle A'B \rangle = 2p_{10} - 1,$$

$$\langle A'B' \rangle = 1 - 2p_{11}, \tag{78}$$

so that Eq. (77) becomes [27]

$$\frac{1}{4}(p_{00} + p_{01} + p_{10} + p_{11}) \leq \frac{3}{4}. \tag{79}$$

If Alice and Bob share quantum entanglement, they still cannot satisfy Eq. (75) for all inputs, but they can achieve an improved success probability compared to the case where they share only randomness. If we suppose that A, A', B, B' are all Hermitian operators with eigenvalues ± 1, and that Alice's operators A and A' commute with Bob's operators B and B', then the quantum-mechanical expectation values obey the Cirel'son inequality [26]

$$\langle AB \rangle + \langle AB' \rangle + \langle A'B \rangle - \langle A'B' \rangle \leq 2\sqrt{2}; \tag{80}$$

the success probability then satisfies

$$\frac{1}{4}(p_{00} + p_{01} + p_{10} + p_{11}) \leq \frac{1}{2} + \frac{1}{2\sqrt{2}} \approx 0.853. \tag{81}$$

Furthermore, the inequality can be saturated if the observables A, A', B, B' are chosen appropriately.

B. A causal operation that is not localizable

Our observations concerning the Cirel'son inequality quickly lead us to a construction of a causal operation that is not localizable.

For a two-qubit state shared by Alice and Bob, consider a superoperator, denoted $\mathcal{E}_{\wedge}$, that can be implemented in two steps. The first step is a complete orthogonal measurement that projects onto the product basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$.

Then in the second step the product state is transformed according to

$$
\begin{matrix}
|00\rangle \\
|01\rangle \} \\
|10\rangle
\end{matrix} \rightarrow \frac{1}{2}(|00\rangle\langle00| + |11\rangle\langle11|),
$$

$$
|11\rangle \rightarrow \frac{1}{2}(|01\rangle\langle01| + |10\rangle\langle10|). \tag{82}
$$

This operation is obviously trace preserving, and since it has an evident operator-sum representation, it is also completely positive. Furthermore, it is causal. Whatever the initial state that Alice and Bob share, each has the final density operator $\rho = I/2$; therefore, neither can receive a signal from the other.

Though causal, $\mathcal{E}_\wedge$ is not localizable—it cannot be implemented by Alice and Bob without communication, even if they share an entangled ancilla. If it were localizable, then Alice and Bob would be able to implement $\mathcal{E}_\wedge$ by applying a local unitary transformation $U_A \otimes U_B$ to the composite system consisting of the input qubits and ancilla, and then throwing some qubits away. Let the input state shared by Alice and Bob be one of the products states $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, let them apply their local unitary transformation to implement $\mathcal{E}_\wedge$, and suppose that each measures her or his output qubit in the basis $\{|0\rangle, |1\rangle\}$ after $\mathcal{E}_\wedge$ is performed. In effect then, Alice subjects the initial state to a measurement of the observable $U_A^{-1} Z_{A,\text{out}} U_A$, and Bob measures $U_B^{-1} Z_{B,\text{out}} U_B$, where $Z_{\cdot,\text{out}}$ denotes a Pauli operator acting on an output qubit. Both observables have eigenvalues ± 1.

Now, the Cirel'son inequality applies to a situation where Alice measures either of two observables in a specified state, and Bob does likewise. Here we are considering a case in which Alice and Bob measure fixed observables, and the initial state can be any of four possible states. But either scheme can be easily related to the other. For example, instead of providing Alice with an input qubit that can be $|0\rangle$ or $|1\rangle$, we can give her the input $|0\rangle$ and instruct her to apply X, or not, before she performs her measurement. In this scenario, Alice receives a classical input bit that instructs her to measure one of the two observables

$$
A = U_A^{-1} Z_{A,\text{out}} U_A,
$$

$$
A' = X_{A,\text{in}} U_A^{-1} Z_{A,\text{out}} U_A X_{A,\text{in}}, \tag{83}
$$

and similarly for Bob.

In this case, then, the Cirel'son inequality constrains how Alice's measurement outcome $|a\rangle$ is correlated with Bob's measurement outcome $|b\rangle$. But if they have really succeeded in implementing the operation $\mathcal{E}_\wedge$, then the outcomes are related to the classical input bits x, y by $a \oplus b = x \wedge y$ with probability 1, a violation of the bound Eq. (81). We conclude that no local protocol implementing $\mathcal{E}_\wedge$ is possible.

However, it is also clear that $\mathcal{E}_\wedge$ is *semilocalizable*—it can be implemented with one-way *classical* communication from Alice to Bob (or from Bob to Alice). Alice measures her qubit in the basis $\{|0\rangle, |1\rangle\}$, and she tosses a coin to decide whether to flip her qubit or not. Then she sends her measurement result and the outcome of her coin toss to Bob. Bob measures his qubit in the basis $\{|0\rangle, |1\rangle\}$, and after reading the data sent by Alice, either flips it or not. Bob arranges that his qubit have the same value as Alice's unless they both measure $|1\rangle$, in which case he arranges for his qubit and Alice's to have opposite values. This procedure implements $\mathcal{E}_\wedge$.

How much communication is necessary? As pointed out to us by Cleve [28], we can obtain a lower bound on the amount of communication needed to implement $\mathcal{E}_\wedge$ from known lower bounds on the quantum communication complexity of the inner product function [29].

