Why initial system-environment correlations do not imply the failure of complete positivity: A causal perspective

David Schmid ⓞ,[1,2,*] Katja Ried,[3] and Robert W. Spekkens[1,2]

[1]*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario N2L 2Y5, Canada*
[2]*Institute for Quantum Computing and Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*
[3]*Institut für Theoretische Physik, Universität Innsbruck, Technikerstraße 21a, 6020 Innsbruck, Austria*

The common wisdom in the field of quantum information theory is that when a system is initially correlated with its environment, the map describing its evolution may fail to be completely positive. If true, this would have practical and foundational significance. We here demonstrate, however, that the common wisdom is mistaken. We trace the error to the standard proposal for how the evolution map ought to be defined. We summarize this standard proposal and then show that it sometimes fails to define a linear map or any map at all. Further, we show that these pathologies persist even in completely classical examples. Drawing inspiration from the framework of classical causal models, we argue that the correct definition of the evolution map is obtained by considering a counterfactual scenario wherein the system is reprepared independently of any systems in its causal past while the rest of the circuit remains the same, yielding a map that is always completely positive. In a postmortem of the standard proposal, we highlight two distinct mistakes that retrospectively become evident (in its application to completely classical examples): (i) The types of constraints to which it appealed are constraints on what one can *infer* about the final state of a system based on its initial state; however, such inferences reflect not just the cause-effect relation between them—which is the basis for defining the correct evolution map—but also the common-cause relation. (ii) In a (retrospectively unnecessary) attempt to introduce variability in the input state, it inadvertently introduced variability in the inference map itself, and then tried to fit the input-output pairs associated to these different maps with a single map.

I. INTRODUCTION

Quantum state evolution is typically represented by a completely positive linear map. There are two justifications that are usually given for this.

The first (the axiomatic justification) is that one always has the freedom to implement a quantum evolution map on a subsystem of some composite system, and complete positivity is then required in order for the state of the composite to remain positive, which is in turn required in order for the Born rule to return positive numbers as the probabilities for the outcomes of future measurements on the system.

The second justification notes that the evolution of an isolated system is always described by a unitary map. This implies that the most general sort of evolution of an open system arises by unitarily coupling the system to an environment (in some fixed state) and then tracing over the environment, and such evolution is always represented by a completely positive map.

However, the common wisdom in the field of quantum information is that there is an exception to the rule that the evolution of quantum states is represented by completely positive maps, namely, when the initial state of the system-environment composite does not factorize. For instance, in Sec. 8.5 of Ref. [1], Nielsen and Chuang state that

a quantum system which interacts with the degrees of freedom used to prepare that system after the preparation is complete will in general suffer a dynamics which is not adequately described within the quantum operations formalism (p. 395)

(here, "quantum operations" refers to completely positive trace-preserving linear maps), and that

it is an interesting problem for further research to study quantum information processing beyond the quantum operations formalism (p. 395).

A large body of literature has arisen to address this problem [2–19]. This literature includes a great diversity of examples to support the claim that quantum state evolution cannot always be described by a completely positive map from initial to final states of the system.

Some recent work in the literature, e.g., Refs. [20,21], has criticized non-completely-positive maps on the grounds that they lack operational significance. These authors then advocate for a more operational approach (building on Refs. [22,23]), introducing an alternative framework that avoids the issue of noncomplete positivity by proposing a different type of map [20,21,24–27] as representing the evolution of the system. The relation between this prior work and our own proposal is discussed in Sec. VII B.

Here, we take a more radical stance against non-completely-positive maps, arguing that they are explicitly at odds with the standard notion of evolution that holds

———
*dschmid@perimeterinstitute.ca

elsewhere in physics. We highlight this fact by taking the perspective of causal modeling. In particular, we do *not* assume that the evolution of a quantum state is constrained by the marginal states of the system before and after the evolution, as has previously been assumed; indeed, we show that this constraint does not hold.

In the framework of classical causal models, the evolution of a system in any setting is described by "do conditionals" [28]. These were introduced to describe the inferences that can be made from one variable to another based solely on their *cause-effect relation*, even when these variables also exhibit correlations due to a common cause. In particular, the correlations that arise between the initial and final versions of a system undergoing Markovian evolution are an example of correlations based solely on a cause-effect relation. If, on the other hand, the system is initially correlated with an environment with which it subsequently interacts, then the correlations between the initial and final versions of the system are due to a combination of cause-effect and common-cause relations. The presence of a nontrivial common-cause relation changes the sorts of *inferences* one can make about the final version of the system based on knowledge of the initial version of the system; however, crucially, the map describing such inferences is not an accurate description of the dynamics under which the system evolves.

Our criticism of the standard proposal for how to define an evolution map rests on the fact that, even for purely classical examples (in the presence of initial system-environment correlations), it contradicts the answer that one obtains by defining an evolution map using a do conditional, as one should do based on the above argument from the framework of causal inference. The thrust of our paper is to argue that, once one separates out inference and influence, one sees that the mathematical relation that the standard proposal focuses on has no relevance for the map which describes the dynamics of the system; it need not constrain the *evolution* map in any way. This is in explicit contradiction with the common wisdom.

Consequently, we argue that the common wisdom is mistaken and that the evolution of a quantum system in time is always represented by a linear and completely positive map. Our definition of the evolution map, therefore, does not lead to any of the pathological conclusions of the standard proposal, such as the failure of complete positivity, the failure of linearity, or the failure to define a map.

When attempting to resolve an (apparent) paradox, a satisfactory resolution should allow one to retrospectively identify why there was an appearance of paradox, as argued by Pearl [29]. The framework of causal modeling allows us to identify two distinct mistaken assumptions in the standard proposal and to understand precisely how the pathological conclusions arise from these mistakes.

We show that the correct evolution map depends on the marginal state of the environment but does *not* depend on the correlations between the system and the environment. This fact is *derived* from our conception of evolution and is not assumed *a priori*. Further, it holds true regardless of the nature of the correlations or the operational procedure by which the initial state was prepared.

Pursuing a quantum generalization of the classical framework of causal modeling has already had many interesting applications in quantum foundations, including revealing a quantum advantage for causal inference [30], uncovering new experimental scenarios wherein there is a gap between quantum and classical correlations [31–37], uncovering a promising approach to achieving a causal explanation of Bell inequality violations without fine-tuning [38–42], expanding the set of experimental configurations wherein one can achieve quantum state pooling [43], and exploring the possibility of quantum uncertainty about the causal structure [44–46].

This article continues this trend by showing that the correct definition of a quantum evolution map in the presence of initial system-environment correlations requires a quantum generalization of a key notion from classical causal modeling, that of the do conditional, thereby demonstrating the conceptual significance of the quantum version of this notion.

A. Outline of the paper

Section II lays out the standard argument for the inadequacy of completely positive maps for describing evolution, abstracted from the various perspectives on the subject found in the literature. The argument is based on a proposal, which we term the *standard proposal*, for how to define the evolution map in the scenarios of interest. We give several examples in which the standard proposal leads to problematic conclusions, and we explain why these conclusions cast doubt on the validity of the proposal.

In Sec. III, we show that the pathological consequences of the standard proposal arise even in purely classical scenarios. In Sec. IV, we describe how to define a classical evolution map in the presence of initial system-environment correlations, using the notion of do conditionals from the framework of classical causal models. In Sec. V, we show that the prescription of the standard proposal generally fails to reproduce it, and we leverage the framework of classical causal models to elucidate the underlying mistaken assumptions within the scope of classical scenarios.

In Sec. VI, we show that the definition of the classical evolution map in terms of do conditionals generalizes naturally to a definition of the quantum evolution map, which is seen to be always completely positive and to avoid the pathologies of the standard proposal. We then assess the mistakes of the standard proposal in the quantum sphere.

Finally, in Sec. VII, we discuss some of the implications of our definition of the quantum evolution map, we explain how it can be extracted from experimental data, and we advocate for the study of open-system dynamics from the causal modeling perspective.

II. PROBLEMS WITH THE STANDARD PROPOSAL

A. Preliminaries

First, we recall some mathematical facts. Let $\mathcal{L}(\mathcal{H}_S)$ denote the space of linear operators on the Hilbert space $\mathcal{H}_S$ describing system S. We denote a map from $\mathcal{L}(\mathcal{H}_{S_1})$ to $\mathcal{L}(\mathcal{H}_{S_2})$ by $\mathcal{E}_{S_2|S_1}$. Such a map is said to be *trace preserving* if $\forall \rho_{S_1} \in \mathcal{L}(\mathcal{H}_{S_1}) : \mathrm{Tr}_{S_2}[\mathcal{E}_{S_2|S_1}(\rho_{S_1})] = \mathrm{Tr}_{S_1}(\rho_{S_1})$. It is said to be *positivity preserving*, or simply *positive*, if it takes all positive
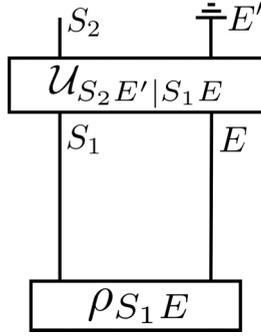
FIG. 1. The scenario of interest. An initially correlated state of system and environment, ρ_{S_1E}, evolves according to a quantum channel $\mathcal{E}_{S_2|S_1E}$.

operators to positive operators, $\forall \rho_{S_1} \geqslant 0 : \mathcal{E}_{S_2|S_1}(\rho_{S_1}) \geqslant 0$. It is said to be *completely positivity preserving*, or simply *completely positive*, if its action on a composite system is also positive; that is, for any ancillary system, denoted E_1 at the initial time and E_2 at the final time, and evolving by the identity map, $\mathrm{id}_{E_2|E_1}$, we have $\forall \rho_{S_1E_1} \geqslant 0$, $(\mathcal{E}_{S_2|S_1} \otimes \mathrm{id}_{E_2|E_1})(\rho_{S_1E_1}) \geqslant 0$.

Suppose that the principal system S is coupled to an ancillary system E by a unitary U_{SE}, that E is prepared at the initial time in the state ρ_E, and that one traces over the ancillary system to obtain the final state of the principal system. The evolution of the principal system is then represented by the map $\mathcal{E}_{S_2|S_1}(\square_{S_1}) : \mathcal{L}(\mathcal{H}_{S_1}) \to \mathcal{L}(\mathcal{H}_{S_2})$ defined by

$$\mathcal{E}_{S_2|S_1}(\square_{S_1}) = \mathrm{Tr}_E\big[U_{S_1E}(\square_{S_1} \otimes \rho_E)U_{S_1E}^\dagger\big]. \tag{1}$$

Note that throughout this article we denote the argument of a map by $\square_A$, where the subscript specifies the type of system at the input. Clearly, such a map is completely positive and trace preserving. It turns out, furthermore, that *any* completely positive trace-preserving map can be realized in this fashion, a result known as the *Stinespring dilation* theorem [47].

B. The standard argument for the inadequacy of completely positive maps

We now review the standard argument for the inadequacy of completely positive maps in describing the evolution of the quantum state of the principal system (which we henceforth simply call "the system") when it is initially correlated with the environment [1–3,5–9,11–14,16,22].

To assume that the system and environment are initially correlated is to assume that their joint state does not factorize, that is, $\rho_{S_1E} \neq \rho_{S_1} \otimes \rho_E$. The system and environment are imagined to subsequently interact according to the map

$$\mathcal{U}_{S_2E'|S_1E}(\square_{S_1E}) := U \square_{S_1E} U^\dagger, \tag{2}$$

where $U : \mathcal{H}_{S_1} \otimes \mathcal{H}_E \to \mathcal{H}_{S_2} \otimes \mathcal{H}_{E'}$ is a unitary operator. This scenario is depicted in Fig. 1.

In our figures, unadorned wires represent the interacting systems (termed the "system" and the "environment'") , while wires adorned with arrows represent classical variables which can either be controlled (if they are inputs) or postselected on (if they are outputs), and the ground symbol represents the trace operation. Boxes represent states, channels, or sets

thereof, in accordance with their associated inputs and outputs (and as denoted by the explicit labels).

The standard argument is predicated on a particular proposal for how to define the evolution map from S_1 to S_2, which we call the *standard proposal*. (We will argue in the following that the standard argument is mistaken precisely because the standard proposal is conceptually misguided.) The standard proposal begins by assuming that the evolution map, which we denote here by $\mathcal{E}_{S_2|S_1}$, is constrained by the following equation:

$$\rho_{S_2} = \mathcal{E}_{S_2|S_1}(\rho_{S_1}), \tag{3}$$

where ρ_{S_1} denotes the marginal on S_1,

$$\rho_{S_1} := \mathrm{Tr}_E(\rho_{S_1E}), \tag{4}$$

and ρ_{S_2} denotes the marginal on S_2,

$$\rho_{S_2} := \mathrm{Tr}_{E'}\big(\mathcal{U}_{S_2E'|S_1E}(\rho_{S_1E})\big). \tag{5}$$

However, Eq. (3) only specifies how a single state of the system, namely ρ_{S_1}, is transformed. This is clearly not sufficient to determine how an arbitrary state on S_1 is transformed, and therefore Eq. (3) does not serve to define a map uniquely.[1]

In order to define a map uniquely, it is critical that its action be specified on many different input states. Toward this end, most articles on the topic do not consider a single joint state on system and environment but rather a set of such states with differing marginal states for the system. The specific means by which this variation is generated differs among proposals: One might apply a transformation on the system, one might apply a joint transformation on the system and the environment, or one might imagine performing some nondestructive measurement on the system. Physically, it must be that some sort of laboratory operation induces variation on the initial system-environment state. We denote the random variable which encodes the setting of this operation by J and the random variable which encodes its outcome by K. For each pair of values (j, k) for these variables, the system-environment composite is prepared in a corresponding state $\rho_{S_1E}^{(j,k)}$.

The general circuit diagram representing the class of scenarios studied in the literature, then, is that of Fig. 2.

The standard proposal asks us to consider the mathematical relation

$$\mathfrak{R} := \big\{\big(\rho_{S_1}^{(j,k)}, \rho_{S_2}^{(j,k)}\big)\big\}_{j,k}, \tag{6}$$

where

$$\rho_{S_1}^{(j,k)} := \mathrm{Tr}_E\big(\rho_{S_1E}^{(j,k)}\big) \tag{7}$$

and

$$\rho_{S_2}^{(j,k)} := \mathrm{Tr}_{E'}\big[\mathcal{U}_{S_2E'|S_1E}\big(\rho_{S_1E}^{(j,k)}\big)\big], \tag{8}$$

which we term the *input-output relation*. The standard proposal asserts that the evolution map $\mathcal{E}_{S_2|S_1} : \mathcal{L}(\mathcal{H}_{S_1}) \to \mathcal{L}(\mathcal{H}_{S_2})$ should satisfy the constraints

$$\forall j, k : \quad \mathcal{E}_{S_2|S_1}\big[\rho_{S_1}^{(j,k)}\big] = \rho_{S_2}^{(j,k)}. \tag{9}$$

---

[1]In particular, any single pair of states ρ_{S_1} and ρ_{S_2} is consistent with the map that ignores the state of S_1 and simply prepares ρ_{S_2}, $\mathcal{E}_{S_2|S_1}(\square_{S_1}) = \rho_{S_2}\mathrm{Tr}_{S_1}(\square_{S_1})$.
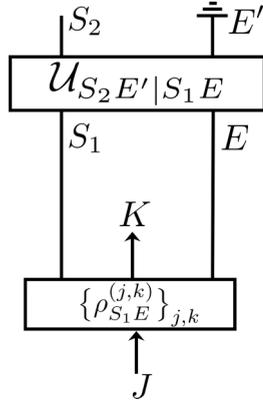
FIG. 2. The most general circuit considered in the standard proposal.

Note that one could only hope to uniquely define a map on all input states in this manner if the map is assumed to be linear and the domain of the input-output relation includes an informationally complete set of states [i.e., the set $\{\rho_{S_1}^{(j,k)}\}_{j,k}$ forms a basis of the operator space $\mathcal{L}(\mathcal{H}_{S_1})$].

The standard argument for the inadequacy of completely positive maps for describing the evolution of the quantum state of the system is concluded by noting that, in many scenarios, the map which is defined by Eq. (9) fails to be completely positive. We provide a simple example of this type in the next section.

In fact, one can find examples wherein the prescription that is endorsed by the standard argument defines a map that is not linear, and other examples where it does not define any map at all. We provide such examples in the two subsequent sections. For each type of failure—the failure of complete positivity, the failure of linearity, and the failure to define a map—we explain why it casts doubt on the standard argument.

Before moving to these examples, we pause to note a problematic feature that lies at the very base of the standard proposal.

The question that we believe ought to be answered is the following one:

Q: What is the evolution map from S_1 to S_2 in the experimental scenario of Fig. 1?

However, the standard proposal immediately substitutes this question for a new one:

Q′: What is the evolution map from S_1 to S_2 in the experimental scenario of Fig. 2?

As noted earlier, the motivation for the substitution is that there seems to be *no way* to answer question Q if one is committed to the standard proposal, that is, if one is committed to defining the evolution map using an input-output relation for the input states that are realized in the scenario at hand. We will ultimately argue that this commitment is mistaken and that there is consequently no need to retreat from Q to Q′. For the moment, however, we wish simply to note a problem with any such retreat. Namely, it implies a violation of a *criterion of universality* that we believe ought to be upheld in such investigations:
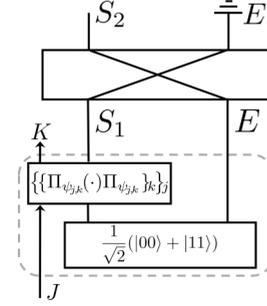
FIG. 3. An example where the input-output relation corresponds to a map that is linear but not completely positive.

Any proposal for the evolution map should be applicable to any scenario.

In particular, it should be applicable to the natural scenario depicted in Fig. 1.

C. An example where the standard proposal implies a map that is not completely positive

This example is motivated by related examples in Refs. [11,48].

Imagine that one achieves a variation over the initial state of the system-environment composite as follows: One first prepares the maximally entangled state between system and environment, $|\phi^+\rangle_{SE} := \frac{1}{\sqrt{2}}(|0\rangle_S|0\rangle_E + |1\rangle_S|1\rangle_E)$, and then, for each value of a setting variable J, one implements upon it the binary-outcome measurement associated to the orthogonal basis $\{|\psi_{j,1}\rangle, |\psi_{j,2}\rangle\}$ with von Neumann–Lüders state update rule and postselects on obtaining the first outcome, $K = 1$. Here, $|\psi_{j,1}\rangle$ and $|\psi_{j,2}\rangle$ form an orthogonal basis for the qubit Hilbert space.