Suppose that Alice has an n-bit classical input string $x = (x_1, x_2, \ldots, x_n)$, not known to Bob, and Bob has an n-bit classical input string $y = (y_1, y_2, \ldots, y_n)$, not known to Alice. Their goal is to compute the mod 2 inner product of their strings,

$$
\mathcal{I}(x,y) = x_1 y_1 \oplus x_2 y_2 \oplus \cdots \oplus x_n y_n. \tag{84}
$$

It is known [29] that even if Alice and Bob share preexisting entanglement, neither party can evaluate $\mathcal{I}(x,y)$ with zero probability of error unless at least $n/2$ qubits are transmitted between the parties. For n even, $n/2$ qubits of communication are also sufficient: Alice can use superdense coding to send x to Bob, and Bob can then evaluate $\mathcal{I}(x,y)$.

But if Alice and Bob were able to implement $\mathcal{E}_\wedge$ "for free," they could use it to evaluate $\mathcal{I}(x,y)$ at a smaller communication cost. Alice prepares the n-qubit state $|x\rangle$ and Bob the n-qubit state $|y\rangle$. Then $\mathcal{E}_\wedge$ is applied to $|x_i, y_i\rangle$ for each $i = 1, 2, \ldots, n$, and Alice and Bob measure their qubits to obtain outputs a_i, b_i for each i. Since $a_i \oplus b_i = x_i y_i$, we see that

$$
\mathcal{I}(x,y) = (a_1 \oplus b_1) \oplus (a_2 \oplus b_2) \oplus \cdots \oplus (a_n \oplus b_n)
$$
$$
= (a_1 \oplus a_2 \oplus \cdots \oplus a_n) \oplus (b_1 \oplus b_2 \oplus \cdots \oplus b_n). \tag{85}
$$

Therefore, Alice can evaluate the sum (mod 2) of her n measurement outcomes, and send the *one-bit* result to Bob. Bob adds Alice's result to the sum of his own measurement outcomes, and so obtains the value of $\mathcal{I}(x,y)$. Just one bit of communication is required.

Suppose that Alice and Bob have a protocol that allows them to implement $\mathcal{E}_\wedge$ with, on average, Q_{av} qubits of quantum communication. (Alice's decision whether to send a qubit could be conditioned on the outcome of a local measurement; therefore the amount of communication required can fluctuate about this average.) Now, if Alice and Bob can implement $\mathcal{E}_\wedge$ n times with Q_n qubits of communication, then since just one additional bit is needed to complete the evaluation of the inner product function, we know that

$$
Q_n + 1 \geq n/2. \tag{86}
$$

For large n, Q_n converges to nQ_{av}, and we conclude that

$$Q_{\text{av}} \geq 1/2. \tag{87}$$

This argument illustrates a general approach to proving that a quantum operation is unlocalizable: if implementing the operation would allow us to reduce the communication complexity of a function below established lower bounds, then no local implementation is possible.

C. Entanglement is stronger than shared randomness

The separation between the CHSH and Cirel'son inequalities allows us to construct a class of operations that can be implemented locally with shared entanglement, but cannot be implemented locally with shared randomness. (The existence of operations with this property was already pointed out in Sec. II C.)

Suppose that Alice and Bob have a shared maximally entangled pair of qubits (qubits 3 and 4), as well as two input qubits (1 and 2) on which the operation is to act. Alice measures qubit 1 and Bob measures qubit 2 in the basis $\{|0\rangle, |1\rangle\}$. Then Alice measures her half of the entangled pair, qubit 3, choosing to measure the observable A if the measured input was $|0\rangle_1$ or A' if the measured input was $|1\rangle_1$. Similarly, Bob measures qubit 4, choosing to measure either B or B' depending on the outcome of his measurement of the input qubit 2.

After measuring A (or A'), Alice rotates qubit 3 to the state $|0\rangle_3$ if she found $A=1$ (or $A'=1$), and rotates it to $|1\rangle_3$ if she found $A=-1$ (or $A'=-1$). Bob does the same to qubit 4. Finally Alice and Bob throw away the input qubits 1 and 2, retaining qubits 3 and 4.

Alice and Bob, then, using their shared entanglement, have locally implemented an operation that acts on a product-state input and produces a product-state output, according to

$$|x,y\rangle \rightarrow |a,b\rangle. \tag{88}$$

Averaged over the four possible product-state inputs, the output of the operation satisfies $a \oplus b = xy$ with a success probability that we will call p. If the observables A, A', B, B' are chosen to saturate the Cirel'son inequality, then $p = \cos^2(\pi/8) \approx .853$.

As is well known [26], probability distributions for quantum measurements of a single qubit can be correctly accounted for by a "hidden-variable theory" (while measurements of entangled qubits cannot be). Therefore, measurements performed by Alice and Bob on a product-input state can be perfectly simulated by a classical probability distribution, so that the measurement results must respect the CHSH inequality, which requires that the success probability p satisfy $p \leq 3/4$. For $3/4 < p \leq \cos^2(\pi/8)$, the operation can be implemented locally with shared entanglement, but not with shared randomness.

VII. UNITARITY AND CAUSALITY

An important special case of an operation is a unitary transformation. In this case, our classification collapses—the classes of causal, localizable, semicausal, and semilocaliz-able unitary transformations all coincide, according to the following theorem.

Theorem 7. A bipartite unitary transformation U_{AB} is semicausal if and only if it is a tensor product, $U_{AB} = U_A \otimes U_B$.