This scenario is depicted in Fig. 3. The part of the circuit that is conceptualized as the preparation of the joint state of system and environment (that is, the part which corresponds to the first gate in Fig. 2) is highlighted by a dashed box. This convention is followed in all of the examples we consider.

Because the environment is initially correlated with the system, the state that one infers for it from the postselection depends on the choice of basis and on the outcome, a phenomenon known as *quantum steering* [49,50]. Specifically, when one learns from the measurement that the quantum state of the system is $|\psi_{j,1}\rangle$, one infers that the state of the environment is $|\psi_{j,1}^T\rangle$, where T denotes transposition in the computational basis (since $\langle\psi|(|\phi^+\rangle\langle\phi^+|)|\psi\rangle = |\psi^T\rangle\langle\psi^T|$ for any ψ).

Thus, by varying the parameter J and conditioning on obtaining the first outcome of the measurement, $K = 1$, one obtains the set $\{\rho_{S_1E}^{(j,1)}\}_j$ where

$$\rho_{S_1E}^{(j,1)} = |\psi_{j,1}\rangle\langle\psi_{j,1}|_{S_1} \otimes |\psi_{j,1}^T\rangle\langle\psi_{j,1}^T|_E. \tag{10}$$

The subsequent system-environment interaction is presumed to implement the swap operation on their states. That is, $\mathcal{U}_{S_2E'|S_1E}$ is defined, via Eq. (2), by the operator

$$U = \text{SWAP}, \tag{11}$$

where $\text{swap}(|\phi\rangle \otimes |\chi\rangle) := |\chi\rangle \otimes |\phi\rangle$ for all χ and ϕ.

By Eq. (8), we infer that the marginal states of the system (for each value of J) after the system-environment interaction are given by

$$\rho_{S_2}^{(j,1)} = \left|\psi_{j,1}^T\right\rangle\!\left\langle\psi_{j,1}^T\right|_{S_2}. \tag{12}$$

The input-output relation in this case, therefore, is

$$\mathfrak{R} = \left\{\left(|\psi_{j,1}\rangle\langle\psi_{j,1}|_{S_1}, \left|\psi_{j,1}^T\right\rangle\!\left\langle\psi_{j,1}^T\right|_{S_2}\right)\right\}_j. \tag{13}$$

If the projectors $\{|\psi_{j,1}\rangle\langle\psi_{j,1}|_{S_1}\}_j$ form a basis for the operator space $\mathcal{L}(\mathcal{H}_{S_1})$ (i.e., an informationally complete set), one can conclude that there is a unique linear map defined [via Eq. (9)] by this relation, namely, the transpose map. This is the canonical example of a map that is positive but not completely positive [1].

a. Problematic implications of the failure of complete positivity. Even though the failure of completely positivity is obviously acknowledged by proponents of the standard argument (it is the reason this field of research even exists), we believe that it already provides good reasons for being suspicious of the argument.

An immediate worry is that such maps could lead to output states on system-ancilla composites that fail to be positive.

The standard response to this worry (found, for instance, in Refs. [8,9]) is that the map describing the evolution of the system in this circumstance is only applicable on a limited domain of input states and that this domain of input states does not include the marginals of the set of entangled states which manifest the failure of the map to be completely positive.

Note, first of all, that this response is a denial of the axiomatic justification of complete positivity which we discussed in the introduction. The problem we see with this denial is that it forces one to give up on the notion that the map describing the evolution can support inferences about counterfactuals. This notion is central to the notion of evolution in physics: Laws of motion are not just *descriptions* of historically actual motions, but *prescriptions* for determining what motion *would* occur for *any* initial condition. The possibility of making inferences about counterfactual scenarios is precisely what makes laws of motion so useful in practice. To entertain the idea that a map *only* describes the evolution of a system *when that system is assured to be in one of a restricted set of states* is to retreat from the usual conception of an evolution map.

D. An example where the standard proposal implies a map that is not linear

The following example is a simplified version of the one presented in Nielsen and Chuang [1].

Here, one begins with the product state $|0\rangle_{S_1}|0\rangle_E$ and, conditioned on the classical control $J \in \{0, 1\}$, one applies a controlled-Hadamard gate on the system and a controlled-NOT gate on the environment. This procedure does not involve any measurement, and hence K is trivial. The system-environment composite is therefore prepared in one of the states $\{\rho_{S_1E}^{(j)}\}_j$ where

$$\rho_{S_1E}^{(0)} = |0\rangle\langle0|_{S_1} \otimes |0\rangle\langle0|_E, \tag{14}$$

$$\rho_{S_1E}^{(1)} = |+\rangle\langle+|_{S_1} \otimes |1\rangle\langle1|_E, \tag{15}$$
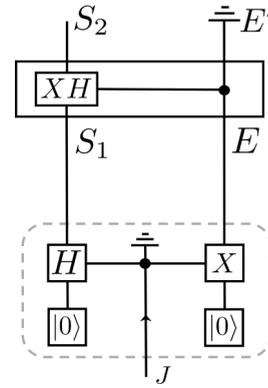
FIG. 4. An example where the input-output relation corresponds to a map that is nonlinear.

with $|+\rangle := \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$.

Next, there is a controlled-XH gate with the environment qubit as control and the system qubit as target,

$$U = \mathbb{1} \otimes |0\rangle\langle0| + \text{XH} \otimes |1\rangle\langle1|, \tag{16}$$

where XH denotes the unitary gate obtained by performing the Hadamard gate H followed by the Pauli gate X.

This circuit is shown in Fig. 4.

By Eq. (8), we infer that the marginal states of the system after the system-environment interaction (for each value of J) are given by

$$\rho_{S_2}^{(0)} = |0\rangle\langle0|_{S_2}, \tag{17}$$

$$\rho_{S_2}^{(1)} = |1\rangle\langle1|_{S_2}. \tag{18}$$

The input-output relation in this case, therefore, is

$$\mathfrak{R} = \left\{\left(|0\rangle\langle0|_{S_1}, |0\rangle\langle0|_{S_2}\right), \right. \tag{19}$$

$$\left. \left(|+\rangle\langle+|_{S_1}, |1\rangle\langle1|_{S_2}\right)\right\}. \tag{20}$$

Note that the map implied by the standard proposal is not completely specified by these constraints. However, any map consistent with this relation must take nonorthogonal states to orthogonal states, and every such map is nonlinear. (Note that once one allows for the failure of linearity, even having a set of input states that span $\mathcal{L}(\mathcal{H}_{S_1})$ becomes insufficient to determine how the map acts on *all* states.)

Problematic implications of the failure of linearity. The fact that the standard proposal does not always define a linear map is troubling, because the linearity of transformations can be justified on numerous physical grounds. For example, representing a process by a nonlinear map violates the principle that processing of a system cannot increase the amount of information it contains about another system. Specifically, every such map violates the data processing inequality [1]. Such violations have physically problematic implications, such as the possibility of superluminal signaling [51–53].

E. An example where the standard proposal implies a relation that is not a map

Most pathologically, the standard proposal can yield a relation which is inconsistent with *any* map whatsoever. This
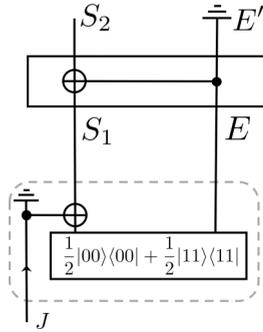
FIG. 5. An example where the input-output relation does not correspond to any map.

is illustrated by a simple example, motivated by one from Ref. [5] and pictured in Fig. 5.

Here, K is again trivial, and for each value of $J \in \{0, 1\}$, one prepares a distinct initial joint state by performing a controlled-NOT gate on the system with J as the control. In this way, the system-environment composite is prepared in one of the states $\{\rho_{S_1 E}^{(j)}\}_j$, namely

$$\rho_{S_1 E}^{(0)} = \tfrac{1}{2}|0\rangle\langle 0|_{S_1} \otimes |0\rangle\langle 0|_E + \tfrac{1}{2}|1\rangle\langle 1|_{S_1} \otimes |1\rangle\langle 1|_E, \quad (21)$$

$$\rho_{S_1 E}^{(0)} = \tfrac{1}{2}|1\rangle\langle 1|_{S_1} \otimes |0\rangle\langle 0|_E + \tfrac{1}{2}|0\rangle\langle 0|_{S_1} \otimes |1\rangle\langle 1|_E. \quad (22)$$

Next, there is a controlled-NOT gate with the environment qubit as control and the system qubit as target, with unitary description

$$U = \mathbb{1} \otimes |0\rangle\langle 0| + X \otimes |1\rangle\langle 1|. \quad (23)$$

By Eq. (8), we infer that the marginal states of the system (for each value of J) after the system-environment interaction are given by

$$\rho_{S_2}^{(0)} = |0\rangle\langle 0|_{S_2}, \quad (24)$$

$$\rho_{S_2}^{(1)} = |1\rangle\langle 1|_{S_2}. \quad (25)$$

The input-output relation defined by this scenario is

$$\mathfrak{R} = \left\{ \left(\tfrac{1}{2}\mathbb{1}_{S_1}, |0\rangle\langle 0|_{S_2} \right), \left(\tfrac{1}{2}\mathbb{1}_{S_1}, |1\rangle\langle 1|_{S_2} \right) \right\}. \quad (26)$$

Because this relation is one to many, it does not define a map from $\mathcal{L}(\mathcal{H}_{S_1})$ to $\mathcal{L}(\mathcal{H}_{S_2})$.

Problematic implications of the failure of the standard proposal to define a map. The standard proposal purports to define the evolution map. However, we have just seen that it sometimes fails to define *any* map whatsoever. This failure strikes us as a decisive criticism. Proponents of the standard proposal have not offered any satisfactory account of how their scheme can be salvaged in the face of this failure.

F. Every pathology can be obtained from every circuit type

We have provided three examples which illustrate distinct pathologies of the standard proposal. Much previous work [2–4,6–9,11–14,16,22] has focused on these distinctions, e.g., by seeking necessary or sufficient conditions for the relation to define a completely positive map, a linear map, and so on. The coming analysis shows that these questions are misguided: Evolution maps are *always* completely positive, and all of the

apparent counterexamples are really just indicative of the fact that the standard proposal is *not* the correct way to define the evolution map.

Our examples differed also in the way in which they introduced variability in the initial state of the system-environment composite. Specifically, the variability was introduced (respectively) by (a) the choice of transformation on the system-environment composite, (b) the choice of measurement on the system and the choice of postselection on its outcomes, and (c) the choice of transformation on the system alone. Roughly, all examples in the literature fit into one of these three special cases of the circuit of Fig. 2. Some articles, however, left the operational (circuit) description of the problem unspecified [5,6,8,9,11–14], and took the problem description to be a (possibly continuous) set of initial states on $S_1 E$ and a unitary system-environment interaction, $\mathcal{U}_{S_2 E'|S_1 E}$. One consequence of our work is to show that this version of the problem is not well posed; knowing the causal structure of the circuit is in fact critical to defining the evolution map.[2]

Apart from their simplicity, there is nothing special about the examples we have chosen. In Appendix A, we show that each of the three circuit types can generate each of the pathologies of the standard proposal. In particular, this means that one cannot evade such pathologies by restricting attention to one of the three special classes of circuits.

As we argued at the end of Sec. II B, any sensible definition of the evolution map should satisfy a criterion of *universality*, namely, that it should be applicable regardless of the scenario. As we already noted in that section, the standard proposal fails to satisfy this criterion because it does not define an evolution map in the simple scenario of Fig. 1. We are also not satisfied with the standard responses to the problematic implications of the standard proposal just outlined, because they too compromise on the criterion of universality. Indeed, they do so in a particularly unsatisfying way. Namely, they assert that their proposed definition of the evolution map is only applicable to a restricted class of circuits, but the restriction is *ad hoc* insofar as no justification is given other than to avoid the pathologies that would otherwise result. For example, Refs. [17,54] recognized that certain scenarios yielded the pathology discussed in Sec. II E and sought to avoid it by excluding such scenarios *by fiat*.

III. PERSISTENCE OF THE PROBLEMS OF THE STANDARD PROPOSAL IN THE CLASSICAL SPHERE

In a classical setting, one can also consider the evolution map for a system when there are initial correlations between the system and environment. In this section, we demonstrate that if one tries to define this classical evolution map using the prescription endorsed by the standard proposal, then one obtains all of the same problematic implications that one saw quantumly. (This should already be evident given that for many of the quantum examples we presented, both above and in the Appendix, all the states and maps could be dephased in the computational basis without affecting our conclusions.)

---

[2]There is, however, some prior acknowledgment of the importance of the form of the circuit in Refs. [22,23].

Therefore, although it is a widely held belief that the surprising form of quantum evolution maps in cases of initial system-environment correlations is just another example of a counterintuitive feature of quantum theory (so-called "quantum weirdness"), we will demonstrate that it should be taken instead as evidence of the fallacy of the standard proposal for how to define the evolution map.

A. Classical preliminaries

A classical system is described by a set of physical states, which can be encoded as values of a random variable S. The statistical state of the classical system is given by a probability distribution P_S over the possible physical states of the system. More precisely, if Λ_S denotes the set of possible values of S (here assumed discrete), then $P_S : \Lambda_S \to [0, 1]$, where $P_S(s)$ denotes the probability that $S = s$. P_S satisfies the normalization condition $\sum_S P_S = 1$, where $\sum_S P_S$ denotes the function on S defined by $\forall s \in \Lambda_S : (\sum_S P_S)(s) := \sum_s P_S(s)$, and where the right-hand side of the condition denotes the function that takes value 1 for all s.

We are interested in the evolution of these probability distributions for a system interacting with an environment. We will denote the space of probability distributions on S by $\mathfrak{P}(S)$, and for later convenience we introduce the following notation:

$$P_A = [a] \text{ means } P_A(a) = 1, \qquad (27)$$

$$P_{AB} = [a][b] \equiv [ab] \text{ means } P_{AB}(a, b) = 1. \qquad (28)$$

A map $\Gamma_{S_2|S_1} : \mathfrak{P}(S_1) \to \mathfrak{P}(S_2)$ is called *stochastic* if there exists a conditional probability distribution $P_{S_2|S_1}$ such that

$$\Gamma_{S_2|S_1}(P_{S_1}) = \sum_{S_1} P_{S_2|S_1} P_{S_1}. \qquad (29)$$

Note that this is a linear map.

The most general manner in which a probability distribution on a classical system can evolve is by a stochastic map. This can be justified, analogously to complete positivity in the quantum case, as follows: If one imagines that an ancillary system is prepared in some fixed distribution, the system-ancilla composite is subjected to a deterministic dynamics, and then the ancilla is ignored, then the resulting map on the probability distribution over the system is always stochastic. Stochastic maps are also the most general type of map that preserve positivity for all input distributions.

B. What the standard proposal stipulates in a classical scenario

The standard proposal applies just as well to classical scenarios as to quantum scenarios, since the former are a strict subset of the latter, where all operators are diagonal in some fixed basis. A classical scenario can be described by the same sort of circuit as a quantum scenario: One simply replaces quantum states by probability distributions and unitary operations by deterministic functions. Hence, we can consider a direct analogy for each of the quantum circuits considered previously.

The classical analog for the general circuit from Fig. 1 is shown in Fig. 6(a). If one followed the standard proposal, one
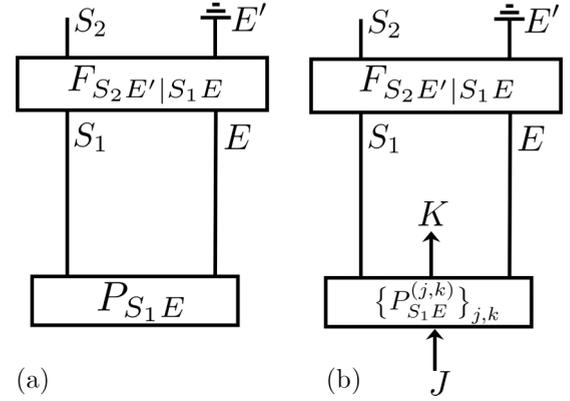
FIG. 6. (a) The classical analog of the circuit in Fig. 1. (b) The classical analog of the circuit in Fig. 2.

would demand that a constraint on the evolution map, which we denote by $\Gamma_{S_2|S_1}$, is

$$P_{S_2} = \Gamma_{S_2|S_1}(P_{S_1}), \qquad (30)$$

with $P_{S_1} := \sum_E P_{S_1 E}$ and $P_{S_2} := \sum_{E'} F_{S_2 E'|S_1 E}(P_{S_1 E})$. One would again be compelled to introduce variability in the input state on S_1 in a manner analogous to what was done in the quantum sphere in Fig. 2.

The resulting scenario, a classical analog for the general circuit from Fig. 2, is shown in Fig. 6(b), where $F_{S_2 E'|S_1 E}$ represents a stochastic map $\mathfrak{P}(S_1) \otimes \mathfrak{P}(E) \to \mathfrak{P}(S_2) \otimes \mathfrak{P}(E')$ induced by a deterministic dynamics, so that

$$F_{S_2 E'|S_1 E}(\square_{S_1 E}) = \sum_{S_1 E} \delta_{S_2, f(S_1, E)} \delta_{E', g(S_1, E)} \square_{S_1 E} \qquad (31)$$

for some functions f and g.

Applying the standard proposal to these classical scenarios, one simply computes

$$P_{S_1}^{(j,k)} = \sum_E P_{S_1 E}^{(j,k)} \qquad (32)$$

and

$$P_{S_2}^{(j,k)} = \sum_{E'} F_{S_2 E'|S_1 E}(P_{S_1 E}^{(j,k)}), \qquad (33)$$

generating an input-output relation

$$\mathfrak{R} = \left\{ \left(P_{S_1}^{(j,k)}, P_{S_2}^{(j,k)} \right) \right\}_{j,k}, \qquad (34)$$

as in the quantum case. The standard proposal dictates that the map describing the evolution should be taken to be one which satisfies the constraints encoded in this relation; that is, a map $\Gamma_{S_2|S_1} : \mathfrak{P}(S_1) \to \mathfrak{P}(S_2)$ for which

$$\forall j, k : \ \Gamma_{S_1|S_1}[P_{S_1}^{(j,k)}] = P_{S_2}^{(j,k)}. \qquad (35)$$

As we now show, this prescription for how to define the evolution map leads to the same sorts of pathologies we saw in the quantum case.

There exist simple physical scenarios that generate relations which imply maps that are not stochastic, others which imply maps that are not linear, and still others which do not define any map at all.

Our examples of each of the three failures are chosen to be analogous to the corresponding quantum examples. Note

that S_1, S_2, and E are now taken to denote classical random variables rather than being mere labels of systems.

C. A classical example where the standard proposal implies a map that is not stochastic

Our example of the failure of stochasticity in a classical system is analogous to our example of the failure of complete positivity in a quantum system (described in Sec. II C) and is realized with a circuit of exactly the same form (shown in Fig. 3).

Consider a system and an environment that each have four possible physical states, so that $S_1, E \in \{0, 1, 2, 3\}$. The pair are first prepared in the joint distribution $P_{S_1 E} := \frac{1}{4}([00] + [11] + [22] + [33])_{SE}$. System S_1 is then subjected to one of three measurements, determined by the value of a ternary variable J, and each measurement having a binary outcome K. The $J = 1$ measurement determines whether $S_1 \in \{2, 3\}$ or not, $J = 2$ whether $S_1 \in \{1, 3\}$ or not, and $J = 3$ whether $S_1 \in \{1, 2\}$ or not. $K = 1$ labels the outcome wherein S_1 is found to be in the given set, and $K = 0$ labels the complementary set. We further imagine that these measurements are not passive but disturb the value of the system variable. Specifically, the update rule is such that the final state on S_1 depends only on J and K: S_1 is prepared in a uniform distribution over the values of S_1 in the complementary set to the one that S_1 was found in. For instance, if the $J = 1$ measurement is done and the outcome $K = 1$ occurs, verifying that $S_1 \in \{2, 3\}$, then S_1 is reprepared in the distribution $\frac{1}{2}[0]_{S_1} + \frac{1}{2}[1]_{S_1}$.

Finally, S_1 undergoes a swap operation with E.

In our example, we consider only the cases where the outcome is found to be $K = 1$. For each possible value j of J, the joint distribution on S_1 and E, $P_{S_1 E}^{(j,1)}$, is as follows:

$$P_{S_1 E}^{(1,1)} = \frac{1}{2}([0] + [1])_{S_1} \frac{1}{2}([2] + [3])_E, \tag{36}$$

$$P_{S_1 E}^{(2,1)} = \frac{1}{2}([0] + [2])_{S_1} \frac{1}{2}([1] + [3])_E, \tag{37}$$

$$P_{S_1 E}^{(3,1)} = \frac{1}{2}([0] + [3])_{S_1} \frac{1}{2}([1] + [2])_E. \tag{38}$$

It follows that after the swap operation on the system and environment, the marginal state of S_2, for each value j of J and for $K = 1$, denoted $P_{S_2}^{(j,1)}$, is as follows:

$$P_{S_2}^{(1,1)} = \frac{1}{2}([2] + [3])_{S_2}, \tag{39}$$

$$P_{S_2}^{(2,1)} = \frac{1}{2}([1] + [3])_{S_2}, \tag{40}$$

$$P_{S_2}^{(3,1)} = \frac{1}{2}([1] + [2])_{S_2}. \tag{41}$$

The input-output relation, therefore, is

$$\mathfrak{R} = \left\{ \left[\frac{1}{2}([0] + [1]), \frac{1}{2}([2] + [3]) \right], \right.$$
$$\left[\frac{1}{2}([0] + [2]), \frac{1}{2}([1] + [3]) \right],$$
$$\left. \left[\frac{1}{2}([0] + [3]), \frac{1}{2}([1] + [2]) \right] \right\}. \tag{42}$$

This relation is consistent with a linear map,

$$\Gamma_{S_2|S_1}[\vec{p}_{S_1}] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \cdot \vec{p}_{S_1}, \tag{43}$$

where the probability distribution over the four physical states of S_1 is expressed as a vector, $\vec{p}_{S_1}$. However, this map is not stochastic (e.g., the columns do not each sum to one), and in fact there is *no* stochastic map consistent with the relation $\mathfrak{R}$ above. To see this, note that the first ordered pair in $\mathfrak{R}$ guarantees that physical state 0 has no probability of mapping to 0 or 1, the second ordered pair guarantees that it has no probability of mapping to 0 or 2, and the third ordered pair guarantees that it has no probability of mapping to 0 or 3. But every stochastic map is certain to map physical state 0 to *some* other physical state, in contradiction with these three constraints.

D. A classical example where the standard proposal implies a map that is not linear

Next, we provide a classical example in which the standard proposal yields a map $\Gamma_{S_2|S_1}$ that fails to be linear. This is analogous to the failure of linearity in the quantum example of Sec. II D, and it is realized in a classical circuit of exactly the same form (shown in Fig. 4).

The system-environment composite is prepared in one of two possible states, depending on the value of J:

$$P_{S_1 E}^{(0)} = [0]_{S_1} \otimes [0]_E, \tag{44}$$

$$P_{S_1 E}^{(1)} = \frac{1}{2}([0]_{S_1} + [1]_{S_1}) \otimes [1]_E. \tag{45}$$

The system-environment composite then undergoes the following joint evolution: If $E = 0$, then $S_2 = S_1$, while if $E = 1$, then $S_2 = S_1 \oplus_3 1$, where $\oplus_d$ denotes summation modulo d. (Note that the principal system has three distinct states in this example.)

The marginal states of the system (for each value of J) after the system-environment interaction are given by

$$P_{S_2}^{(0)} = [0]_{S_2}, \tag{46}$$

$$P_{S_2}^{(1)} = \frac{1}{2}([1] + [2])_{S_2}. \tag{47}$$

The input-output relation defined by this transformation is

$$\mathfrak{R} = \left\{ ([0], [0]), \left[\frac{1}{2}([0] + [1]), \frac{1}{2}([1] + [2]) \right] \right\}. \tag{48}$$

Any map consistent with this relation must take overlapping distributions to nonoverlapping distributions, and every such map is nonlinear.

E. A classical example where the standard proposal does not define a map

Finally, we provide a classical example in which the standard proposal fails to yield any map at all. This is an *exact* analog of the quantum example of Sec. II E, so it is realized in a classical circuit of exactly the same form (shown in Fig. 5).

As in the quantum example, K is trivial, and for each value of $J \in \{0, 1\}$, one prepares a distinct initial joint state by performing a controlled-NOT gate on the system. In this way, the system-environment composite is prepared in one of the distributions $\{P_{S_1 E}^{(j)}\}_j$, namely,

$$P_{S_1 E}^{(0)} = \frac{1}{2}[0]_{S_1} \otimes [0]_E + \frac{1}{2}[1]_{S_1} \otimes [1]_E, \tag{49}$$

$$P_{S_1 E}^{(1)} = \frac{1}{2}[1]_{S_1} \otimes [0]_E + \frac{1}{2}[0]_{S_1} \otimes [1]_E. \tag{50}$$

As before, the system-environment interaction is a controlled-NOT gate with the environment qubit as control and the system qubit as target, so that if $E = 0$, then $S_2 = S_1$, while if $E = 1$, then $S_2 = S_1 \oplus_2 1$.

The marginal states of the system (for each value of J) after the system-environment interaction are given by

$$P_{S_2}^{(0)} = [0]_{S_2}, \tag{51}$$

$$P_{S_2}^{(1)} = [1]_{S_2}. \tag{52}$$

The corresponding input-output relation is

$$\mathfrak{R} = \left\{ \left[\tfrac{1}{2}([0] + [1]), [0] \right], \left[\tfrac{1}{2}([0] + [1]), [1] \right] \right\}. \tag{53}$$

Since it is one to many, there is no map consistent with this relation.

F. Every classical pathology can be obtained from every classical circuit type

We have designed these specific classical examples to make the point as simply as possible (while maintaining a close analogy with the three quantum examples presented previously). As we show in Appendix B, however, there are infinite families of examples which exhibit each of the three pathologies, no matter which type of operational scenario [(a), (b), and (c) in Sec. II F above] one considers. These generic classical examples are analogous to the generic quantum examples in Appendix A.

IV. THE CORRECT DEFINITION OF THE EVOLUTION MAP IN THE CLASSICAL SPHERE

In the quantum literature, some of the problematic examples we described earlier led researchers to question the notion that the evolution of quantum states is always described by a completely positive map. By contrast, no one has previously seen fit to cite the kinds of problematic classical examples that we have just described as a reason to question the claim that evolution of classical probability distributions is always described by a stochastic map. Why not? Could it simply be a failure to recognize the existence of problematic classical examples?

No. The reason no one has questioned the adequacy of stochastic maps is because in the classical sphere no one was inclined to endorse the classical analog of the standard proposal for how to define the evolution map. Rather, the framework of causal modeling and the so-called *do calculus* was developed [28], and this provided a scheme for inferring the evolution map on the system even in the presence of initial correlations between system and environment. It is found to differ significantly from that of the standard proposal, and in particular it avoids all of the problematic implications. In this section, we will present this scheme, explain why it clearly yields the correct notion of an evolution map in the classical realm, and identify the mistaken assumptions in the standard proposal.

The correct definition of an evolution map in the classical sphere generalizes easily to the quantum sphere, as we will show in Sec. VI.
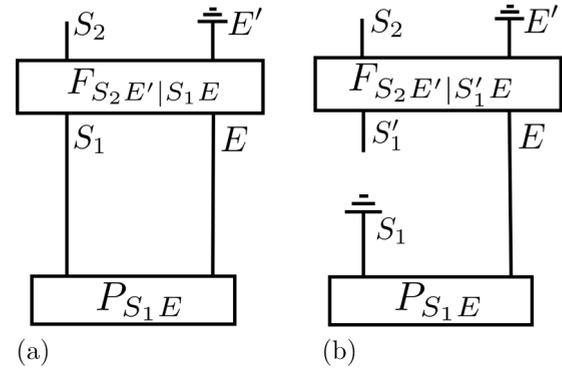
FIG. 7. (a) The classical analog of the circuit in Fig. 1. (b) The hypothetical circuit which aids in defining the evolution map for the situation in panel (a).

A. The causal perspective on defining an evolution map

We begin with some reflections on the notion of an evolution map in a classical statistical theory. These are the ideas that underlie the calculus for causal reasoning that has been developed by Pearl [28].[3]

An evolution map is not a mere description of the *actual* statistical states of the input and the output of a process, but a prescription for determining the statistical state of the output for *any* statistical state of the input, in analogy to what distinguishes a law of motion from a historical account (as we already noted in Sec. II C a). Just as a law of motion is autonomous from the initial conditions, an evolution map is autonomous from the state at its input: For any variation of the state of its input, the map is unchanged.

To define the evolution map from A to B in an arbitrary circuit, therefore, one considers a counterfactual scenario involving the minimal modification to the circuit which allows one to freely vary the statistical state of A, while keeping the rest of the circuit unchanged. This minimal modification consists of altering the causal mechanism that determines A, while keeping every other causal mechanism in the circuit unchanged. (We have here made explicit use of the fact that the different causal mechanisms in the circuit are autonomous.) Because the way that B depends on A is only a function of these other causal mechanisms (and not a function of the mechanism that determines A), it is the same in the counterfactual circuit as it is in the actual circuit. Therefore, the evolution map from A to B in the actual circuit can be identified with the map from A to B in the counterfactual circuit. We denote the evolution map from A to B by $\Gamma_{B|\mathrm{do}A}(\square_A)$, in deference to the notion of a *do conditional* from classical causal modeling.

Recall the very first scenario we introduced, shown in Fig. 1, and consider its classical analog, shown in Fig. 7(a). For such a circuit, one computes the evolution map from S_1 to S_2 by applying the prescription just given, as follows. One imagines that the naturally occurring state of S_1 is

---

[3]Classical causal models are typically defined in terms of directed acyclic graphs, rather than circuit diagrams. We use the latter representation because it facilitates the comparison with the quantum case.

ignored, and instead a new state for S_1 is prepared, in a manner uncorrelated with the systems that would under natural circumstances be its causal parents. The counterfactual scenario being imagined is depicted in Fig. 7(b). The wire labeled S_1 in the original circuit is replaced, in the counterfactual circuit, by a pair of wires that are disconnected. The first, which inherits the label S_1, maintains the causal ancestry of the original, but is marginalized over (denoted by the ground symbol from electronics). The second, which is labeled by S_1', maintains the causal descendants of the original, and is an input to the counterfactual circuit.

The evolution map from S_1 to S_2 in the original circuit of Fig. 7(a), $\Gamma_{S_2|\mathrm{do}(S_1)} : \mathfrak{P}_{S_1} \to \mathfrak{P}_{S_2}$, is defined as the map which is isomorphic (under the identification of S_1 and S_1') to the map $\Gamma_{S_2|S_1'} : \mathfrak{P}_{S_1'} \to \mathfrak{P}_{S_2}$ in the counterfactual circuit of Fig. 7(b), that is,

$$\Gamma_{S_2|\mathrm{do}S_1}\left(\square_{S_1}\right) := \Gamma_{S_2|S_1'}\left(\square_{S_1'}\right). \tag{54}$$

Given that

$$\Gamma_{S_2|S_1'}\left(\square_{S_1'}\right) = \sum_{E'} F_{S_2E'|S_1'E}\left(\square_{S_1'} \otimes P_E\right), \tag{55}$$

where $P_E := \sum_{S_1} P_{S_1E}$, it follows that the evolution map for the scenario of Fig. 7(a) is

$$\Gamma_{S_2|\mathrm{do}S_1}\left(\square_{S_1}\right) := \sum_{E'} F_{S_2E'|S_1E}\left(\square_{S_1} \otimes P_E\right). \tag{56}$$

The classical causal models framework typically focuses on the do conditional $P_{S_2|\mathrm{do}S_1}$, rather than the evolution map $\Gamma_{S_2|\mathrm{do}S_1}$, but the latter is simply related to the former via

$$\Gamma_{S_2|\mathrm{do}S_1}\left(\square_{S_1}\right) := \sum_{S_1} P_{S_2|\mathrm{do}S_1}\square_{S_1}. \tag{57}$$

Equation (56) implies that the evolution map can be computed directly from the identity of the circuit elements in the original circuit of Fig. (1). This is because it depends only on $F_{S_2E'|S_1E}$ and P_E (which is obtained from P_{S_1E} by marginalization).

It follows, in particular, that there is no need to consider an input-ouput relation.

This is in contrast to the standard proposal, which (because of its insistence on starting with the input-output relation) *cannot* define a map uniquely unless the scenario explicitly involves a *set* of initial states. This is why, as we noted in Sec. II, the standard proposal cannot define a map for the circuit of Fig. 1 and why it is forced to consider circuits of the form of Fig. 2, where there are variables J and/or K that allow one to introduce variation in the input state.

Note that we are here discussing how to *define* the evolution map when one has a complete description of the form of the circuit and the identity of each of its elements. The question of how one obtains such a description is not relevant to the definitional question, and it is important not to confuse the two questions. We will discuss the question of how to infer the quantum evolution map from experimental data in Sec. VII D.

B. Contrasting the evolution map with the inference map

In the causal modeling framework [28], the primary motivation for introducing do conditionals was to distinguish them from standard conditionals. We pause here to describe the distinction because it is critical to our analysis of the mistakes of the standard proposal.

Suppose that one considers the map defined by the standard conditional $P_{S_2|S_1}$, namely,

$$\Gamma_{S_2|S_1}\left(\square_{S_1}\right) := \sum_{S_1} P_{S_2|S_1}\square_{S_1}. \tag{58}$$

The correct way of interpreting the map defined in Eq. (58) is as an *inference* map. It answers the following question: Given a particular state of knowledge of S_1, what is the appropriate state of knowledge to assign to S_2? In particular, if one updates one's description of S_1 (based on passively observing it, for instance), this map specifies how one should update one's description of S_2. If S_1 and S_2 are related causally by both a cause-effect connection and a common cause of the two, then the inference map describes what S_1 can teach you about S_2 through *either* causal pathway. By contrast, the evolution map specifies *only* what S_1 can teach you about S_2 through the cause-effect pathway.

This distinction is often illustrated by Simpson's paradox. This is a scenario in which there is a positive correlation between a treatment variable and a recovery variable, even though the causal influence of the treatment is to *reduce* the probability of recovery (see Sec. 6 of Ref. [28]). The positive correlation is the result of there also being a common cause acting on the two variables. For instance, if men are more likely than women to seek the treatment and are also more likely than women to recover regardless of treatment, then gender acts as a common cause. In this circumstance, learning that an individual in the sample population got the treatment warrants assigning a higher likelihood to the proposition that that individual recovered, simply because learning that someone got the treatment is positive evidence for them being male, which in turn is positive evidence for them recovering.

To see this mathematically, let R denote the recovery variable and let T denote the treatment variable. If one naïvely computes $P_{R|T} = P_{RT}/P_R$, or equivalently, the map $\Gamma_{R|T}$ [via Eq. (58)], it is clear that this does *not* represent the cause-effect relation that exists between T and R and hence cannot be used to make assessments of the effectiveness of the treatment on recovery. Only the map $\Gamma_{R|\mathrm{do}T}$, computed from $P_{R|\mathrm{do}T}$ [via Eq. (57)], wherein the intervention on T rules out the possibility of inference via a common cause, represents state updating based purely on the cause-effect relation.

In short, Simpson's paradox reminds us that correlation does not imply causation. To draw conclusions about whether a given treatment contributes causally to recovery, one must consider what would occur in a trial wherein the value of the treatment variable is assigned at random (drug or placebo, for instance), independently of any preferences of the individual. The do conditional describes what would occur in such randomized trials.

Hence, we see that the evolution map is generally distinct from the inference map.

Finally, note that—just as we saw for the evolution map—the inference map in classical scenarios can be computed directly from the original circuit elements, so that one need not consider any sort of input-output relation. To see this, first note that the conditional $P_{S_2|S_1}$ can be expressed as

$$P_{S_2|S_1} = \sum_E P_{S_2|S_1 E} P_{E|S_1} \tag{59}$$

$$= \sum_{E'E} P_{S_2 E'|S_1 E} P_{E|S_1}. \tag{60}$$

The assumption that the circuit elements in Fig. 7(a) are known implies that $F_{S_2 E'|S_1 E}$ and $P_{S_1 E}$ are known. $P_{S_2 E'|S_1 E}$ in the above equation is just the conditional associated to the stochastic map $F_{S_2 E'|S_1 E}$, while $P_{E|S_1}$ can be computed from $P_{S_1 E}$ by $P_{E|S_1} = P_{S_1 E}/P_{S_1}$, where $P_{S_1} = \sum_E P_{S_1 E}$. The inference map $\Gamma_{S_2|S_1}$ associated to the conditional $P_{S_2|S_1}$ is then found, via Eq. (58), to be

$$\Gamma_{S_2|S_1}(\square_{S_1}) := \sum_{S_1 EE'} P_{S_2 E'|S_1 E} P_{E|S_1} \square_{S_1} = \sum_{E'} F_{S_2 E'|S_1 E} \left(P_{E|S_1} \square_{S_1} \right). \tag{61}$$

V. THE MISTAKEN ASSUMPTIONS OF THE STANDARD PROPOSAL IN THE CLASSICAL SPHERE

With the causal point of view in mind and the correct definition of the evolution map in hand, it is very instructive to revisit the standard proposal in the classical sphere and to isolate the mistaken assumptions therein.