Proof. It is obvious that a tensor product transformation is causal. The nontrivial content of the theorem is that if U_{AB} is not a tensor product, then we can devise protocols whereby Alice can signal Bob and Bob can signal Alice; hence U_{AB} is not semicausal.

To prove this, we first recall that since linear operators are a vector space with a Hilbert-Schmidt inner product, a bipartite operator (whether unitary or not) can be Schmidt decomposed [30]. We may write

$$U_{AB} = \sum_{\mu} \lambda_{\mu} A_{\mu} \otimes B_{\mu}, \tag{89}$$

where the λ_{μ}'s are non-negative real numbers, and the operator bases $\{A_{\mu}\}$ and $\{B_{\mu}\}$ are orthogonal,

$$\text{tr}(A_{\mu}^{\dagger} A_{\nu}) = N_A \delta_{\mu\nu}, \quad \text{tr}(B_{\mu}^{\dagger} B_{\nu}) = N_B \delta_{\mu\nu}; \tag{90}$$

N_A is the dimension of $\mathcal{H}_A$ and N_B is the dimension of $\mathcal{H}_B$—we have chosen this normalization so that unitary operators are properly normalized.

If U_{AB} is not a tensor product, than more than one λ_{μ} is strictly positive. We will show that if this is true, then U_{AB} allows Bob to signal Alice. (A similar argument shows that Alice can signal Bob.) Suppose that Alice introduces a *reference system* $\mathcal{H}_R$ and that she prepares a maximally entangled state of $\mathcal{H}_R \otimes \mathcal{H}_A$

$$|\Phi\rangle_{RA} = \sum_i |i\rangle_R \otimes |i\rangle_A. \tag{91}$$

(Because it will be convenient later on, we have chosen an unconventional normalization of the state $|\Phi\rangle_{RA}$.) Meanwhile, Bob prepares a pure state $|\psi\rangle_B$. When U_{AB} acts, the density operator becomes

$$\rho_{RAB} = \sum_{\mu\nu} \lambda_{\mu} \lambda_{\nu} (I \otimes A_{\mu}) |\Phi\rangle \langle \Phi| (I \otimes A_{\nu}^{\dagger}) \otimes B_{\mu} |\psi\rangle \langle \psi| B_{\nu}^{\dagger}. \tag{92}$$

After tracing out Bob's system, the density operator of Alice's system becomes

$$\rho_{RA} = \sum_{\mu\nu} \lambda_{\mu} \lambda_{\nu} (I \otimes A_{\mu}) |\Phi\rangle \langle \Phi| (I \otimes A_{\nu}^{\dagger}) \langle \psi| B_{\nu}^{\dagger} B_{\mu} |\psi\rangle. \tag{93}$$

Bob can signal Alice if the density operator ρ_{RA} depends on Bob's initial state $|\psi\rangle$. It follows from Eq. (90) that the states $\{(I \otimes A_{\mu}) |\Phi\rangle\}$ are mutually orthogonal; therefore signaling is possible if there exist two states $|\psi\rangle$ and $|\psi'\rangle$ such that

$$\langle \psi| B_{\nu}^{\dagger} B_{\mu} |\psi\rangle \neq \langle \psi'| B_{\nu}^{\dagger} B_{\mu} |\psi'\rangle. \tag{94}$$

for some μ and ν.

Now we distinguish two cases.

(1) Suppose that for some μ, B_μ is not unitary. Then, in order to satisfy the normalization condition Eq. (90), $B_\mu^\dagger B_\mu$ must have (at least) two distinct eigenvalues. Choose $|\psi\rangle$ and $|\psi'\rangle$ to be the corresponding eigenvectors. Then Eq. (94) is satisfied for $\nu=\mu$.

(2) Suppose that B_μ and B_ν are both unitary for $\mu\neq\nu$. Then $B_\nu^\dagger B_\mu$ is nonzero and [according to Eq. (90)] has vanishing trace; therefore it has (at least) two distinct eigenvalues. Thus Eq. (94) is satisfied, where $|\psi\rangle$ and $|\psi'\rangle$ are the corresponding eigenvectors.

In either case, Alice's density operator depends on how Bob's initial state is chosen; hence Bob can signal Alice. A similar argument shows that Alice can signal Bob. Therefore U_{AB} is not semicausal. This completes the proof of Theorem 7.

It follows immediately from Theorem 7 that if a bipartite unitary transformation is semicausal, it is also localizable, and therefore fully causal.

VIII. GENERAL OPERATIONS: CRITERIA FOR SEMICAUSALITY

To show that an operation $\mathcal{E}$ is *not* fully causal, it suffices to exhibit a protocol whereby the operation can be used to send a signal in one direction, and to show that it is not semicausal, it suffices to exhibit protocols for signaling in both directions. On the other hand, to show that it *is* fully causal (or semicausal), we must prove that no such signaling protocols exist. To settle whether a particular operation is causal, it is very helpful to have a simpler criterion that can be checked with a straightforward calculation. We will now develop such a criterion for semicausal superoperators.