A. The mistake of confusing the evolution map with the inference map

Consider what the standard proposal entails for the simple circuit of Fig. 7(a). There is only a single distribution P_{S_1} and a single distribution P_{S_2}, and so the input-output relation consists of a *single* input-output pair, namely, $\{(P_{S_1}, P_{S_2})\}$. The standard proposal is committed to the notion that this input-output pair can be interpreted as a constraint on the evolution map, namely, that the evolution map acting on P_{S_1} must yield P_{S_2}.

Classical probability theory dictates that the relationship which holds between P_{S_1} and P_{S_2} is

$$P_{S_2} = \sum_{S_1} P_{S_2|S_1} P_{S_1}, \tag{62}$$

and consequently, if $\Gamma_{S_2|S_1}$ is the stochastic map associated to the conditional $P_{S_2|S_1}$ by Eq. (58), then

$$P_{S_2} = \Gamma_{S_2|S_1}(P_{S_1}). \tag{63}$$

But $\Gamma_{S_2|S_1}$ is the *inference* map from S_1 to S_2, which can differ from the evolution map $\Gamma_{S_2|\text{do}S_1}$, as discussed in Sec. IV B. Consequently, the input-output pair $\{(P_{S_1}, P_{S_2})\}$ is only guaranteed to be consistent with the inference map and is *not* guaranteed to be consistent with the evolution map. Indeed, if there is a common cause acting on S_1 and S_2, as is the case in the circuit of Fig. 7(a), then generically the evolution map and the inference map *do* differ, and the input-output pair $\{(P_{S_1}, P_{S_2})\}$ is *only* consistent with the inference map and not the evolution map.

To assume, as the standard proposal does, that the input-output pair $\{(P_{S_1}, P_{S_2})\}$ constitutes a constraint on the evolution map is to make a mistake akin to inferring causation from correlation alone. For instance, it is akin to inferring, from a positive correlation between treatment and recovery, that treatment has a positive causal influence on recovery even though there is a common cause (such as gender) that could account for this positive correlation.

This is the first mistake of the standard proposal.

As we discuss in Appendix C, the inference map and the evolution map coincide if and only if E and S_1 are marginally independent, $P_{S_1 E} = P_{S_1} \otimes P_E$. This is ironic, because proponents of the standard proposal were interested in characterizing precisely those scenarios which violated this condition.

B. The mistake of taking input-output pairs from different inference maps as constraints on a single map

Proponents of the standard proposal sought to define a map through an input-output relation $\mathfrak{R}$ where variability in the input state of the system S_1 (and therefore also the output state of S_2) was introduced via variation in the values of J and K in a circuit of the form of Fig. 6(b).

Since, as noted in the previous section, an individual input-output pair in $\mathfrak{R}$ is not a constraint on the evolution map, the *set* of such pairs obviously does not constrain the evolution map either. This would seem to leave open the possibility that the set of input-output pairs might still constrain the inference map. However, this is not the case either. If confusing inference with evolution had been the *only* mistake of the standard proposal, then because every inference map is a stochastic map, one would never have found any failure of the input-output relation to define a stochastic map, contrary to what is found in the pathological examples.

The standard proposal not only fails to yield the evolution map but also fails to yield the inference map, due to a second mistaken assumption, which we now discuss.

As we argued in Sec. IV B, the inference map in classical scenarios can be computed directly from the circuit, without introducing a set of initial states indexed by J and K. Therefore, introducing explicit variation in the initial state was unnecessary. Moreover, it is the root of all the pathological implications of the standard proposal.

The means by which variability was introduced in the input state *inadvertently* led to variability in the inference map as well. That is, for generic examples of circuits to which the standard proposal has been applied [those of the form of Fig. 6(b)], the variation in the values of J and K leads *not only* to variation in the marginal state of the system S_1 but also to variation in $P_{E|S_1}$, and hence, given Eq. (61), to variation in the inference map.

To show this, we determine the formula that relates the conditional probability $P_{S_1|JK}$ to $P_{S_2|JK}$ for circuits of the form of Fig. 6(b).

By definition,

$$P_{S_2|JK} = \sum_{S_1 E} P_{S_2 S_1 E|JK}. \tag{64}$$

To express $P_{S_2|JK}$ as a function of $P_{S_1|JK}$, we simply make a repeated application of an identity from Bayesian probability

theory, namely, $P_{AB|C} = P_{A|BC} P_{B|C}$. We infer that $P_{S_2 S_1 E|JK} = P_{S_2|S_1 EJK} P_{E|S_1 JK} P_{S_1|JK}$ and consequently that

$$P_{S_2|JK} = \sum_{S_1 E} P_{S_2|S_1 EJK} P_{E|S_1 JK} P_{S_1|JK}. \tag{65}$$

For a circuit of the form of Fig. 6(b), the causal structure ensures that S_2 is conditionally independent of JK given $S_1 E$, so that $P_{S_2|S_1 EJK} = P_{S_2|S_1 E}$, and we conclude that

$$P_{S_2|JK} = \sum_{S_1 E} P_{S_2|S_1 E} P_{E|S_1 JK} P_{S_1|JK}. \tag{66}$$

This defines a JK-dependent map from distributions on S_1 to distributions on S_2. Specifically, if we define $P_{E|S_2}^{(j,k)}$ by

$$\forall e, s, j, k : P_{E|S_2}^{(j,k)}(e|s) := P_{E|S_2 JK}(e|sjk), \tag{67}$$

the map is

$$\Gamma_{S_2|S_1}^{(j,k)} \big(\square_{S_1} \big) = \sum_{S_1 E} P_{S_2|S_1 E} P_{E|S_1}^{(j,k)} \square_{S_1}. \tag{68}$$

Rewriting in terms of the circuit elements given in our scenario [as was done in Eq. (61)], this map has the form

$$\Gamma_{S_2|S_1}^{(j,k)} \big(\square_{S_1} \big) = \sum_{E'} F_{S_2 E'|S_1 E} \big(P_{E|S_1}^{(j,k)} \square_{S_1} \big). \tag{69}$$

For one who knows that $J = j$ and $K = k$, the latter map is the correct way of making inferences from S_1 to S_2. The inference map clearly depends on the values of J and K.

It follows from Eqs. (66) and (68) that if $P_{S_1}^{(j,k)}$ and $P_{S_2}^{(j,k)}$ are defined by

$$\forall s, j, k : P_{S_1}^{(j,k)}(s) := P_{S_1|JK}(s|jk), \tag{70}$$

$$\forall s, j, k : P_{S_2}^{(j,k)}(s) := P_{S_2|JK}(s|jk), \tag{71}$$

then the input-output pair $\{(P_{S_1}^{(j,k)}, P_{S_2}^{(j,k)})\}$ is consistent with the map $\Gamma_{S_2|S_1}^{(j,k)}$ in the sense that

$$P_{S_2}^{(j,k)} = \Gamma_{S_2|S_1}^{(j,k)} \big(P_{S_1}^{(j,k)} \big). \tag{72}$$

However, the prescription of the standard proposal was to find a *single* map $\Gamma_{S_2|S_1}$ such that

$$\forall j, k : P_{S_2}^{(j,k)} = \Gamma_{S_2|S_1} \big(P_{S_1}^{(j,k)} \big). \tag{73}$$

This last step is the origin of the pathologies of the standard proposal. As Eq. (72) shows, in general the input-output pairs for different values of JK describe the input and output states of *different* maps. But the prescription of the standard proposal asks us to collect all of these input-output pairs into a single set, the input-output relation, and to try and find a *single* map that is consistent with all of them. Given the origin of these pairs, there is no guarantee that there is *any* such map, and even if there does happen to be one, there is no guarantee that it is linear or stochastic.

To summarize, the second mistake of the standard proposal is to have inadvertently introduced variability in the inference map and then to have tried to define a *unique* map from the input-ouput pairs that are associated to these different inference maps.

C. The correct evolution map(s) for the scenarios wherein the standard proposal led to pathologies

In the previous section, we noted that the inference map can be JK dependent in the scenario of Fig. 6(b). It turns out that the evolution map can also be JK dependent in this scenario. Generally, therefore, *there is no single evolution map* to be characterized in this scenario. Rather, following the prescription of Sec. IV A, one finds that for each valuation (j, k) of J and K, there is a (generally) distinct marginal state $P_E^{(j,k)}$ of the environment, which leads to distinct evolution maps, namely,

$$\Gamma_{S_2|\text{do}S_1}^{(j,k)} \big(\square_{S_1} \big) = \sum_{E'} F_{S_2 E'|S_1 E} \big(\square_{S_1} \otimes P_E^{(j,k)} \big). \tag{74}$$

At first glance, this may seem problematic, but in fact the knowledge dependence of evolution maps is ubiquitous in both classical and quantum physics. We give several examples in Appendix D.

D. Illustrating the mistakes with the classical examples considered previously

In this section, we illustrate the general discussion just given by explicitly analyzing classical examples introduced previously, highlighting interesting features along the way.

Consider again the third classical example, discussed in Sec. III E, where the input-output relation failed to identify any map at all.

Recall that the example has no K variable, so only J is relevant. Consider the inference map from S_1 to S_2 for a particular value j of J. Specializing Eq. (69), this has the form

$$\Gamma_{S_2|S_1}^{(j)} \big(\square_{S_1} \big) = \sum_{E'} F_{S_2 E'|S_1 E} \big(P_{E|S_1}^{(j)} \square_{S_1} \big). \tag{75}$$

The coupling of system and environment implies that $S_2 = S_1 \oplus_2 E$; that is, it implies that S_2 tracks the *parity* of S_1 and E. The controlled-NOT operation from J to S_1 toggles this parity in a J-dependent way, but without changing the marginal state of S_1. Thus, as one varies J, one has variability in the state on S_2, but no variability in the state on S_1, leading to the one-to-many relation that fails to correspond to any map.

One way to understand the fact that different values of J lead to different states on S_2 is that E is correlated with J given S_1, so that $\exists e, s : P_{E|S_1 J}(e|s, 0) \neq P_{E|S_1 J}(e|s, 1)$, or equivalently [given Eq. (70)], $P_{E|S_1}^{(0)} \neq P_{E|S_1}^{(1)}$, and consequently the inference map in Eq. (75) becomes J dependent.

Indeed, as we show in Appendix E, the inference maps for the two values of J are

$$\Gamma_{S_2|S_2}^{(j)} \big(\square_{S_1} \big) = \delta_{S_2, j}, \tag{76}$$

so that for $J = 0$ ($J = 1$), one updates one's description of S_2 to the state $[0]_{S_2}$ ($[1]_{S_2}$) *regardless* of one's state of knowledge of S_1.

Recall the input-output relation for this example, Eq. (53).

Clearly, the first ($J = 0$) input-output pair is consistent with the inference map for $J = 0$, $\Gamma_{S_2|S_1}^{(0)}$, while the second ($J = 1$) input-output pair is consistent with the inference map for $J = 1$, $\Gamma_{S_2|S_2}^{(1)}$. However, if one mistakenly considers *both* input-output pairs to be associated to a *single* map, one finds

a contradiction. This illustrates the second mistake of the standard proposal.

The first mistake of the standard proposal is also illustrated in this example.

The correct evolution map is straightforward to identify using the prescription of Sec. IV A. For both values of J, the marginal state of the environment is uniformly random, so the evolution map is J independent and equal to the randomizing map

$$\Gamma_{S_2|\text{do}S_1}\big(\Box_{S_1}\big) = \tfrac{1}{2}\big([0]_{S_2} + [1]_{S_2}\big), \qquad (77)$$

which takes any input probability distribution on S_1 to a uniformly random distribution on S_2. This is distinct from either of the inference maps $\Gamma^{(0)}_{S_2|S_1}$ or $\Gamma^{(1)}_{S_2|S_1}$ and so confirms that the evolution map is not constrained *at all* by the input-output pairs considered in the standard proposal.

We now turn our attention to the second classical example, discussed in Sec. III E.

Again, only the J variable is nontrivial in this example. However, whereas J was a cause of S_1 alone in the third classical example, here it is a cause of E as well. In fact, J is the complete common cause of S_1 and E and consequently E is conditionally independent of S_1 given J, $P_{E|S_1J} = P_{E|J}$. Thus, Eq. (69) specializes to the expression

$$\Gamma^{(j)}_{S_2|S_1}\big(\Box_{S_1}\big) = \sum_{E'} F_{S_2E'|S_1E}\big(P^{(j)}_E \otimes \Box_{S_1}\big) \qquad (78)$$

for the inference map from S_1 to S_2 in this example. Different values of J lead to different marginals on E and therefore different inference maps.

The inference map for $J = j$ is

$$\Gamma^{(j)}_{S_2|S_1}\big(\Box_{S_1}\big) = \delta_{S_2,(S_1\oplus_3 j)}, \qquad (79)$$

as shown in Appendix E. For any state of knowledge of S_1, one should assign the same state of knowledge to S_2 if $J = 0$, and the same state of knowledge modulo an increase of the value by 1 (in arithmetic modulo 3) if $J = 1$.

Again, one can check explicitly that in this example an individual input-output pair for $J = j$ [as in Eqs. (44) and (46)] is a constraint on the inference map for $J = j$. However, if one mistakenly considers both input-output pairs to be associated to a single map, then the only maps consistent with the constraint are nonlinear.

In this example, it happens that the evolution map for a particular value j of J coincides with the inference map for that same value,

$$\Gamma^{(j)}_{S_2|\text{do}S_1} = \Gamma^{(j)}_{S_2|S_1}. \qquad (80)$$

The reason is that conditioning on J makes S_1 and E independent, so that in the presence of this conditioning, one need not intervene on the system to achieve this independence.

The analysis of the first classical example is similar and is also provided in Appendix E.

The situation in the classical sphere can be summarized as follows. The only sensible notion of an evolution map classically is the one defined by a do conditional, but the information used in the standard proposal to try to infer the system's evolution—input-output pairs of statistical states—does not constrain the evolution map but instead only constrains the
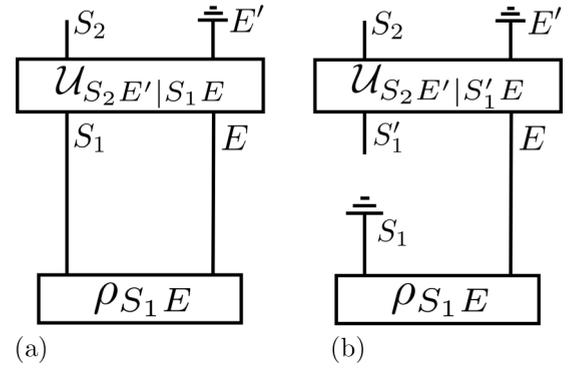
FIG. 8. (a) Repeat of Fig. 1, for ease of reference. (b) The hypothetical circuit which aids in defining the evolution map for the situation in panel (a).

inference map. Furthermore, *even if* one were content to try to identify the inference map rather than the evolution map, the prescription of the standard proposal does not provide a means of doing so because it takes a number of input-output pairs for *different* inference maps and mistakenly takes them all to be constraints on a single map.

VI. THE CORRECT DEFINITION OF THE EVOLUTION MAP IN THE QUANTUM SPHERE

The definition of a classical evolution map, presented in Sec. IV, generalizes naturally to the definition of a quantum evolution map. It suffices to substitute the quantum analogs of the relevant classical notions (quantum states for statistical states and unitary dependences for functional dependences) in the definition provided there.

The ideas underlying the definition are the same as those outlined in Sec. IV A: An evolution map for quantum states of a system is a prescription for determining the quantum state of the output for *any* quantum state of the input, and it is autonomous from the state at its input.

Just as in the classical case, one imagines the counterfactual scenario with the minimal modification to the circuit which allows one to freely vary the quantum state of A, while keeping the rest of the circuit unchanged; the evolution map from A to B in the actual circuit can then be identified with the map from A to B in the modified circuit. We denote an evolution map from quantum system A to quantum system B by $\mathcal{E}_{B|\text{do}A}(\Box_A)$, where the "do$A$" on the right of the conditional parallels the classical notation and is a reminder that the definition requires contemplating the counterfactual scenario just described.

Let us return to the most basic quantum circuit wherein the system and environment are initially correlated, that of Fig. 1, reproduced here as Fig. 8(a). For this case, the relevant counterfactual scenario is depicted in Fig. 8(b). We again denote the version of the system which is varied counterfactually by S'_1, and here the ground symbol from electronics represents the trace over a subsystem.

The evolution map from S_1 to S_2 in the circuit of Fig. 8(a), denoted $\mathcal{E}_{S_2|\text{do}S_1} : \mathcal{L}(\mathcal{H}_{S'_1}) \to \mathcal{L}(\mathcal{H}_{S_2})$, is defined as the map which is isomorphic (under the identification of S_1 and S'_1) to the map $\mathcal{E}_{S_2|S'_1} : \mathcal{L}(\mathcal{H}_{S'_1}) \to \mathcal{L}(\mathcal{H}_{S_2})$ in the counterfactual

circuit of Fig. 8(b); that is,

$$\mathcal{E}_{S_2|\mathrm{do}(S_1)} := \mathcal{E}_{S_2|S_1'}. \tag{81}$$

This prescription unambiguously leads one to identify the evolution map to be

$$\mathcal{E}_{S_2|\mathrm{do}(S_1)}(\square_{S_1}) = \mathrm{Tr}_{E'}\big(\mathcal{U}_{S_2E'|S_1E}(\square_{S_1} \otimes \rho_E)\big), \tag{82}$$

where $\rho_E := \mathrm{Tr}_{S_1}(\rho_{S_1E})$. This is the quantum analog of Eq. (56). It is always completely positive.

The quantum evolution map can be deduced from the identity of the circuit elements in the original circuit of Fig. 8(a) because it depends only on $\mathcal{U}_{S_2E'|S_1E}$ and on ρ_E (which is obtained from ρ_{S_1E} by taking a partial trace). No input-output relation is needed.

Furthermore, like its classical counterpart, the quantum evolution map is knowledge dependent insofar as it depends on one's information about the environment, encoded in the quantum state ρ_E.[4] As we discuss in Appendix D, this knowledge dependence of one's description of evolution is ubiquitous in physics, including textbook quantum mechanics.

For completeness, in Appendix F we determine the correct evolution map for each of the three quantum examples from Sec. II. We also show that the most general circuit considered in the standard proposal, Fig. 2, is described by a distinct evolution map for each valuation (j, k) of J and K:

$$\mathcal{E}_{S_2|\mathrm{do}(S_1)}^{(j,k)}(\square_{S_1}) = \mathrm{Tr}_{E'}\big(\mathcal{U}_{S_2E'|S_1E}(\square_{S_1} \otimes \rho_E^{(j,k)})\big). \tag{83}$$

A. What is an inference map quantumly?

Once one takes a causal point of view, the description of quantum evolution is scarcely more complicated than the description of classical evolution. Devising a complete framework for describing *inference* in quantum theory, however, is a much more complicated venture, and remains an open problem [39,43,55]. (Note that, in keeping with the framework laid out in Sec. IV, both evolution and inference are formalized as maps from states of S_1 to states of S_2, and it is this type of object that we are interested in here. We discuss maps with a more general input type in Sec. VII B).

The root of the problem lies in how one acquires the information about S_1 that is input into the map. Classically, passive observation of a variable does not change the dependence of that variable on its causal parents. So, there is a way to update one's knowledge of S_1, and therefore the distribution P_{S_1}, without changing any of the causal mechanisms that relate S_1 to the other variables of interest. In the quantum realm, however, it is unclear whether there is an appropriate analog of passive observation given that every attempt to gain information about a system changes its state. Consequently,

---

[4] Some previous works (e.g., Refs. [5,11]) sought to define the evolution map as a function of the marginal state of the environment *and also* as a function of the initial system-environment correlations present in the initial state of the composite. This prescription is distinct from that of the standard proposal, but it does not correspond to our prescription. In particular, it does not reproduce the correct evolution map [of Eq. (82)] because the latter depends *only* on the marginal state of the environment.

it would seem to be impossible to update one's knowledge of S_1, and therefore the state ρ_{S_1}, without changing any of the causal mechanisms that relate S_1 to the other systems of interest. (For example, measuring a quantum system generally leads to a different postmeasurement state, altering the way in which the system affects its causal children, while preparing the system in a known state nullifies the influence of its causal parents.)

It is worth noting that, even classically, it is only for a limited set of probing schemes on S_1 (which includes "passive" measurements, which are nondisturbing and maximally informative) that the question of what one can infer about S_2 (given the outcome of the probing scheme) can be answered by an inference map that takes statistical states on S_1 as input. Since this is mathematically a limiting case of the full quantum treatment, one can see already that for quantum systems, too, only a *limited set* of probing schemes on S_1 could possibly admit inferences about S_2 that can be expressed by a map with quantum states on S_1 as input. Moreover, since quantum states generally contain richer information about a system than their classical limit, the set of scenarios that admit such an inference map is likely to be even more restricted. Nonetheless, it is conceivable that there are *intrinsically quantum* scenarios—those that cannot be reduced to an effectively classical description—wherein such an inference map can be defined.

For *generic* probing schemes (both classical and quantum), there is no such map which takes a state of S_1 as input.

B. The mistakes of the standard proposal in the quantum sphere

The set of quantum states and evolution maps includes the set of classical states and evolution maps as special cases (wherein all operators are diagonal in some fixed basis). The fact that the standard proposal fails to identify the correct evolution map in various classical examples, therefore, implies that it is not a valid prescription for identifying the correct evolution map in the quantum sphere. Furthermore, the mistaken assumptions of the standard proposal that we have identified by considering classical examples remain mistaken assumptions in the quantum sphere.

Outside the classical subtheory, the story is more subtle, but analogs of the two mistaken assumptions can be identified. We begin with the quantum analog of the first mistake. Although we have left open the question of whether one can make sense of the notion of an inference map from S_1 to S_2 in such examples, it is still clear that, whenever one has initial system-environment correlations, the quantum systems S_1 and S_2 are causally related not only as cause and effect but by a common cause as well. As such, an individual input-output pair of states in such a scenario will generally not reflect the cause-effect relation alone. But given that the correct definition of the evolution map depends *only* on the cause-effect relation, there is no reason to think that an individual input-output pair of states is a constraint on the evolution map when there are initial system-environment correlations.

Indeed, as we show in Appendix G, the necessary and sufficient condition for an individual input-output pair to be a constraint on the evolution map is that the joint state

FIG. 9. Composition of system-environment interactions, leading to non-Markovian evolution on the system.

on S_1E factorizes, $\rho_{S_1E} = \rho_{S_1} \otimes \rho_E$. The ironic conclusion (reached also in our discussion of the classical case) is that the only circumstance in which the first mistake of the standard proposal would be innocuous is the case of *no* initial system-environment correlations.

If there is (as speculated in the previous section) a subset of intrinsically quantum scenarios where an inference map can be defined, and the scenario in question is presumed to be within this subset, then *perhaps* an input-output pair of states for a given value of JK could be taken as a constraint on the inference map. But it would still be the case that the inference map would vary with JK, and consequently it would be inappropriate to consider all of these input-output pairs as constraints on a single map. This is the second mistake of the standard proposal.

VII. DISCUSSION AND FUTURE WORK

We here summarize some of the lessons of our analysis. First, one cannot define an evolution map simply by listing a set of marginal states at its input and at its output, since if there is a common cause acting on the input and the output, then such pairs of states do not constitute constraints on the evolution map. Second, one should not restrict the input domain of the evolution map, as this would violate the counterfactual conception of evolution, embodied for instance in the idea that a law of motion should be autonomous from the initial conditions. Further, such restrictions are not needed to preserve complete positivity. Third, the marginal state of the environment is relevant to the evolution map, while the system-environment correlations are *not*. This latter fact is not assumed but rather is derived from our conception of evolution and holds true in all possible operational scenarios and for all possible initial states. Fourth, because the marginal state one assigns to the environment describes one's information about the environment, the evolution map one assigns will depend on one's state of knowledge. If a scenario includes a variable that contains information about the environment, then that scenario is associated to a *set* of evolution maps rather than a single evolution map, one for each value of the variable.

We hope that these lessons might be valuable outside the scope of this work.

A. Composition of evolution maps

The definition of the quantum evolution map, Eq. (82), has the following feature: For a circuit of the form of Fig. 9,

the composition of the evolution map from S_0 to S_1 with the evolution map from S_1 to S_2 does not generally yield the evolution map from S_0 to S_2; that is, the compositional property $\mathcal{E}_{S_2|\mathrm{do}S_0} = \mathcal{E}_{S_2|\mathrm{do}S_1} \circ \mathcal{E}_{S_1|\mathrm{do}S_0}$ fails to hold.

This compositional property only holds if S_1 constitutes the *complete* causal mediary between S_0 and S_2. For the circuit in Fig. 9, however, E_1 is also such a causal mediary, and so the only sequential decomposition of the evolution map from S_0 to S_2 that holds is the decomposition into an evolution map from S_0 to S_1E_1 and an evolution map from S_1E_1 to S_2, that is, $\mathcal{E}_{S_2|\mathrm{do}S_0} = \mathcal{E}_{S_2|\mathrm{do}S_1\mathrm{do}E_1} \circ \mathcal{E}_{S_1E_1|\mathrm{do}S_0}$.

It is only with the correct definition of the evolution map that one can properly pose the question of how it can be decomposed into a temporal sequence of evolution maps. Therefore, we expect that our results will have relevance to questions about the divisibility of quantum channels and the detection of nonMarkovianity [56,57].

B. Comparison with prior work

As mentioned in the introduction, some prior work has also criticized the standard argument for the inadequacy of completely positive maps [20,21]. These works have argued against the standard proposal on the grounds that it lacks operational meaning, for instance, by pointing out that one cannot vary the marginal state of the system while keeping the system-environment correlations fixed. In contrast, our criticism of the standard proposal is that (i) it fails to satisfy the criterion of universality, e.g., it only seeks to answer question Q′ (concerning Fig. 2) rather than question Q of Sec. II B (concerning Fig. 1), and even for question Q′ it sometimes fails to identify *any* map as the description of the evolution of the system, and (ii) in the classical limit, it fails to reproduce the evolution map that is implied by the framework of classical causal modeling.

This prior work, e.g., Refs. [20,21,24–27], has also proposed a framework for analyzing dynamics in the presence of system-environment correlations, where one focuses on an altogether different type of map that *does* have a clear operational meaning and which is always completely positive. Specifically, whereas maps purporting to describe the dynamics of system S typically take states on S_1 as input and have states on S_2 as output, these new maps take instruments from S_1 to S_1' as input and have states on S_2 as output. In other words, the domain of the map is the space of operators on *two* rather than one copy of the system's Hilbert space, that is, on $\mathcal{H}_{S_1} \otimes \mathcal{H}_{S_1'}$ rather than on $\mathcal{H}_{S_1}$.

One can motivate this perspective by considering the circuit in Fig. 10(a), which is clearly a special case of Fig. 2, and reconceptualizing the experiment not as a preparation of a set of initial joint states on $S_1'E$, but rather as an interventional probing scheme on the system. This idea is represented in Fig. 10(b), where the intervention on the system, highlighted by a dashed box, is associated with a map $\mathcal{E}_{S_1'|S_1}^{(j,k)} : \mathcal{L}(\mathcal{H}_{S_1}) \to \mathcal{L}(\mathcal{H}_{S_1'})$ for each value k of K and j of J, whereas the circuit fragment outside the dashed box is associated with a map $\mathcal{E}_{S_2S_1|S_1'} : \mathcal{L}(\mathcal{H}_{S_1'}) \to \mathcal{L}(\mathcal{H}_{S_1} \otimes \mathcal{H}_{S_2})$. The circuit fragment defines a map from each element of the instrument describing the intervention, $\mathcal{E}_{S_1'|S_1}^{(j,k)}$, to a state $\rho_{S_2}^{(j,k)}$ using the link product
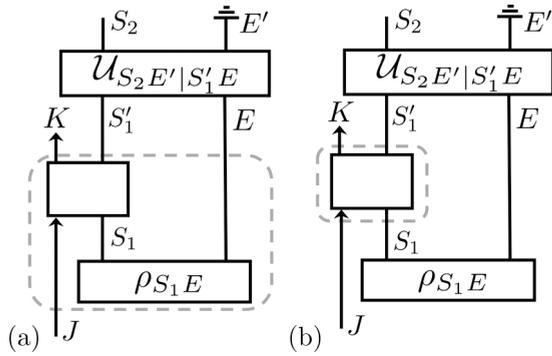
FIG. 10. (a) The subset of circuits considered in the standard proposal wherein the setting variable J determines the measurement on the system and one postselects on the measurement outcome K. (b) The same operational scenario as in panel (a), but reconceptualized as an intervention with setting J and outcome K on the system, rather than a preparation of a set of states labeled by J and K.

of Ref. [58], $\mathcal{E}_{S_2 S_1 | S_1'} \star \mathcal{E}_{S_1' | S_1}^{(j,k)} = \rho_{S_2}^{(j,k)}$, where the S_1 output of the circuit fragment is fed into the S_1 input of the instrument and the S_1' output of the instrument is fed into the S_1' input of the circuit fragment. This sort of map has been studied in many formalisms [20,44,58–64] and is often termed a *quantum comb* [58] or a *process matrix* [61]. When it was introduced for the study of open quantum systems in Ref. [20], this object was termed the M map.

The evolution map can be recovered from a specification of the M map (see the discussion surrounding Eq. (84) and in Sec. VII C). Nonetheless, the formalism of M maps differs from our own framework in two key respects.

First, the M map and our evolution map seem to have been proposed as answers to different questions: The M map was introduced to describe aspects of the system's dynamics in an experiment of the form of Fig. 10(b), where an intervention is actually made on the system, whereas our proposal is intended to describe the evolution of the system in the natural experiment of Fig. 1. In other words, whereas we aim to answer question Q from Sec. II B (concerning Fig. 1), the M map is proposed as an answer to a question of the form of Q′, but with the experiment given by Fig. 10(b) (rather than Fig. 2).

Second, we have defined the evolution map to be a complete description of the inferences one can make from S_1 to S_2 based purely on the cause-effect relation between the two, while the M map does not capture only those types of inferences but rather includes (as noted in, e.g., Ref. [20]) information about the initial correlations between the system and environment and hence describes one's inferences based on both the cause-effect and on the common-cause relation between S_1 and S_2. One consequence of this distinction is that our notion of the quantum evolution map reduces in the classical limit (where all operators are diagonal in some fixed basis) to the classical evolution map, which we identify with Pearl's do conditional, whereas the M map does not reduce to the do conditional in the classical limit. Indeed, from our perspective, the very fact that the M map takes an instrument from S_1 to S_1' as input rather than a state on S_1 precludes it from being a candidate for the evolution map.

Reference [65] also asks a question of type Q′; namely, it asks what map one should use to describe scenarios in which certain special types of interventions are performed on the system. One such set of interventions considered in Ref. [65] and called "stochastic preparations" consists of a trace operation on S_1 and a repreparation of S_1' in one of an informationally complete set of states. For such a set of interventions, the map that is tomographically reconstructed from data is

$$\mathcal{E}_{S_2 | S_1'} (\square_{S_1'}) = \mathrm{Tr}_{S_1} \circ \mathcal{E}_{S_2 S_1 | S_1'} (\square_{S_1'}). \tag{84}$$

Based on the notation used for this object in Ref. [20], we call it the L map. Unlike the M map, it is identical in form to the evolution map we have introduced in Eq. (82). However, just as we noted in our discussion of the M map, we read this prior work as asserting that the L map describes the dynamics only for a very limited scope of experiments, namely those wherein a "stochastic preparation" was actually implemented, and not that it does so for the general case involving no intervention, depicted in Fig. 1.

This prior work appears to be motivated by a type of operationalism wherein an entity is only meaningful insofar as one can specify the experimental procedure that would allow one to measure it. In particular, it defines both the M map and the L map in terms of the tomographic procedure by which one could identify it. By contrast, the causal modeling perspective that we adopt in this work is not wedded to this positivist notion of meaningfulness and therefore aims to answer question Q directly, without recourse to interventions but rather through the use of counterfactuals.

Note, however, that if one adopts the positivist attitude merely as a means for discovering the form of the L map, but one then kicks away this empiricist ladder and simply treats the L map as the correct description of dynamics even in the case of Fig. 1 (where there are no interventions on which any operational definitions could be based), then one comes to the same conclusion regarding the correct definition of the evolution map as we have reached here by adopting the causal modeling perspective.

Similarly, one might appeal to such a ladder-kicking procedure to justify considering the M map as pertinent not just to the class of experiments shown in Fig. 10(b) but also to the general case of Fig. 1. Even reconceptualized in this way, however, the M map is not an evolution map in our sense, for the reasons described above. So what does the map $\mathcal{E}_{S_2 S_1 | S_1'}$ describe from the causal modeling perspective? We address this question in the next section.

C. The status of the quantum comb from the causal point of view

Viewed as a description of the general case of Fig. 1, Eq. (84) shows how the map $\mathcal{E}_{S_2 S_1 | S_1'}$ can be used to recover the evolution map, which describes inferences that can be made solely on the basis of the cause-effect relation between S_1 and S_2. Additionally, $\mathcal{E}_{S_2 S_1 | S_1'}$ can be used to recover an inference map that describes inferences that can be made solely on the basis of the common cause relation between S_1 and S_2. In order to be assured that one's inferences about S_2 are based purely on knowledge of S_1, and not on knowledge of S_1',

one presumes a state of complete ignorance regarding S_1', represented by the completely mixed state $\rho_{S_1'} = \frac{1}{d_{S_1'}}\mathbb{1}_{S_1'}$, where $d_{S_1'}$ is the dimension of $\mathcal{H}_{S_1'}$. The object that allows one to make inferences based solely on the common cause relation, therefore, is the joint state

$$\rho_{S_1 S_2} := \mathcal{E}_{S_2 S_1 | S_1'}\left(\frac{1}{d_{S_1'}}\mathbb{1}_{S_1'}\right). \tag{85}$$

In particular, if one implements a measurement on S_1 and one obtains an outcome $X = x$ associated to a positive operator $E_{S_1}^{(x)}$, then one should update one's description of S_2 to

$$\rho_{S_2}^{(x)} = \mathcal{N}\mathrm{Tr}_{S_1}\big(\rho_{S_1 S_2}\big(E_{S_1}^{(x)} \otimes \mathbb{1}_{S_2}\big)\big), \tag{86}$$

where $\mathcal{N}$ is a normalization constant. This defines a map from effects on S_1 to states on S_2 which has been termed the *steering map* and discussed in Ref. [30]. It is linear and co-CP (that is, the composition of the map with the tranpose operation is completely positive).

Based on these facts, one might hope that the map $\mathcal{E}_{S_2 S_1 | S_1'}$ simply represents the elusive quantum inference map that we discussed in Sec. VI A. However, it does not.

Here, one must again distinguish between two sorts of maps that might describe such inferences—one whose inputs are operators on a single copy of S_1 and one whose inputs are operators on two copies of S_1. It is the possibility of the former that we discussed in Sec. VI A, whereas the map $\mathcal{E}_{S_2 S_1 | S_1'}$, if it is to be interpreted as an inference map, is of the latter variety. Even if it turns out that the only sort of inference map which can sensibly be defined in generic quantum scenarios has as input a state on $\mathcal{H}_{S_1} \otimes \mathcal{H}_{S_1'}$,[5] the natural inference map would still be of the form $\mathcal{E}_{S_2 | S_1 S_1'}$ rather than $\mathcal{E}_{S_2 S_1 | S_1'}$. Defining such a map likely requires a theory of quantum Bayesian inversion (e.g., to determine the correct input state on $\mathcal{H}_{S_1} \otimes \mathcal{H}_{S_1'}$ from a specification of an instrument from S_1 to S_1').

D. Experimentally determining quantum combs and evolution maps

Our focus herein has been on elucidating the correct *definition* of the evolution map when the circuit is completely specified. A secondary problem concerns deducing the evolution map from experimental data. (It is secondary insofar as it can only be addressed once one has the correct definition of the evolution map.)

It is clear that if one can experimentally characterize the different components of the circuit (the state $\rho_{S_1 E}$ and the unitary map $\mathcal{U}_{S_2 E' | S_1 E}$), then one can deduce the evolution map from Eq. (82). Although this is sufficient, it is not necessary for characterizing the evolution map. The question arises, therefore, of what is the *minimal* experimental effort that suffices.

_____

[5]Which is a reasonable conjecture: Even in the classical sphere, disturbing interventions are most naturally described by an inference map from states of knowledge on *two* copies of random variable S_1 to S_2.

To begin with, consider the idealization wherein one can perfectly implement any desired laboratory operation. The causal map $\mathcal{E}_{S_2 S_1 | S_1'}$ has input space $\mathcal{L}(\mathcal{H}_{S_1'})$ and output space $\mathcal{L}(\mathcal{H}_{S_1} \otimes \mathcal{H}_{S_2})$ Hence, if one intervenes on the system by performing an informationally complete measurement on S_1 and then repreparing one of an informationally complete set of states on S_1', and one also implements (for each state on S_1') an informationally complete measurement on S_2, then one can tomographically reconstruct the causal map. This was termed *causal tomography* in Ref. [30].