First, we recall that, although in our definition of semicausality we allowed the initial state shared by Alice and Bob to be entangled, we could without loss of generality restrict their initial state to be a product state (Theorem 1). Next, we note that a helpful tool in our analysis of the causality properties of unitary transformations, entanglement with a reference system, can also be fruitfully applied to the general case. To give a useful restatement of the criterion for semicausality, suppose again that Alice prepares the maximally entangled state $|\Phi\rangle_{RA}$ of her reference system $\mathcal{H}_R$ with her system $\mathcal{H}_A$. Then for each state $|\varphi\rangle_A \in \mathcal{H}_A$, there is a corresponding "relative state" $|\varphi^*\rangle_R \in \mathcal{H}_R$, chosen so that

$$_R\langle\varphi^*|\Phi\rangle_{RA}=|\varphi\rangle_A. \tag{95}$$

We can easily see that an operation $\mathcal{E}$ is semicausal (Bob is unable to signal Alice) if and only if

$$\mathrm{tr}_B[\mathcal{E}(|\Phi\rangle\langle\Phi|\otimes|\psi\rangle\langle\psi|)] \tag{96}$$

is independent of Bob's state $|\psi\rangle$. (Here of course $\mathcal{E}$ really denotes the operation $I_R\otimes\mathcal{E}_{AB}$ acting on the RAB system.) If the expression in Eq. (96) depends on $|\psi\rangle$, then obviously Bob can signal Alice. Conversely, if Bob can signal Alice,

then there is a signaling protocol in which the initial state is a product of pure states; there are states $|\varphi\rangle$, $|\psi\rangle$, and $|\psi'\rangle$, such that

$$\begin{aligned}
&_R\langle\varphi^*|\mathrm{tr}_B[\mathcal{E}(|\Phi\rangle\langle\Phi|\otimes|\psi\rangle\langle\psi|)]|\varphi^*\rangle_R\\
&=\mathrm{tr}_B[\mathcal{E}(|\varphi\rangle\langle\varphi|\otimes|\psi\rangle\langle\psi|)]\\
&\neq\mathrm{tr}_B[\mathcal{E}(|\varphi\rangle\langle\varphi|\otimes|\psi'\rangle\langle\psi'|)]\\
&=_R\langle\varphi^*|\mathrm{tr}_B[\mathcal{E}(|\Phi\rangle\langle\Phi|\otimes|\psi'\rangle\langle\psi'|)]|\varphi^*\rangle_R.
\end{aligned} \tag{97}$$

Since we have found a particular matrix element of $\mathrm{tr}_B[\mathcal{E}(|\Phi\rangle\langle\Phi|\otimes|\psi\rangle\langle\psi|)]$ that depends on $|\psi\rangle$, evidently so does $\mathrm{tr}_B[\mathcal{E}(|\Phi\rangle\langle\Phi|\otimes|\psi\rangle\langle\psi|)]$ itself.

Now, let us provide Bob with a reference system S, and suppose that he prepares a maximally entangled state of $\mathcal{H}_B\otimes\mathcal{H}_S$

$$|\Phi'\rangle_{BS}=\sum_i |i\rangle_B\otimes|i\rangle_S \tag{98}$$

(again we have chosen an unconventional normalization for convenience). We are ready to state and prove our new criterion for semicausality.

Theorem 8. Let $|\Phi\rangle_{RA}$ be a maximally entangled state of system A with the reference system R, and let $|\Phi'\rangle_{BS}$ be a maximally entangled state of system B with reference system S. Then the bipartite superoperator $\mathcal{E}$ acting on AB is semicausal (Bob cannot signal Alice) if and only if

$$\mathrm{tr}_B\{(\mathcal{E}_{AB}\otimes I_{RS})[(|\Phi\rangle\langle\Phi|)_{RA}\otimes(|\Phi'\rangle\langle\Phi'|)_{BS}]\} \tag{99}$$

is proportional to the product $\rho_{RA}\otimes I_S$, where I_S denotes the identity on S.

Proof. If $\mathrm{tr}_B[\mathcal{E}(|\Phi\rangle\langle\Phi|\otimes|\Phi'\rangle\langle\Phi'|)]$ is proportional to $\rho_{RA}\otimes I_S$, then by evaluating the matrix element between relative states $|\psi^*\rangle_S$, we see that $\mathrm{tr}_B[\mathcal{E}(|\Phi\rangle\langle\Phi|\otimes|\psi\rangle\langle\psi|)]$ is independent of $|\psi\rangle$. Therefore Bob cannot signal Alice. Conversely, suppose that Bob cannot signal Alice. Then

$$\begin{aligned}
&\mathrm{tr}_B[\mathcal{E}(|\Phi\rangle\langle\Phi|\otimes|\psi\rangle\langle\psi|)]\\
&=_S\langle\psi^*|\mathrm{tr}_B[\mathcal{E}(|\Phi\rangle\langle\Phi|\otimes|\Phi'\rangle\langle\Phi'|)]|\psi^*\rangle_S
\end{aligned} \tag{100}$$

is independent of $|\psi^*\rangle$. It follows that

$$\mathrm{tr}_B[\mathcal{E}(|\Phi\rangle\langle\Phi|\otimes|\Phi'\rangle\langle\Phi'|)]$$

is proportional to I_S. This proves Theorem 8.

IX. CONCLUSIONS

We have studied the constraints on quantum operations that are imposed by relativistic causality. In the bipartite setting where no classical communication is permitted, we find a hierarchy of operations: (1) operations that can be implemented with no shared resources; (2) operations that can be implemented with shared randomness; (3) operations that can be implemented with shared entanglement (localizable

operations); (4) causal operations; and (5) acausal operations.

Our central observation is that the classes (3) and (4) do not coincide: there are operations that respect causality, but are nonetheless forbidden by the rules of local quantum physics.

Our work can be regarded as a useful step toward the broader goal of characterizing the physically realizable operations in relativistic quantum-field theory. However, as noted in Sec. I, to apply our results to field theory one must accept the idealization that the resources shared by the parties are external probes not themselves described by the field theory.

In a separate paper, we have also discussed causality constraints that apply to non-Abelian gauge theories [17]; we have shown that the *nondemolition* measurement of a spacelike Wilson-loop operator is an acausal operation (confirming a speculation of Sorkin [14]), and is therefore surely not localizable. On the other hand, a *destructive* measurement of a Wilson loop is possible—spacelike separated parties can perform a POVM from which the value of the Wilson loop can be inferred, but this POVM will damage Wilson-loop eigenstates.

The compatibility of quantum mechanics with special relativity is highly nontrivial; in fact, it is something of a miracle. Because relativistic quantum-field theories are so highly constrained, it is tempting to speculate that "quantum mechanics is the way it is because any small changes in quantum mechanics would lead to absurdities [31]."

From this perspective, the existence of causal operations that are not localizable comes as a surprise. We seem to have the freedom to relax the rules of quantum theory by allowing more general operations, without encountering unacceptable physical consequences. Nontrivial support for this notion is provided by the semigroup property of the causal operations. It is reasonable to insist that the operations allowed at a given time ought not to depend on the previous history of the system; since the composition of two causal operations is causal, a theory that admits more general causal operations than those allowed in local quantum theory could adhere to this proviso.

One wonders whether there are further principles, beyond relativistic causality, that will restrict the class of allowed operations to those and only those that are truly realizable in Nature. If so, these principles might lead us to an understanding of why quantum mechanics has to be the way it is. What might these principles be?

We do not know. But the discussion in Sec. VI invites us to contemplate the fundamental limitations on the correlations among the parts of a physical system. Experimentally confirmed violations of the CHSH inequality demonstrate that the correlations are stronger than those allowed by any local hidden variable theory. Operations that are causal but not localizable produce correlations that are stronger still, and violate the Cirel'son inequality. What criteria point toward a description of Nature that incorporates violation of the CHSH inequality, but not violation of the Cirel'son inequality?

Or could it be that Nature really does allow more general operations, and that the conventional framework of local

quantum physics needs revision? Ultimately, only experiment can decide.

Note added. After this paper appeared, a proof of the conjecture that semicausal superoperators are semilocalizable was found by Eggeling, Schlingemann, and Werner [32].

APPENDIX A: PROOF OF THEOREM 3

Theorem 3. Consider a bipartite complete orthogonal measurement superoperator of the form

$$\mathcal{E}(\rho) = \sum_a |a\rangle\langle a|\rho|a\rangle\langle a|, \qquad (A1)$$

where $\{|a\rangle\}$ is an orthonormal basis for $\mathcal{H}_A \otimes \mathcal{H}_B$, and let $\sigma_A^a = \mathrm{tr}_B(|a\rangle\langle a|)$. Then $\mathcal{E}$ is semicausal if and only if the following property holds: For each pair of operators $\{\sigma_A^a, \sigma_A^b\}$, either $\sigma_A^a = \sigma_A^b$ or $\sigma_A^a \sigma_A^b = 0$.

Proof. We begin by observing that if $\sigma_A^a \sigma_A^b = 0$, then

$$\langle a|I_A \otimes U_B|b\rangle = 0 \qquad (A2)$$

for any unitary U_B. To show this, we Schmidt decompose the states $|a\rangle$ and $|b\rangle$:

$$|a\rangle = \sum_i \sqrt{\lambda_{a,i}}|a,i\rangle_A \otimes |a,i\rangle_B,$$

$$|b\rangle = \sum_i \sqrt{\lambda_{b,i}}|b,i\rangle_A \otimes |b,i\rangle_B, \qquad (A3)$$

where $\{|a,i\rangle_A\}$, $\{|b,i\rangle_A\}$ are orthonormal bases of $\mathcal{H}_A$, $\{|a,i\rangle_B\}$, $\{|b,i\rangle_B\}$ are orthonormal bases of $\mathcal{H}_B$, and the $\lambda_{a,i}$'s, $\lambda_{b,i}$'s are all non-negative. In terms of these bases, we find

$$\sigma_A^a = \sum_i \lambda_{a,i}|a,i\rangle_{A\ A}\langle a,i|,$$

$$\sigma_A^b = \sum_i \lambda_{b,i}|b,i\rangle_{A\ A}\langle b,i|; \qquad (A4)$$

therefore $\sigma_A^a \sigma_A^b = 0$ iff $_A\langle a,i|b,j\rangle_A = 0$ for each i and j with $\lambda_{a,i}\lambda_{b,j} \neq 0$. Equation (A2) follows immediately.

Now suppose that for each a and b, either $\sigma_A^a = \sigma_A^b$ or $\sigma_A^a \sigma_A^b = 0$. Let $|\psi\rangle$ be an arbitrary pure state in $\mathcal{H}_A \otimes \mathcal{H}_B$. We will show that

$$\mathrm{tr}_B\mathcal{E}(|\psi\rangle\langle\psi|) = \mathrm{tr}_B\mathcal{E}((I_A \otimes U_B)|\psi\rangle\langle\psi|(I_A \otimes U_B^\dagger)) \quad (A5)$$

for any unitary U_B, which according to Theorem 1 suffices to show that $\mathcal{E}$ is semicausal. To prove Eq. (A5), we expand the state $|\psi\rangle$ in the basis $\{|b\rangle\}$ as

$$|\psi\rangle = \sum_b \alpha_b |b\rangle, \qquad (A6)$$

and so obtain

$$\mathrm{tr}_B \mathcal{E}((I_A \otimes U_B)|\psi\rangle\langle\psi|(I_A \otimes U_B^\dagger))$$

$$= \sum_{a,b,c} \alpha_b \alpha_c^* \langle a|I_A \otimes U_B|b\rangle\langle c|I_A \otimes U_B^\dagger|a\rangle \sigma_A^a. \qquad (A7)$$