If one has experimentally determined the causal map, then the evolution map can be computed directly from it via Eq. (84). A less demanding experimental method, however, is to simply ignore S_1 (rather than measuring it) and prepare S_1' in one of an informationally complete set of states. For each such state, one then implements an informationally complete measurement on S_2. This achieves process tomography on the evolution map. Note that this corresponds to experimentally implementing the intervention that is contemplated in the counterfactual circuit that defines the evolution map. It is analogous to measuring a classical do conditional by implementing a randomized trial.

What about experimentally determining the causal map or evolution map when the experimenter does *not* have ideal laboratory operations? For example, what if they are not able to localize their operations to particular systems, or to characterize their laboratory operations accurately? Because such limitations can result in inadvertently preparing initial system-environment correlations, they have previously served as a motivation for this field. Specifically, it has been suggested that if no completely positive map fits the relation of the standard proposal, this should be taken as evidence that one has inadvertently introduced system-environment correlations [22,66].

Given our demonstration that the prescription of the standard proposal does not yield the evolution map, we advocate against analyzing experimental data in the manner it proposes, even if only as a diagnostic for imperfections in one's laboratory operations. The solution to the problem of imperfect laboratory operations is to use a form of tomography in which various features of the probing scheme (such as the identities of the laboratory operations and the dimensionality and nature of the systems being probed) are not *presupposed* but are rather taken to be hypotheses whose plausibility is assessed on the basis of the data. In identifying the evolution map that provides the best fit to the experimental data, the maps over which one varies must be *constrained* to be completely positive (by virtue of the definition of an evolution map), and if the quality of the fit is poor (as is indicated, for instance, by a bad p value), then the correct reaction is not to entertain the possibility that completely positive maps are inadequate for describing evolution, as has previously been suggested [22,66], but to reject one or more of the assumptions about features of the probing scheme. A demonstration of how to implement tomography in this way is provided in Ref. [67].

Finally, we note that in classical scenarios, it is sometimes possible to identify the evolution map from purely passive observations if these are made on the right systems. (The more general problem of identifying the do conditional for some pair of variables embedded in a given causal structure is

known as the "identifiability problem" [28].) The significance of this fact is that although evolution maps are defined in terms of hypothetical nonpassive interventions, these nonpassive interventions need not necessarily be performed in order to identify the map. Whether there is a sensible quantum analog of the identifiability problem and whether it is possible to sometimes deduce the quantum evolution map without non-passive interventions remains unclear. In particular, to even pose the question, one must explore whether there is anything in quantum theory that ought to be considered an analog of passive observation.

VIII. CONCLUSION

We have argued that, for a general circuit wherein a system interacts with its environment such as the one in Fig. 1, the evolution map on the system should be defined as in Eq. (82). This map is always linear and completely positive, *even* in the presence of arbitrary initial system-environment correlations. Thus, we have shown that the common wisdom that such initial correlations constitute an exception to the rule of complete positivity is mistaken.

Our results provide good reasons to abandon a host of questions that have previously been the primary focus of research in this field. There is no reason to find necessary or sufficient conditions under which one can find a completely positive map (or a linear map, a positive map, etc.) that is consistent with the input-output relation of the standard proposal, because this relation has no bearing on the evolution map. There is also no reason to worry about the physical meaning of evolution maps that are not completely positive or not linear, because these never arise.

Instead, we advocate for pursuing a new set of questions, inspired by the drive to generalize the do calculus of Ref. [28] to the quantum realm [41,68]. Are there circumstances in which one can define an intrinsically quantum inference map? Can one sometimes identify the evolution map even without implementing the hypothetical intervention invoked in its definition? What are the implications of our results for continuous-time dynamics of open quantum systems, and in particular, non-Markovian dynamics?

APPENDIX A: GENERIC FAMILIES OF QUANTUM EXAMPLES IN WHICH THE STANDARD PROPOSAL FAILS

We here illustrate the breadth of the pathologies of the standard proposal and show that there is nothing special about the examples we have given.

In Sec. II F, we argued that there exist operationally realizable relations $\mathfrak{R} = \{(\sigma_j, \tau_j)\}_j$ for which the following is true:

(1) There is no completely positive map $\mathcal{E}$ such that $\forall j : \mathcal{E}(\sigma_j) = \tau_j$.

(2) There is no linear map $\mathcal{E}$ such that $\forall j : \mathcal{E}(\sigma_j) = \tau_j$.

(3) There is no map $\mathcal{E}$ such that $\forall j : \mathcal{E}(\sigma_j) = \tau_j$.

For notational simplicity, we have dropped the system labels on the quantum states in the relation; throughout, note that $\sigma_j \in \mathcal{L}(\mathcal{H}_{S_1})$ and $\tau_j \in \mathcal{L}(\mathcal{H}_{S_2})$.

We further asserted that all three of these failures can arise in each of the three operational circuit types described in Sec. II F; namely, circuits where the initial state of the system is varied by

(a) the choice of transformation on the system-environment composite,

(b) the choice of measurement on the system (and the choice of postselection on its outcome), and

(c) the choice of transformation on the system alone.

All of these are, of course, special cases of the general circuit in Fig. 2. We have highlighted these three specific circuit types because they cover the range of operational scenarios considered in the literature. Roughly, scenarios of type (a) are considered in Refs. [20,23–26]; of type (b) in Refs. [3,5,14,16,23], and of type (c) in Refs. [20,23–26]. Furthermore, *all* examples in the literature, to our knowledge, fit within one of these three categories.

We now provide these nine families of examples (one for each pairing of 1,2,3 and a,b,c), in order to cover the full diversity of examples in the literature and to demonstrate explicitly that one cannot salvage the standard proposal simply by restricting to one of the families of operational circuits (a,b,c).

For simplicity of presentation, we represent the system-environment interaction in these examples by a general quantum channel $\mathcal{F}_{S_2|S_1E}$ rather than a unitary channel $\mathcal{U}_{S_2E'|S_1E}$. (Of course, any such quantum channel could be dilated to recover a unitary description of the same example, although this dilation might require an increase in the dimensionality of the environment to accommodate the ancilla required to achieve the dilation.)

Examples of types 1,2,3 will be given, respectively, by providing an operationally realizable relation such that the following is true:

(1) The set $\{\sigma_j\}_j$ is informationally complete, and $\forall j : \tau_j := T(\sigma_j)$, where T denotes the transpose map (relative to some basis), so that $\mathcal{E}$ is this transpose map and consequently is not completely positive.

(2) The $\{\tau_j\}_j$ are more distinguishable[6] than the $\{\sigma_j\}_j$, so that $\mathcal{E}$ is necessarily a nonlinear map.

---

[6]That is, if there exists a pair (τ_j, τ_k) which has a smaller fidelity than the corresponding pair (σ_j, σ_k).

(3) The σ_j are independent of j (so $\forall j : \sigma_j := \bar{\sigma}$ for some fixed $\bar{\sigma}$) but τ_j varies nontrivially with j, so that the relation is one to many, and there can be no map $\mathcal{E}$ consistent with it.

1. Operational scenarios in which the setting variable *J* affects the system and the environment and *K* is trivial

Here, we provide quantum examples of types (a, 1), (a, 2), and (a, 3).

If the setting variable J influences both the principal system and the environment, one can prepare any initial joint state on the system and environment. Then, one can easily generate any relation $\mathfrak{R}$ *at all*, including all three types from above. The following arguments can be seen as a generalization of arguments originating with Pechukas [2].

For example, by taking the environment system to have dimension equal to the cardinality of the set $\{\sigma_j\}_j$, then for any set $\{\sigma_j\}_j$ one can prepare the system-environment joint state $\sigma_j \otimes |j\rangle\langle j|$ when $J = j$. The system and environment may then interact via the controlled channel $\mathcal{F}_{S_2|S_1 E}(\square_{S_1 E}) = \sum_j \langle j|(\mathrm{Tr}_{S_1}(\square_{S_1 E}))|j\rangle_E (\tau_j)_{S_2}$, for any set of states $\{\tau_j\}_j$ on S_2. Hence, for $J = j$ the final marginal state of the system is τ_j. Since both sets $\{\sigma_j\}_j$ and $\{\tau_j\}_j$ are completely unconstrained, one can certainly satisfy any of the conditions articulated in 1, 2, and 3.

2. Operational scenarios in which the setting variable *J* determines the measurement on the system and one postselects on the measurement outcome *K*

Here, we provide quantum examples of types (b, 1), (b, 2), and (b, 3).

Suppose that the initial state of the system-environment composite is a maximally entangled state. By the Hughston-Josza-Wootters theorem [69], it is possible, by implementing a j-dependent measurement on the system and postselecting on outcome k, to steer the environment to any arbitrary state $\tau_{j,k}$.[7] Further, the update rule of the measurement (which affects only the system) can be arbitrary. Given an outcome k of the jth measurement, one can find an update rule which ignores the state of the system and simply reprepares it in the state $\sigma_{j,k}$. If the subsequent system-environment interaction is a SWAP gate, one can generate the input-output relation $\mathfrak{R} = \{(\sigma_{j,k}, \tau_{j,k})\}_{j,k}$ for completely unconstrained sets $\{\sigma_{j,k}\}_{j,k}$ and $\{\tau_{j,k}\}_{j,k}$. Hence, one can certainly satisfy the conditions of 1, 2, and 3 (where j is replaced with j, k).

3. Operational scenarios in which the setting variable *J* affects only the system and *K* is trivial

Here, we provide quantum examples of types (c, 1), (c, 2), and (c, 3).

Suppose that one of four possible states are prepared by implementing a transformation on the system, as fol-

─────────

[7]Note that for a given choice j of measurement, the ensemble to which one steers is not entirely arbitrary: The weighted average of the states in the ensemble is fixed. However, we can avoid this restriction by postselecting on a particular outcome for each choice of measurement.

lows: One applies a "preparation" channel $\mathcal{G}_j \otimes \mathrm{id}_E$, for $j \in \{1, 2, 3, 4\}$ and for $\mathcal{G}_j : \mathcal{L}(\mathcal{H}_{S_0}) \to \mathcal{L}(\mathcal{H}_{S_1})$, to the fiducial state $|\Phi^+\rangle_{S_0 E} := 2^{-1/2}(|0\rangle_{S_0}|0\rangle_E + |1\rangle_{S_0}|1\rangle_E)$, where

$$\mathcal{G}_j(\square_{S_0}) = X_j(\square_{S_0})X_j^\dagger \tag{A1}$$

and where $\{X_j\}_j$ is the standard set of four Pauli matrices. In this case, there are four possible system-environment states, corresponding to the elements of the Bell basis,

$$(\rho_j)_{S_1 E} = (X_j \otimes \mathbb{1})|\Phi^+\rangle\langle\Phi^+|_{S_0 E}(X_j^\dagger \otimes \mathbb{1}). \tag{A2}$$

Each of these has the same marginal on S_1—the completely mixed state—so

$$\forall j : \sigma_j = \tfrac{1}{2}\mathbb{1}. \tag{A3}$$

Now suppose that the system-environment interaction consists of a measurement of the Bell basis on the system-environment composite and then a repreparation of the system in a state depending on the outcome that was obtained,

$$\mathcal{F}_{S_2|S_1 E}(\square_{S_1 E}) = \sum_j \langle\phi^+|(X_j \otimes \mathbf{1})(\square_{S_1 E})(X_j^\dagger \otimes \mathbf{1})$$
$$\times |\phi^+\rangle_{S_1 E}(\tau_j)_{S_2}, \tag{A4}$$

where $\{\tau_j\}_j$ denotes an arbitrary set of four states on S_2.

The input-output relation in this case, therefore, is easily verified to be

$$\mathfrak{R} = \left\{ \left(\tfrac{1}{2}\mathbb{1}, \tau_j\right) \right\}_j. \tag{A5}$$

If the τ_j depend nontrivially on j, then the relation is one to many, so we have an example of type 3.

We can modify this example slightly to get one of type 2. Simply let the initial preparation channel on the system, $\mathcal{G}_j$, be a Pauli unitary with probability $(1 - \epsilon)$ and a repreparation of an arbitrary state $\tilde{\sigma}_j$ with probability ϵ:

$$\mathcal{G}_j(\square_{S_1}) = (1 - \epsilon)X_j(\square_{S_1})X_j^\dagger + \epsilon\tilde{\sigma}_j, \tag{A6}$$

where the $\tilde{\sigma}_j$ can be drawn from an arbitrary set of four states $\{\tilde{\sigma}_j\}_j$. In this case, the four initial system-environment states one generates are

$$(\rho_j)_{S_1 E} = (1 - \epsilon)(X_j \otimes \mathbb{1})|\Phi^+\rangle\langle\Phi^+|_{S_0 E}(X_j^\dagger \otimes \mathbb{1})$$
$$+ \epsilon\tilde{\sigma}_j \otimes \tfrac{1}{2}\mathbb{1}. \tag{A7}$$

The four corresponding marginals on S_1 are

$$\sigma_j = (1 - \epsilon)\tfrac{1}{2}\mathbb{1} + \epsilon\tilde{\sigma}_j. \tag{A8}$$

The measurement of the Bell basis on the system-environment composite gives a uniform distribution over its outcomes if the joint state is of the form $\tilde{\sigma}_j \otimes \tfrac{1}{2}\mathbb{1}$, so that applying Eq. (A4) with a set $\{\tilde{\tau}_j\}_j$ rather than $\{\tau_j\}_j$ to the state in Eq. (A7) gives the four final marginal states of the system $\tau_j = (1 - \epsilon)\tilde{\tau}_j + \epsilon\tfrac{1}{4}\sum_j \tilde{\tau}_j$. Taking $\tfrac{1}{4}\sum_j \tilde{\tau}_j = \tfrac{1}{2}\mathbb{1}$ as a simple special case, we have

$$\tau_j = (1 - \epsilon)\tilde{\tau}_j + \epsilon\tfrac{1}{2}\mathbb{1}. \tag{A9}$$

The input-output relation in this case is therefore

$$\mathfrak{R} = \left\{ \left[(1 - \epsilon)\tfrac{1}{2}\mathbb{1} + \epsilon\tilde{\sigma}_j, (1 - \epsilon)\tilde{\tau}_j + \epsilon\tfrac{1}{2}\mathbb{1} \right] \right\}_j. \tag{A10}$$

Since both sets $\{\tilde{\sigma}_j\}_j$ and $\{\tilde{\tau}_j\}_j$ are completely unconstrained sets of four states, one can certainly choose them (for any ϵ)

to satisfy condition 2. One can also choose them to satisfy condition 1, since a set of four states can be informationally complete for a two-dimensional Hilbert space. The simplest such construction would take the set $\{\tilde{\sigma}_j\}_j$ to be an informationally complete set of pure states, $\epsilon = \frac{1}{2}$, and $\tilde{\tau}_j = T(\tilde{\sigma}_j)$ for all j, where T is the transpose operation.

APPENDIX B: GENERIC FAMILIES OF CLASSICAL EXAMPLES IN WHICH THE STANDARD PROPOSAL FAILS

All nine families of quantum examples from Appendix A have close classical analogs.

For simplicity of presentation, we represent the system-environment interaction in these examples by a conditional probability distribution $P_{S_2|S_1E}$ rather than the deterministic map $F_{S_2E'|S_1E}$. (Of course, the stochastic map induced by $P_{S_2|S_1E}$ can always be dilated to a deterministic map $F_{S_2E'|S_1E}$.)

We show that [for probability distributions $\omega_j \in \mathfrak{P}(S_1)$ and $\nu_j \in \mathfrak{P}(S_2)$] there exist operationally realizable input-output relations $\mathfrak{R} = \{(\omega_j, \nu_j)\}_j$ for which the following is true:

(1) There is no stochastic map Γ such that $\forall j : \Gamma(\omega_j) = \nu_j$.

(2) There is no linear map Γ such that $\forall j : \Gamma(\omega_j) = \nu_j$.

(3) There is no map Γ such that $\forall j : \Gamma(\omega_j) = \nu_j$.

Each point is demonstrated, respectively, by providing an operationally realizable input-output relation such that the following is true:

(1) The set $\{\omega_j\}_j$ is informationally complete (that is, forms a basis for the space of functions on S_1), and the convex hull of the $\{\omega_j\}_j$ has symmetries that are not shared by the full simplex of probability distributions on S_1, so that any map implementing such a symmetry transformation cannot be a mixture of permutations of the physical states and consequently cannot be a stochastic map.[8]

(2) The $\{\nu_j\}_j$ are more distinguishable than the $\{\omega_j\}_j$,[9] so that Γ is necessarily a nonlinear map.

(3) The ω_j are independent of j (so $\forall j : \omega_j = \bar{\omega}$ for some fixed $\bar{\omega}$) but ν_j varies nontrivially with j, so that the relation is one to many, and there can be no map Γ consistent with it.

We again demonstrate that all three of these failures occur in each of the three operational circuit types (listed as a,b,c in Appendix A).

1. Classical operational scenarios in which the setting variable J affects the system and the environment and K is trivial

Here, we provide classical examples of types (a, 1), (a, 2), and (a, 3).

If the setting variable J influences both the principal system and the environment, one can prepare any initial joint proba-

---

[8]This holds because transformations generated by symmetries of the full space of probability distributions $\mathfrak{P}(S_1)$ are reversible, and the only reversible transformations on the physical states correspond to permutations—that is, symmetries of the simplex; further, all stochastic maps can be expressed as a mixture of permutations. We give a specific example of this form in Appendix B 1.

[9]That is, there exists a pair (ν_j, ν_k) which has a smaller classical fidelity than the corresponding pair (μ_j, μ_k).

bility distribution on the system and environment. Then, as in the quantum case, one can easily generate any relation $\mathfrak{R}$ *at all*, including all three types from above.

For example, taking the environment system to have dimension equal to the cardinality of the set $\{\omega_j\}_j$, then for any set $\{\omega_j\}_j$ one can prepare the system-environment joint state $\omega_j \otimes [j]$ when $J = j$. The system and environment may then interact via the channel $\Gamma_{S_2|S_1E}(\square_{S_1E}) = \sum_j \sum_{S_1E} \delta_{j,E} \square_{S_1E}(\nu_j)_{S_2}$, for any set of states $\{\nu_j\}_j$. Hence, for $J = j$ the final marginal state of the system is ν_j. Since both sets $\{\omega_j\}_j$ and $\{\nu_j\}_j$ are completely unconstrained, one can certainly satisfy the conditions of 1, 2, and 3.

2. Classical scenarios in which the setting variable J determines the measurement on the system and one postselects on the measurement outcome K

Here, we provide classical examples of types (b, 1), (b, 2), and (b, 3).

Suppose that the initial state of the system-environment composite is a maximally correlated probability distribution. By the natural classical analog of the Hughston-Josza-Wootters theorem, it is possible, by implementing a j-dependent measurement on the system and postselecting on outcome k, to steer the environment to any arbitrary probability distribution $\nu_{j,k}$ (see footnote 7). Further, the update rule of the measurement (which affects only the system) can be arbitrary. Given an outcome k of the jth measurement, one can find an update rule which ignores the physical state of the system and simply repreprares the system in the state $\omega_{j,k}$. If the subsequent system-environment interaction is a swap gate, one can again generate the input-output relation $\mathfrak{R} = \{(\omega_{j,k}, \nu_{j,k})\}_{j,k}$ for completely unconstrained sets $\{\omega_{j,k}\}_{j,k}$ and $\{\nu_{j,k}\}_{j,k}$. Hence, one can certainly satisfy the conditions of 1, 2, and 3 (where j is replaced with j, k).

3. Classical operational scenarios in which the setting variable J affects only the system and K is trivial

Here, we provide classical examples of types (c, 1), (c, 2), and (c, 3).

Suppose that one of four possible states are prepared by implementing a transformation on the system, as follows. For setting $j \in \{1, 2, 3, 4\}$, one applies a transformation $\Gamma_{\Pi_j} \otimes \mathcal{I}$ to the fiducial state $P_{S_1E} := \frac{1}{4}([11] + [22] + [33] + [44])$, where $\mathcal{I}$ denotes the identity map and each Γ_{Π_j} is the map on probability distributions over system S_1 induced by the corresponding permutation of ontic states (in cycle notation):

$$\Pi_1 := (1)(2)(3)(4),$$
$$\Pi_2 := (12)(34),$$
$$\Pi_3 := (13)(24),$$
$$\Pi_4 := (14)(23). \tag{B1}$$

In this case, there are four possible system-environment distributions, corresponding to four maximally correlated probability distributions,

$$(P_j)_{S_1E} = \left(\Gamma_{\Pi_j} \otimes \mathcal{I}\right)\left(P_{S_1E}\right), \quad j = \{1, 2, 3, 4\}. \tag{B2}$$

Each of these has the same marginal on S_1, namely, the uniform probability distribution

$$\omega_j = \tfrac{1}{4}([1] + [2] + [3] + [4]) := \bar{\omega}. \tag{B3}$$

Now suppose that the system-environment interaction consists of a measurement distinguishing the supports of the maximally correlated states defined by Eq. (B2) (that is, a measurement whose outcome j can be associated to the response function $\delta_{\Pi_j(S_1),E}$), followed by a repreparation of the system in the state ν_j, thereby enacting the stochastic map

$$\Gamma_{S_2|S_1E}\big(\square_{S_1E}\big) = \sum_j \sum_{S_1E} \delta_{\Pi_j(S_1),E} \, \square_{S_1E}(\nu_j)_{S_2}, \tag{B4}$$

where $\{\nu_j\}_j$ denotes an arbitrary set of four probability distributions.

The input-output relation in this case is

$$\mathfrak{R} = \{(\bar{\omega}, \nu_j)\}_j. \tag{B5}$$

If the ν_j depend nontrivially on j, then the relation is one to many, so we have an example of type 3.

We can modify this example slightly to get one of type 2. Simply let the initial transformation Γ_{Π_j} on the system be generated by the permutation Π_j of Eq. (B1) with probability $(1-\epsilon)$ and a repreparation of an arbitrary state $\tilde{\omega}_j$ with probability ϵ, so that

$$\Gamma_j\big(\square_{S_1}\big) = (1-\epsilon)\Gamma_{\Pi_j}\big(\square_{S_1}\big) + \epsilon\tilde{\omega}. \tag{B6}$$

In this case, the four initial system-environment states are

$$(P_j)_{S_1E} = (1-\epsilon)\big(\Gamma_{\Pi_j} \otimes \mathcal{I}\big)\big(P_{S_1E}\big) + \epsilon\tilde{\omega} \otimes \bar{\omega}. \tag{B7}$$

The marginals on S_1 are

$$\omega_j = (1-\epsilon)\bar{\omega} + \epsilon\tilde{\omega}. \tag{B8}$$

The measurement defined above gives a uniform distribution over its outcomes if the joint probability distribution is of the form $\omega_j \otimes \bar{\omega}$, so that applying Eq. (B4) with ν_j replaced by $\tilde{\nu}_j$ to the state in Eq. (B7) gives the four final marginal states of the system $\nu_j = (1-\epsilon)\tilde{\nu}_j + \epsilon\frac{1}{4}\sum_j \tilde{\nu}_j$. Taking $\frac{1}{4}\sum_j \tilde{\nu}_j = \bar{\omega}$ as a simple special case, we have

$$\nu_j = (1-\epsilon)\tilde{\nu}_j + \epsilon\bar{\omega}. \tag{B9}$$

The input-output relation in this case, therefore, is

$$\mathfrak{R} = \{[(1-\epsilon)\bar{\omega} + \epsilon\tilde{\omega}_j, (1-\epsilon)\tilde{\nu}_j + \epsilon\bar{\omega}]\}_j. \tag{B10}$$

Since both sets $\{\tilde{\omega}_j\}_j$ and $\{\tilde{\nu}_j\}_j$ are completely unconstrained sets of four states, one can certainly choose them (for any ϵ) to satisfy condition 2.

One can also choose them to satisfy condition 1. As one example, let $\epsilon = \frac{1}{2}$, and let the $\tilde{\omega}_j$ be

$$\tilde{\omega}_1 := \tfrac{1}{2}[1] + \tfrac{1}{2}[2],$$
$$\tilde{\omega}_2 := \tfrac{1}{2}[1] + \tfrac{1}{2}[3],$$
$$\tilde{\omega}_3 := \tfrac{1}{2}[1] + \tfrac{1}{2}[4],$$
$$\tilde{\omega}_4 := \tfrac{1}{2}[3] + \tfrac{1}{2}[4]. \tag{B11}$$

The set $\{\tilde{\omega}_j\}$ is a basis for the space of distributions over the four physical states. Let the ν_j be $\tilde{\nu}_j = 2\bar{\omega} - \tilde{\omega}_j$, so

that

$$\tilde{\nu}_1 := \tfrac{1}{2}[3] + \tfrac{1}{2}[4],$$
$$\tilde{\nu}_2 := \tfrac{1}{2}[2] + \tfrac{1}{2}[4],$$
$$\tilde{\nu}_3 := \tfrac{1}{2}[2] + \tfrac{1}{2}[3],$$
$$\tilde{\nu}_4 := \tfrac{1}{2}[1] + \tfrac{1}{2}[2]. \tag{B12}$$

The relation in such circumstances is

$$\mathfrak{R} = \Big\{\big[\tfrac{1}{2}\big(\bar{\omega} + \tfrac{1}{2}[1] + \tfrac{1}{2}[2]\big), \tfrac{1}{2}\big(\bar{\omega} + \tfrac{1}{2}[3] + \tfrac{1}{2}[4]\big)\big],$$
$$\big[\tfrac{1}{2}\big(\bar{\omega} + \tfrac{1}{2}[1] + \tfrac{1}{2}[3]\big), \tfrac{1}{2}\big(\bar{\omega} + \tfrac{1}{2}[2] + \tfrac{1}{2}[4]\big)\big],$$
$$\big[\tfrac{1}{2}\big(\bar{\omega} + \tfrac{1}{2}[1] + \tfrac{1}{2}[4]\big), \tfrac{1}{2}\big(\bar{\omega} + \tfrac{1}{2}[2] + \tfrac{1}{2}[3]\big)\big],$$
$$\big[\tfrac{1}{2}\big(\bar{\omega} + \tfrac{1}{2}[3] + \tfrac{1}{2}[4]\big), \tfrac{1}{2}\big(\bar{\omega} + \tfrac{1}{2}[1] + \tfrac{1}{2}[2]\big)\big]\Big\}. \tag{B13}$$

There is a linear map which fits the input-output relation: It is given by the matrix

$$\begin{bmatrix} -\nicefrac{1}{2} & \nicefrac{1}{2} & \nicefrac{1}{2} & \nicefrac{1}{2} \\ \nicefrac{1}{2} & -\nicefrac{1}{2} & \nicefrac{1}{2} & \nicefrac{1}{2} \\ \nicefrac{1}{2} & \nicefrac{1}{2} & -\nicefrac{1}{2} & \nicefrac{1}{2} \\ \nicefrac{1}{2} & \nicefrac{1}{2} & \nicefrac{1}{2} & -\nicefrac{1}{2} \end{bmatrix} \tag{B14}$$

(acting on probability distributions over the four physical states of S_1, expressed as four-component vectors). This linear map is not stochastic, since the elements in the matrix are not all positive. Because the set of input distributions in the relation form a basis for the space of all distributions over physical states, this is the *unique* linear map satisfying the relation. Hence, in such an example, the only linear map consistent with the standard proposal is nonstochastic.

Geometrically, the convex hull of the distributions in Eqs. (B11) and (B12) forms an octahedron inside the 4-simplex of all probability distributions. The evolution map one finds by the standard proposal [defined by Eq. (B13)] corresponds to a reflection through the origin, which is a reversible injective map. However, this map (extended uniquely by linearity to the whole space) does not correspond to a symmetry of the 4-simplex. Hence, the linear map in this example does not correspond to any permutation of the physical states.

APPENDIX C: WHEN DOES THE EVOLUTION MAP COINCIDE WITH THE INFERENCE MAP CLASSICALLY?

Consider the classical circuit of Fig. 7(b). Generally, it is possible to have inferences along a common cause pathway and therefore possible that the inference map might differ from the evolution map. What are the necessary and sufficient conditions on P_{S_1E} for them to be the same for all $F_{S_2E'|S_1E}$? The answer is simply that E and S_1 must be marginally independent,

$$P_{S_1E} = P_{S_1} \otimes P_E. \tag{C1}$$

A sufficient condition for this independence to hold is that S_1 and E have no common ancestors in the causal structure. However, the independence relation can sometimes hold when this causal condition fails, for instance, using fine-tuned choices of circuit elements [28,38].

If one is considering a circuit that contains some variables which one is conditioning upon (like J and K in the examples

we considered previously), then the sufficient condition for equality of the inference and evolution maps is simply that one must have S_1 independent of E after conditioning on these variables. It is worth noting that this provides another example (besides the possibility of fine-tuning) of how a statistical independence can hold in spite of the variables failing to have disconnected causal ancestry—in classical example 2, for instance, J is a common cause of S_1 and E, but if one conditions upon it, then S_1 and E become independent.

Therefore, the precise circumstances under which confusing inference and evolution has no consequences are when there are no initial system-environment correlations. In other words, the first mistake of the standard proposal, conflating inference and evolution, is problematic precisely because the standard proposal sought to address scenarios with such correlations.

APPENDIX D: KNOWLEDGE DEPENDENCE OF EVOLUTION MAPS

We mentioned in Secs. V C and VI that the classical and quantum evolution maps depend on one's knowledge of the state of the environment. We now provide several textbook examples of this phenomena, simply to show that it is not a cause for concern.

Consider encoding a single bit using a Vernam cypher (one-time pad). Let S be the bit to be encoded (the plain text) and E be the key, which is shared between the sender and the receiver. The value of E is sampled uniformly at random. To encode the plain text, a controlled-NOT is implemented on S with the key E as control, as shown in Fig. 11. Because the users of the cypher know the value of E, they describe the evolution of S, conditional on the value of E, as follows: If $E = 0$, the evolution map is the identity, while if $E = 1$, it is the bit flip. An eavesdropper assigns to E the uniformly random distribution $\frac{1}{2}([0] + [1])$ and therefore describes the evolution of S as the randomization channel (the equal mixture of identity and bit-flip channels).

The knowledge dependence of one's description of evolution is also seen in many protocols for quantum cryptography. For example, a private quantum channel on a qubit [70] can be implemented as follows. The sender and receiver share a uniformly random two-bit key. The sender draws a Pauli matrix from the set $\{I, X, Y, Z\}$ based on the value of the key and implements it on the system. To decode, the user applies the same Pauli. The users of the channel, who know the key's value, describe the evolution as $\rho \to \sigma_i \rho \sigma_i^T$. An eavesdropper who does *not* know the key's value but knows only that the
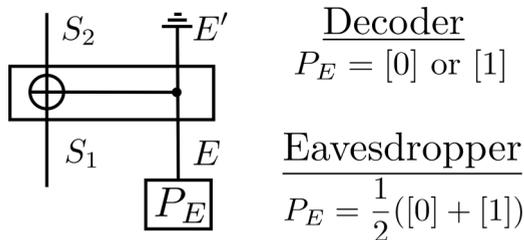
key was drawn uniformly at random, describes the evolution as $\rho \to \frac{1}{2}\mathbb{1}$.

As a final example, consider the transformation that a quantum state undergoes when it is measured. The map which one uses to update the state of the measured system depends on how much one knows about the outcome of the measurement. Consider for simplicity a rank-1 projective measurement $\{\Pi_k\}_k$ performed on some initial state ρ. If one has no information about which outcome occurred, then the correct update map is given by the nonselective update rule, $\rho \to \sum_k \Pi_k \rho \Pi_k$. On the other hand, if one knows that outcome k occurred, then the correct update map is given by the selective update rule, $\rho \to \Pi_k \rho \Pi_k$.

APPENDIX E: ANALYSIS OF THE THREE CLASSICAL EXAMPLES OF SEC. III

We now provide a derivation of the claims of Sec. V D, completing the analysis of the three classical examples from Sec. III and the mistakes of the standard proposal when applied to them. For convenience, we represent the system-environment interaction in each case by the probabilistic dependence $P_{S_2|S_1E} = \sum_{E'} P_{S_2E'|S_1E}$, where $P_{S_2E'|S_1E}$ is the conditional associated to the stochastic map $F_{S_2E'|S_1E}$; we also indicate conditioning on $J = j$ or $K = k$ in the subscripts [e.g., $P_{E|J=0}(e) := P_{E|J}(e|0)$].

1. Classical example 3, from Sec. III E

Consider again the third classical example, introduced in Sec. III E and discussed in Sec. V D.

As indicated by Eq. (75), the inference map from S_1 to S_2 for a particular value j of J is

$$\Gamma_{S_2|S_1}^{(j)}(\square_{S_1}) = \sum_{S_1 E} P_{S_2|S_1E} P_{E|S_1J=j} \square_{S_1}. \tag{E1}$$

Given the nature of the coupling of system and environment, we have

$$P_{S_2|S_1E} = \delta_{S_2, S_1 \oplus_2 E}. \tag{E2}$$

Denote the early version of S_1 by S_0. Then, because the joint distribution over S_0 and E is a state of perfect positive correlation with uniform marginals,

$$P_{S_0E} = \frac{1}{2}[0]_{S_0} \otimes [0]_E + \frac{1}{2}[1]_{S_0} \otimes [1]_E$$
$$= \left(\frac{1}{2}[0]_{S_0} + \frac{1}{2}[1]_{S_0}\right)\delta_{E, S_0}, \tag{E3}$$

and because the controlled-NOT from J to S_0 is modeled by the conditional

$$P_{S_1|S_0} = \delta_{S_1, S_0 \oplus_2 J}, \tag{E4}$$

a simple application of Bayesian probability theory implies that

$$P_{E|S_1} = \sum_{S_0} P_{E|S_0} P_{S_0|S_1}$$
$$= \sum_{S_0} \delta_{E, S_0} \delta_{S_0, S_1 \oplus_2 J}$$
$$= \delta_{E, S_1 \oplus_2 J}. \tag{E5}$$

FIG. 11. (a) A one-time pad. The decoder and eavesdropper have distinct information about the key (E).

Finally, substituting Eqs. (E2) and (E5) into Eq. (E1), we conclude that

$$\Gamma^{(j)}_{S_2|S_1}\big(\square_{S_1}\big) = \sum_{S_1 E} \delta_{S_2, S_1 \oplus_2 E} \delta_{E, S_1 \oplus_2 j} \square_{S_1} \tag{E6}$$

$$= \delta_{S_2, j} \tag{E7}$$

$$= [j]_{S_2}. \tag{E8}$$

This constitutes the proof of Eq. (76).

The correct evolution map in this example is straightforward to identify using the prescription of Sec. IV A. For both values of J, the marginal state of the environment is uniformly random, so there is a unique evolution map. It is easily seen to be the randomization channel, which takes any input probability distribution P_{S_1} to a uniformly random distribution on S_2, as stated in Eq. (77).

2. Classical example 2, from Sec. III D

We start with Eq. (78), in the form

$$\Gamma^{(j)}_{S_2|S_1}\big(\square_{S_1}\big) = \sum_{S_1 E} P_{S_2|S_1 E} P_{E|J=j} \square_{S_1}. \tag{E9}$$

Recalling the details of the example, one has

$$P_{S_2|S_1 E} = \delta_{S_2, (S_1 \oplus_3 E)}. \tag{E10}$$

Furthermore, because E simply tracks J, we have

$$P_{E|J} = \delta_{E, J}. \tag{E11}$$

Substituting these expressions into Eq. (E9), we conclude that

$$\Gamma^{(j)}_{S_2|S_1}\big(\square_{S_1}\big) = \sum_{S_1 E} \delta_{S_2, (S_1 \oplus_3 E)} \delta_{E, J} \square_{S_1} \tag{E12}$$

$$= \sum_{S_1} \delta_{S_2, (S_1 \oplus_3 j)} \square_{S_1}, \tag{E13}$$

confirming Eq. (79)

3. Classical example 1, from Sec. III C

Here, both J and K are nontrivial. The causal structure is the same as that of Fig. 3 and ensures that E is conditionally independent of S_1 given JK, $P_{E|S_1 JK} = P_{E|JK}$. Thus, the inference map for $J = j$ and $K = k$ is

$$\Gamma^{(j,k)}_{S_2|S_1}\big(\square_{S_1}\big) = \sum_{S_1 E} P_{S_2|S_1 E}\big(P_{E|J=j, K=k} \otimes \square_{S_1}\big). \tag{E14}$$

Because of the perfect correlation between S and E in the initial joint state, whatever one infers about S from learning that $j = j$ and $K = k$, one learns the same thing about E as well. Consequently,

$$P_{E|J=1, K=1} = \tfrac{1}{2}[2]_E + \tfrac{1}{2}[3]_E, \tag{E15}$$

$$P_{E|J=2, K=1} = \tfrac{1}{2}[1]_E + \tfrac{1}{2}[3]_E, \tag{E16}$$

$$P_{E|J=3, K=1} = \tfrac{1}{2}[1]_E + \tfrac{1}{2}[2]_E. \tag{E17}$$

The system and environment interact via a swap, so

$$P_{S_2|S_1 E} = \delta_{S_2, E}. \tag{E18}$$

Substituting these expressions into Eq. (E14), we have

$$\Gamma^{(1,1)}_{S_2|S_1}\big(\square_{S_1}\big) = \tfrac{1}{2}[2]_{S_2} + \tfrac{1}{2}[3]_{S_2}, \tag{E19}$$

$$\Gamma^{(2,1)}_{S_2|S_1}\big(\square_{S_1}\big) = \tfrac{1}{2}[1]_{S_2} + \tfrac{1}{2}[3]_{S_2}, \tag{E20}$$

$$\Gamma^{(3,1)}_{S_2|S_1}\big(\square_{S_1}\big) = \tfrac{1}{2}[1]_{S_2} + \tfrac{1}{2}[2]_{S_2}. \tag{E21}$$

The inference map for $K = 1$ and for the different values of J is simply the map that ignores the distribution on S_1 and outputs whatever distribution one infers for E as the distribution on S_2. Clearly, therefore, the inference map is JK dependent in this example.