Now we use the property that either $\sigma_A^a = \sigma_A^b$ or $\sigma_A^a \sigma_A^b = 0$ for each a and b: because $\sigma_A^a \sigma_A^b = 0$ implies that $\langle a|I_A \otimes U_B|b\rangle = 0$, we can replace σ_A^a by σ_A^b in Eq. (A7) without altering the sum. After this replacement, we use the property $\sum_a |a\rangle\langle a| = I$ and the unitarity of U_B to find

$$\mathrm{tr}_B \mathcal{E}((I_A \otimes U_B)|\psi\rangle\langle\psi|(I_A \otimes U_B^\dagger)) = \sum_b |\alpha_b|^2 \sigma_A^b, \qquad (A8)$$

which is independent of U_B, proving Eq. (A5) and hence the "if" part of Theorem 3.

To prove the "only if" part of Theorem 3, we suppose that for some a and b, $\sigma_A^a \sigma_A^b \neq 0$, and $\sigma_A^a \neq \sigma_A^b$; we must show that Bob can signal Alice. It suffices to show that a basis element $|b\rangle$ and a unitary transformation U_B can be chosen so that

$$\mathrm{tr}_B \mathcal{E}((I_A \otimes U_B)|b\rangle\langle b|(I_A \otimes U_B^\dagger)) = \sum_a |\langle a|I_A \otimes U_B|b\rangle|^2 \sigma_A^a$$

$$= \sigma_A^b = \mathrm{tr}_B \mathcal{E}(|b\rangle\langle b|). \qquad (A9)$$

If Eq. (A9) holds, then Bob can signal Alice by the following protocol: Alice and Bob prepare in advance the shared state $|b\rangle$. Just before $\mathcal{E}$ acts, Bob either applies U_B to the state or he does nothing. Equation (A9) says that Alice's density operator after $\mathcal{E}$ acts depends on the action chosen by Bob; therefore, Bob can signal Alice.

We will prove in two steps that U_B and $|b\rangle$ exist such that Eq. (A9) is satisfied. The first step is to show that for $\sigma_A^a \sigma_A^b \neq 0$, there is a unitary U_B such that

$$\langle b|I_A \otimes U_B|a\rangle \neq 0. \qquad (A10)$$

In terms of the Schmidt bases defined in Eq. (A3), what is to be shown can be rewritten as

$$\sum_{i,j} \sqrt{\lambda_{b,j}\lambda_{a,i}} \, {}_A\langle b,j|a,i\rangle_A \, {}_B\langle b,j|U_B|a,i\rangle_B \neq 0. \qquad (A11)$$

Now recall that $\sigma_A^a \sigma_A^b \neq 0$ implies that ${}_A\langle b,j|a,i\rangle_A \neq 0$ for some i and j. By labeling the Schmidt bases appropriately we can ensure that ${}_A\langle b,1|a,1\rangle_A \neq 0$. By adopting suitable phase conventions, we can ensure that each ${}_A\langle b,i|a,i\rangle_A$ is real and non-negative, and we can choose U_B so that ${}_B\langle b,j|U_B|a,i\rangle_B = \delta_{ij}$. Thus

$$\sum_{i,j} \sqrt{\lambda_{b,j}\lambda_{a,i}} \, {}_A\langle b,j|a,i\rangle_{A} \, {}_B\langle b,j|U_B|a,i\rangle_B$$

$$= \sum_i \sqrt{\lambda_{b,i}\lambda_{a,i}} \, {}_A\langle b,i|a,i\rangle_A \qquad (A12)$$

is a sum of non-negative terms, at least one of which is nonzero; therefore the sum is surely nonzero, as we wished to show.

Now we have seen that U_B and $|b\rangle$ can be chosen so that the sum in Eq. (A9) contains a term other than σ_A^b. The second step of the argument will establish that we can, in fact, choose U_B and $|b\rangle$ such that the sum is not equal to σ_A^b.

For this purpose, consider the set S^b containing all σ_A^a such that $\sigma_A^a \sigma_A^b \neq 0$. Suppose that S^b contains at least two elements, and that σ_A^b is an *extremal* element of S^b—that is, σ_A^b cannot be expressed as a nontrivial convex combination of other elements of S^b. Then since the sum in Eq. (A9) is a convex combination of elements of S^b, and since we can choose U_B so that the sum contains some $\sigma_A^a \neq \sigma_A^b$ with a nonvanishing coefficient, the inequality in Eq. (A9) follows from the extremality of σ_A^b in S^b.

Finally, it only remains to show that $|b\rangle$ can be chosen so that σ_A^b is extremal in S^b. For this purpose, of all σ_A^b such that S^b contains two or more elements, choose one with maximal Hilbert-Schmidt norm {i.e., with maximal $\mathrm{tr}[(\sigma_A^b)^2]$}. We claim that this σ_A^b must be extremal in S^b.

To see that σ_A^b is extremal in S^b, we appeal to the following property: Let $\{v_i\}$ be a finite set of vectors, and let $\|v\|_{\max}$ be the maximum value of $\|v_i\|$. Then the *strict* inequality

$$\left\| \sum_i p_i v_i \right\| < \|v\|_{\max} \qquad (A13)$$

holds for any nontrivial convex combination of the v_i's (one with two or more nonvanishing p_i's). Applying Eq. (A13) to S^b, the Hilbert-Schmidt norm of our selected σ_A^b is on the right-hand side, which is strictly greater than the left-hand side, the norm of any nontrivial convex combination of elements of S^b. Therefore σ_A^b is extremal in S^b.

This completes the proof of Theorem 3.

APPENDIX B: PROOF OF THEOREM 5

Theorem 5. If $\mathcal{E}$ is a localizable superoperator on $\mathcal{H}_A \otimes \mathcal{H}_B$, and $|\psi\rangle$, $A \otimes I|\psi\rangle$, and $I \otimes B|\psi\rangle$ are all eigenstates of $\mathcal{E}$ (where A and B are invertible operators), then $A \otimes B|\psi\rangle$ is also an eigenstate of $\mathcal{E}$.

Proof. If the superoperator $\mathcal{E}$ is localizable, its action on a pure state $|\eta\rangle_{AB}$ can be realized by a tensor product unitary transformation $U_{RA} \otimes V_{BS}$ acting on $|\eta\rangle_{AB}|\varphi\rangle_{RS}$, where $|\varphi\rangle_{RS}$ is a suitable ancilla state shared by Alice and Bob. By hypothesis, this unitary transformation acting on $|\psi\rangle_{AB}|\varphi\rangle_{RS}$ preserves $|\psi\rangle_{AB}$ and rotates only the ancilla state:

$$U_{RA} \otimes V_{BS}|\psi\rangle_{AB}|\varphi\rangle_{RS} = |\psi\rangle_{AB}|\varphi_0\rangle_{RS}, \qquad (B1)$$

for some state of the ancilla $|\varphi_0\rangle_{RS}$. Similarly, by hypothesis, we have

$$U_{RA}\otimes V_{BS}(A_A\otimes I_B)|\psi\rangle_{AB}|\varphi\rangle_{RS}=(A_A\otimes I_B)|\psi\rangle_{AB}|\varphi_A\rangle_{RS},$$

$$U_{RA}\otimes V_{BS}(I_A\otimes B_B)|\psi\rangle_{AB}|\varphi\rangle_{RS}=(I_A\otimes B_B)|\psi\rangle_{AB}|\varphi_B\rangle_{RS},$$

$$(B2)$$

for states of the ancilla $|\varphi_A\rangle_{RS}$, $|\varphi_B\rangle_{RS}$.

Now consider the transformation $\Delta_{RA}=UAU^{-1}A^{-1}$. By construction, Δ_{RA} acts only on Alice's system RA. In fact, we can show that when acting on the state $(A_A\otimes I_B)|\psi\rangle_{AB}|\varphi_0\rangle_{RS}$, Δ_{RA} acts trivially on A and nontrivially only on Alice's ancilla R; we observe that

$$(A_A\otimes I_B)|\psi\rangle_{AB}|\varphi_A\rangle_{RS}=(U_{RA}\otimes V_{BS})(A_A\otimes I_B)|\psi\rangle_{AB}|\varphi\rangle_{RS}$$

$$=(\Delta_{RA}\otimes I_B)(A_A\otimes I_B)(U_{RA}\otimes V_{BS})$$

$$\times|\psi\rangle_{AB}|\varphi\rangle_{RS}$$

$$=(\Delta_{RA}\otimes I_B)(A_A\otimes I_B)|\psi\rangle_{AB}|\varphi_0\rangle_{RS}.$$

$$(B3)$$

Therefore, acting on $(A_A\otimes I_B)|\psi\rangle_{AB}|\varphi_0\rangle_{RS}$, we may replace $\Delta_{RA}\otimes I_B$ by $R_R\otimes I_B$, where R_R is a (unitary) transformation acting on R alone that rotates $|\varphi_0\rangle_{RS}$ to $|\varphi_A\rangle_{RS}$. We then have

$$(U_{RA}A_A\otimes V_{BS})|\psi\rangle_{AB}|\varphi\rangle_{RS}$$

$$=(\Delta_{RA}\otimes I_B)(A_AU_{RA}\otimes V_{BS})|\psi\rangle_{AB}|\varphi\rangle_{RS}$$

$$=(\Delta_{RA}\otimes I_B)(A_A\otimes I_B)|\psi\rangle_{AB}|\varphi_0\rangle_{RS}$$

$$=(R_R\otimes I_B)(A_A\otimes I_B)|\psi\rangle_{AB}|\varphi_0\rangle_{RS}$$

$$=(R_R\otimes I_B)(A_A U_{RA}\otimes V_{BS})|\psi\rangle_{AB}|\varphi\rangle_{RS}, \quad (B4)$$

and multiplying both sides by $I_A\otimes V_{BS}^{-1}$ gives

$$(U_{RA}A_A\otimes I_B)|\psi\rangle_{AB}|\varphi\rangle_{RS}$$

$$=(R_R\otimes I_B)(A_A U_{RA}\otimes I_B)|\psi\rangle_{AB}|\varphi\rangle_{RS}; \quad (B5)$$

that is, acting on the state $|\psi\rangle_{AB}|\varphi\rangle_{RS}$, we may replace $U_{RA}A_A$ by $R_R A_A U_{RA}$. A similar argument shows that

$$(I_A\otimes V_{BS}B_S)|\psi\rangle_{AB}|\varphi\rangle_{RS}$$

$$=(I_A\otimes S_S)(I_A\otimes B_S V_{BS})|\psi\rangle_{AB}|\varphi\rangle_{RS}, \quad (B6)$$

where S_S is a unitary transformation acting on Bob's ancilla.