Recalling Eq. (42), one sees that each of the input-output pairs (for a given value of J when $K = 1$) is consistent with the corresponding inference map in Eqs. (E19)–(E21). But if one tries to find a *single* map that is consistent with *all three* input-output pairs, as the standard proposal mistakenly suggests we do, one finds that no stochastic map can do the job.

As with example 2, it happens that the evolution map for any given values of J and K coincides with the inference map for those values because the causal structure is such that S_1 and E are conditionally independent given JK, so that an intervention on S_1 does not change the map after conditioning on JK,

$$\Gamma^{(j,k)}_{S_2|\mathrm{do}S_1} = \Gamma^{(j,k)}_{S_2|S_1}. \tag{E22}$$

APPENDIX F: THE CORRECT QUANTUM EVOLUTION MAPS FOR SCENARIOS WITH NONTRIVIAL J AND K

Consider the scenario at play in the standard proposal, shown in Fig. 2. Following the prescription of Sec. VI, one obtains (for $J = j$ and $K = k$) a map in the modified circuit given by

$$\mathcal{E}^{(j,k)}_{S_2|S_1'}\big(\square_{S_1'}\big) = \mathrm{Tr}_{E'}\big(\mathcal{U}_{S_2 E'|S_1' E}\big(\square_{S_1'} \otimes \rho^{(j,k)}_E\big)\big) \tag{F1}$$

where $\rho^{(j,k)}_E = \mathrm{Tr}_{S_1}(\rho^{(j,k)}_{S_1 E})$. Thus, by Eq. (81), the evolution map from S_1 to S_2 in the circuit of Fig. 2, given that the setting variable has value j and the outcome was found to be k, is

$$\mathcal{E}^{(j,k)}_{S_2|\mathrm{do}(S_1)}\big(\square_{S_1}\big) = \mathrm{Tr}_{E'}\big(\mathcal{U}_{S_2 E'|S_1 E}\big(\square_{S_1} \otimes \rho^{(j,k)}_E\big)\big). \tag{F2}$$

This is the quantum analog of Eq. (74).

As was the case classically, the evolution map generally depends on the values of the parameters JK, through the marginal state of the environment appearing in Eq. (F2).

1. Resolution of the three quantum examples in Sec. II

Using Eq. (F2), we can compute the correct evolution maps for the three quantum examples of Sec. II.

In the first quantum example (Sec. II C), the evolution map for $J = j$ and $K = 1$ is

$$\mathcal{E}^{(j,1)}_{S_2|\mathrm{do}(S_1)}\big(\square_{S_1}\big) = \big|\psi^T_{j,1}\big\rangle\big\langle\psi^T_{j,1}\big|_{S_2}. \tag{F3}$$

In other words, for each value j of J (which corresponded to preparing $|\psi_{j,1}\rangle_{S_1}$ as the marginal state of the system), one has a distinct map, which ignores the input on the system and reprepares the fixed state $|\psi^T_{j,1}\rangle_{S_2}$.

In the second quantum example (Sec. II D), the evolution maps for $J = 0$ and $J = 1$ are

$$\mathcal{E}^{(0)}_{S_2|\mathrm{do}(S_1)}(\square_{S_1}) = \square_{S_1},\tag{F4}$$

$$\mathcal{E}^{(1)}_{S_2|\mathrm{do}(S_1)}(\square_{S_1}) = (\mathrm{XH})\square_{S_1}(\mathrm{XH})^\dagger.\tag{F5}$$

That is, for $J = 0$ the controlled operation is not activated, while for $J = 1$ it is.

For these two quantum examples, if one considers specific values for J and K, the input-output pair associated to that valuation is consistent with the evolution map for that valuation. This is because in both examples conditioning on J and K makes S_1 and E marginally independent in the sense that

$$\rho_{S_1E|J=j,K=k} = \rho_{S_1|J=j,K=k} \otimes \rho_{E|J=j,K=k}\tag{F6}$$

and therefore satisfies the condition of Eq. (G1) for the equivalence of the evolution map and the inference map.

In the third quantum example (Sec. II E), there is a unique evolution map $\mathcal{E}^{(j,k)} = \mathcal{E}$, since $\rho_E^{(j,k)} = \rho_E$. It is

$$\mathcal{E}_{S_2|\mathrm{do}(S_1)}(\square_{S_1}) = \tfrac{1}{2}\square_{S_1} + \tfrac{1}{2}Z\,\square_{S_1}Z^\dagger,\tag{F7}$$

the dephasing map in the Z basis. This is intuitively clear from the fact that we are performing a flip of the system in the Z basis, conditioned on an environment whose marginal state is completely mixed.

In this example, each individual input-output pair for a given value of J and K is *not* consistent with the evolution map. This is because, for a given value of J, if you came to know something about S_1, it would inform you about S_2 *both* by the cause-effect connection between S_1 and S_2 *and* by the common-cause connection between S_1 and S_2 that is mediated by E. The precise formalism for modeling this sort of inference in quantum theory, however, is not yet clear.

APPENDIX G: WHEN DO INDIVIDUAL INPUT-OUTPUT PAIRS OF QUANTUM STATES CONSTRAIN THE QUANTUM EVOLUTION MAP?

Here, we show that the necessary and sufficient condition for an individual input-output pair appearing in the standard proposal to be a constraint on the evolution map is that S_1 and E are marginally independent,

$$\rho_{S_1E} = \rho_{S_1} \otimes \rho_E.\tag{G1}$$

The proof is by contradiction. If the joint state on S_1E did not factorize, then there would be a choice of unitary in the circuit such that S_2 depended explicitly on the correlations between S_1 and E for a given value of JK. The input-output pair for that value of JK would then reflect this dependence. On the other hand, by the definition of the quantum evolution map, its output on S_2 *cannot* depend on the correlations between S_1 and E, and so cannot yield the same input-output pair.

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, 10th ed. (Cambridge University Press, New York, 2011).

[2] P. Pechukas, Reduced Dynamics Need not be Completely Positive, Phys. Rev. Lett. **73**, 1060 (1994).

[3] R. Alicki, Comment on "Reduced Dynamics Need not be Completely Positive," Phys. Rev. Lett. **75**, 3020 (1995).

[4] P. Pechukas, Pechukas Replies, Phys. Rev. Lett. **75**, 3021 (1995).

[5] P. Štelmachovič and V. Bužek, Dynamics of open quantum systems initially entangled with environment: Beyond the Kraus representation, Phys. Rev. A **64**, 062106 (2001).

[6] H. Hayashi, G. Kimura, and Y. Ota, Kraus representation in the presence of initial correlations, Phys. Rev. A **67**, 062109 (2003).

[7] K. M. Fonseca Romero, P. Talkner, and P. Hänggi, Is the dynamics of open quantum systems always linear? Phys. Rev. A **69**, 052109 (2004).

[8] T. F. Jordan, A. Shaji, and E. C. G. Sudarshan, Dynamics of initially entangled open quantum systems, Phys. Rev. A **70**, 052110 (2004).

[9] A. Shaji and E. Sudarshan, Who's afraid of not completely positive maps? Phys. Lett. A **341**, 48 (2005).

[10] T. F. Jordan, A. Shaji, and E. C. G. Sudarshan, Maps for Lorentz transformations of spin, Phys. Rev. A **73**, 032104 (2006).

[11] H. A. Carteret, D. R. Terno, and K. Życzkowski, Dynamics beyond completely positive maps: Some properties and applications, Phys. Rev. A **77**, 042113 (2008).

[12] C. A. Rodríguez-Rosario, K. Modi, A. M. Kuah, A. Shaji, and E. C. G. Sudarshan, Completely positive maps and classical correlations, J. Phys. A: Math. Theor. **41**, 205301 (2008).

[13] A. Shabani and D. A. Lidar, Vanishing Quantum Discord is Necessary and Sufficient for Completely Positive Maps, Phys. Rev. Lett. **102**, 100402 (2009); Erratum: Vanishing Quantum Discord is Necessary and Sufficient for Completely Positive Maps [Phys. Rev. Lett. **102**, 100402 (2009)], **116**, 049901(E) (2016).

[14] C. A. Rodríguez-Rosario, K. Modi, and A. Aspuru-Guzik, Linear assignment maps for correlated system-environment states, Phys. Rev. A **81**, 012313 (2010).

[15] M. E. Cuffaro and W. C. Myrvold, On the debate concerning the proper characterization of quantum dynamical evolution, Philos. Sci. **80**, 1125 (2013).

[16] A. Brodutch, A. Datta, K. Modi, A. Rivas, and C. A. Rodríguez-Rosario, Vanishing quantum discord is not necessary for completely positive maps, Phys. Rev. A **87**, 042301 (2013).

[17] J. M. Dominy, A. Shabani, and D. A. Lidar, A general framework for complete positivity, Quant. Info. Proc. **15**, 465 (2015).

[18] E. Chitambar, A. Abu-Nada, R. Ceballos, and M. Byrd, Restrictions on initial system-environment correlations based on the dynamics of an open quantum system, Phys. Rev. A **92**, 052110 (2015).

[19] J. M. Dominy and D. A. Lidar, Beyond complete positivity, Quant. Info. Proc. **15**, 1349 (2016).

[20] K. Modi, Operational approach to open dynamics and quantifying initial correlations, Sci. Rep. **2**, 581 (2012).

[21] S. Milz, F. A. Pollock, and K. Modi, An introduction to operational quantum dynamics, Open Syst. Inf. Dyn. **24**, 1740016 (2017).

[22] K. Modi and E. C. G. Sudarshan, Role of preparation in quantum process tomography, Phys. Rev. A **81**, 052119 (2010).

[23] K. Modi, Preparation of states in open quantum mechanics, Open Syst. Inf. Dyn. **18**, 253 (2011).

[24] M. Ringbauer, C. J. Wood, K. Modi, A. Gilchrist, A. G. White, and A. Fedrizzi, Characterizing Quantum Dynamics with Initial System-Environment Correlations, Phys. Rev. Lett. **114**, 090402 (2015).

[25] F. A. Pollock, C. Rodríguez-Rosario, T. Frauenheim, M. Paternostro, and K. Modi, Non-Markovian quantum processes: Complete framework and efficient characterization, Phys. Rev. A **97**, 012127 (2018).

[26] F. A. Pollock, C. Rodríguez-Rosario, T. Frauenheim, M. Paternostro, and K. Modi, Operational Markov Condition for Quantum Processes, Phys. Rev. Lett. **120**, 040405 (2018).

[27] S. Milz, F. A. Pollock, T. P. Le, G. Chiribella, and K. Modi, Entanglement, non-Markovianity, and causal non-separability, New J. Phys. **20**, 033033 (2018).

[28] J. Pearl, *Causality: Models, Reasoning, and Inference* (Cambridge University Press, New York, 2000).

[29] J. Pearl and D. Mackenzie, *Book of Why: The New Science of Cause and Effect* (Penguin Books, New York, 2019).

[30] K. Ried, M. Agnew, L. Vermeyden, D. Janzing, R. W. Spekkens, and K. J. Resch, A quantum advantage for inferring causal structure, Nat. Phys. **11**, 414 (2015).

[31] T. Fritz, Beyond Bell's theorem: Correlation scenarios, New J. Phys. **14**, 103001 (2012).

[32] J. Henson, R. Lal, and M. F. Pusey, Theory-independent limits on correlations from generalized Bayesian networks, New J. Phys. **16**, 113043 (2014).

[33] R. Chaves, L. Luft, and D. Gross, Causal structures from entropic information: Geometry and novel scenarios, New J. Phys. **16**, 043001 (2014).

[34] R. Chaves, C. Majenz, and D. Gross, Information-theoretic implications of quantum causal structures, Nat. Commun. **6**, 5766 (2015).

[35] M. Weilenmann and R. Colbeck, Inability of the entropy vector method to certify nonclassicality in linelike causal structures, Phys. Rev. A **94**, 042112 (2016).

[36] J. Pienaar, Which causal structures might support a quantum-classical gap? New J. Phys. **19**, 043021 (2017).

[37] R. Chaves, G. Carvacho, I. Agresti, V. Di Giulio, L. Aolita, S. Giacomini, and F. Sciarrino, Quantum violation of an instrumental test, Nat. Phys. **14**, 291 (2018).

[38] C. J. Wood and R. W. Spekkens, The lesson of causal discovery algorithms for quantum correlations: Causal explanations of Bell-inequality violations require fine-tuning, New J. Phys. **17**, 033002 (2015).

[39] M. S. Leifer and R. W. Spekkens, Towards a formulation of quantum theory as a causally neutral theory of Bayesian inference, Phys. Rev. A **88**, 052130 (2013).

[40] E. G. Cavalcanti and R. Lal, On modifications of Reichenbach's principle of common cause in light of Bell's theorem, J. Phys. A: Math. Theor. **47**, 424018 (2014).

[41] J.-M. A. Allen, J. Barrett, D. C. Horsman, C. M. Lee, and R. W. Spekkens, Quantum Common Causes and Quantum Causal Models, Phys. Rev. X **7**, 031021 (2017).

[42] E. G. Cavalcanti, Classical Causal Models for Bell and Kochen-Specker Inequality Violations Require Fine-Tuning, Phys. Rev. X **8**, 021018 (2018).

[43] M. S. Leifer and R. W. Spekkens, A Bayesian approach to compatibility, improvement, and pooling of quantum states, J. Phys. A: Math. Theor. **47**, 275301 (2014).

[44] J.-P. W. MacLean, K. Ried, R. W. Spekkens, and K. J. Resch, Quantum-coherent mixtures of causal relations, Nat. Commun. **8**, 15149 (2017).

[45] D. Ebler, S. Salek, and G. Chiribella, Enhanced Communication with the Assistance of Indefinite Causal Order, Phys. Rev. Lett. **120**, 120502 (2018).

[46] A. Feix and C. Brukner, Quantum superpositions of "common-cause" and "direct-cause" causal structures, New J. Phys. **19**, 123028 (2017).

[47] W. F. Stinespring, Positive functions on C*-algebras, Proc. Am. Math. Soc. **6**, 211 (1955).

[48] K. S. Ried, Causal models for a quantum world, Ph.D. thesis, University of Waterloo, Waterloo, Canada, 2016.

[49] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete? Phys. Rev. **47**, 777 (1935).

[50] E. Schrodinger, Discussion of probability relations between separated systems, Math. Proc. Cambridge Philos. Soc. **31**, 555 (1935).

[51] D. S. Abrams and S. Lloyd, Nonlinear Quantum Mechanics Implies Polynomial-Time Solution for *Np*-Complete and # *p* Problems, Phys. Rev. Lett. **81**, 3992 (1998).

[52] A. Peres, Nonlinear Variants of Schrödinger's Equation Violate the Second Law of Thermodynamics, Phys. Rev. Lett. **63**, 1114 (1989).

[53] N. Gisin, Weinberg's non-linear quantum mechanics and supraluminal communications, Phys. Lett. A **143**, 1 (1990).

[54] K. K. Sabapathy, J. S. Ivan, S. Ghosh, and R. Simon, Quantum discord plays no distinguished role in characterization of complete positivity: Robustness of the traditional scheme, arXiv:1304.4857.

[55] D. Horsman, C. Heunen, M. F. Pusey, J. Barrett, and R. W. Spekkens, Can a quantum state over time resemble a quantum state at a single time? Proc. R. Soc. London, Ser. A **473**, 2205 (2017).

[56] M. M. Wolf and J. I. Cirac, Dividing quantum channels, Commun. Math. Phys. **279**, 147 (2008).

[57] M. M. Wolf, J. Eisert, T. S. Cubitt, and J. I. Cirac, Assessing Non-Markovian Quantum Dynamics, Phys. Rev. Lett. **101**, 150402 (2008).

[58] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Theoretical framework for quantum networks, Phys. Rev. A **80**, 022339 (2009).

[59] M. S. Leifer, Quantum dynamics as an analog of conditional probability, Phys. Rev. A **74**, 042310 (2006).

[60] Y. Aharonov, S. Popescu, J. Tollaksen, and L. Vaidman, Multiple-time states and multiple-time measurements in quantum mechanics, Phys. Rev. A **79**, 052110 (2009).

[61] O. Oreshkov, F. Costa, and Č. Brukner, Quantum correlations with no causal order, Nat. Commun. **3**, (2012).

[62] R. Silva, Y. Guryanova, N. Brunner, N. Linden, A. J. Short, and S. Popescu, Pre- and postselected quantum states: Density matrices, tomography, and Kraus operators, Phys. Rev. A **89**, 012121 (2014).

[63] J. F. Fitzsimons, J. A. Jones, and V. Vedral, Quantum correlations which imply causation, Sci. Rep. **5**, 18281 (2015).

[64] L. Hardy, The operator tensor formulation of quantum theory, Philos. Trans. R. Soc. London A **370**, 3385 (2012).

[65] A.-M. Kuah, K. Modi, C. A. Rodríguez-Rosario, and E. C. G. Sudarshan, How state preparation can affect a quantum experiment: Quantum process tomography for open systems, Phys. Rev. A **76**, 042113 (2007).

[66] Y. S. Weinstein, T. F. Havel, J. Emerson, N. Boulant, M. Saraceno, S. Lloyd, and D. G. Cory, Quantum process tomography of the quantum Fourier transform, J. Chem. Phys. **121**, 6117 (2004).

[67] M. D. Mazurek, M. F. Pusey, K. J. Resch, and R. W. Spekkens, Experimentally bounding deviations from quantum

theory in the landscape of generalized probabilistic theories, arXiv:1710.05948.

[68] F. Costa and S. Shrapnel, Quantum causal modelling, New J. Phys. **18**, 063032 (2016).

[69] L. P. Hughston, R. Jozsa, and W. K. Wootters, A complete classification of quantum ensembles having a given density matrix, Phys. Lett. A **183**, 14 (1993).

[70] A. Ambainis, M. Mosca, A. Tapp, and R. De Wolf, Private quantum channels, in *Foundations of Computer Science, 2000. Proceedings. 41st Annual Symposium on* (IEEE, 2000), pp. 547–553.