Now we can use the commutation properties Eqs. (B5) and (B6) to determine how the superoperator $\mathcal{E}$ acts on $(A\otimes B)|\psi\rangle_{AB}$:

$$(U_{RA}\otimes V_{BS})(A_A\otimes B_B)|\psi\rangle_{AB}|\varphi\rangle_{RS}$$

$$=(I_A\otimes V_{BS}B_S)(U_{RA}A_A\otimes I_B)|\psi\rangle_{AB}|\varphi\rangle_{RS}$$

$$=(R_R\otimes I_B)(I_A\otimes V_{BS}B_S)(A_A U_{RA}\otimes I_B)|\psi\rangle_{AB}|\varphi\rangle_{RS}$$

$$=(R_R\otimes I_B)(A_A U_{RA}\otimes I_B)(I_A\otimes V_{BS}B_S)|\psi\rangle_{AB}|\varphi\rangle_{RS}$$

$$=(R_R\otimes S_S)(A_A U_{RA}\otimes I_B)(I_A\otimes B_S V_{BS})|\psi\rangle_{AB}|\varphi\rangle_{RS}$$

$$=(R_R\otimes S_S)(A_A\otimes B_B)(U_{RA}\otimes V_{BS})|\psi\rangle_{AB}|\varphi\rangle_{RS}$$

$$=(R_R\otimes S_S)(A_A\otimes B_B)|\psi\rangle_{AB}|\varphi_0\rangle_{RS}$$

$$=[(A_A\otimes B_B)|\psi\rangle_{AB}][(R_R\otimes S_S)|\varphi_0\rangle_{RS}]. \quad (B7)$$

We have shown that $(A_A\otimes B_B)|\psi\rangle_{AB}$ is an eigenstate of $\mathcal{E}$, which completes the proof of Theorem 5.

---

[1] M. A. Nielsen and I. L Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[2] J. Preskill, http://www.theory.caltech.edu/people/preskill/ph229.

[3] R. Haag, *Local Quantum Physics: Fields, Particles, Algebras* (Springer-Verlag, Berlin, 1992).

[4] D. DiVincenzo (private communication).

[5] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Oxford University Press, Oxford, 1958).

[6] E. P. Wigner, Z. Phys. **133**, 101 (1952).

[7] H. Araki and M. M. Yanase, Phys. Rev. **120**, 622 (1960).

[8] M. A. Nielsen, Phys. Rev. Lett. **79**, 2915 (1997); e-print quant-ph/9706006.

[9] L. Landau and R. Peierls, Z. Phys. **69**, 56 (1931).

[10] N. Bohr and L. Rosenfeld, Mat. Fys. Medd. K. Dan. Vidensk. Selsk. **12**, 8 (1933).

[11] Y. Aharonov, D. Z. Albert, and L. Vaidman, Phys. Rev. D **34**, 1805 (1986).

[12] Y. Aharonov and D. Z. Albert, Phys. Rev. D **21**, 3316 (1980).

[13] Y. Aharonov and D. Z. Albert, Phys. Rev. D **24**, 359 (1981).

[14] R. D. Sorkin, in *Directions in General Relativity*, edited by B. L. Hu and T. A. Jacobson (Cambridge University Press, Cambridge, 1993), Vol. 2; e-print gr-qc/9302018.

[15] S. Popescu and L. Vaidman, Phys. Rev. A **49**, 4331 (1994).

[16] D. Beckman, Ph.D. thesis, California Institute of Technology, 2001.

[17] D. Beckman, D. Gottesman, A. Kitaev, and J. Preskill (unpublished).

[18] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996); e-print quant-ph/9604024.

[19] M. Zukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert, Phys. Rev. Lett. **71**, 4287 (1993).

[20] M. O. Scully and K. Drühl, Phys. Rev. A **25**, 2208 (1982).

[21] D. Gottesman, Phys. Rev. A **54**, 1862 (1996), e-print quant-ph/9604038.

[22] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, Phys. Rev. Lett. **78**, 405 (1997); e-print quant-ph/9605005.

[23] A. R. Calderbank and P. W. Shor, Phys. Rev. A **54**, 1098 (1996); e-print quant-ph/9512032.

[24] A. Steane, Proc. R. Soc. London, Ser. A **452**, 2551 (1996); e-print quant-ph/9601029.

[25] A. Steane, Phys. Rev. Lett. **78**, 2252 (1997); e-print quant-ph/9611027.

[26] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, London, 1993).

[27] H. Buhrman, R. Cleve, and W. van Dam, SIAM J. Comput. **30**, 1829 (2001); e-print quant-ph/9705033.

[28] R. Cleve (private communication).

[29] R. Cleve, W. van Dam, M. Nielsen, and A. Tapp, Lecture Notes Comput. Sci. **1509**, 61 (1999); e-print quant-ph/9708019.

[30] M. A. Nielsen, Ph. D. thesis, University of New Mexico, 1998; e-print quant-ph/0011036.

[31] S. Weinberg, *Dreams of a Final Theory* (Vintage, New York, 1994).

[32] T. Eggeling, D. Schlingemann, and R. F. Werner, e-print quant-ph/0104027.