Quantum superpositions of causal orders as an operational resource

Márcio M. Taddei [1,*] Ranieri V. Nery,[1,2] and Leandro Aolita[1]

[1]*Instituto de Física, Federal University of Rio de Janeiro, 21941-972, P. O. Box 68528, Rio de Janeiro, Brazil*
[2]*International Institute of Physics, Federal University of Rio Grande do Norte, 59078-970, P. O. Box 1613, Natal, Brazil*

Causal nonseparability refers to processes where events take place in a coherent superposition of different causal orders. These may be the key resource for experimental violations of causal inequalities and have been recently identified as resources for concrete information-theoretic tasks. Here, we take a step forward by deriving a complete operational framework for causal nonseparability as a resource. Our first contribution is a formal definition for the specific notion of quantum control of causal orders, a stronger form of causal nonseparability—with the celebrated quantum switch as best-known example—where the causal orders of events for a target system are coherently controlled by a control system. We then build a resource theory—for both generic causal nonseparability as well as quantum control of causal orders—with a physically motivated class of free operations, based on process-matrix concatenations. We present the framework explicitly in the mindset with a control register. However, our machinery is totally versatile, being directly applicable also to scenarios with a target register alone. Moreover, an important subclass of our operations is free not only with respect to causal nonseparability and quantum control of causal orders but it also preserves the very causal structure of causal processes. Hence, our treatment contains, as a built-in feature, the basis of a resource theory of quantum causal networks too. As applications, first, we establish a simple sufficient condition for pure-process free convertibility. This imposes a hierarchy of quantum control of causal orders with the quantum switch at the top. Second, we prove that causal-nonseparability distillation exists. More precisely, we show how to convert multiple copies of a process with arbitrarily little causal nonseparability into fewer copies of a quantum switch. Our findings reveal conceptually new, unexpected phenomena, with both fundamental and practical implications.

I. INTRODUCTION

The study of physical processes with events without a predefined, fixed causal order is ultimately motivated by general relativity, whereby the dynamical distribution of energy has a bearing on whether events are time- or spacelike separated. In fact, it has been conjectured [1–3] that quantum gravity may require a theory where a dynamical causal order between events plays an important role. In this context, quantum-mechanical effects on causal orders cannot be disregarded. For instance, this is particularly relevant when one considers the space-time warping caused by spatial quantum superpositions of a massive body [4].

On a more down-to-earth plane, processes with events in indefinite causal orders have sparked a great deal of interest in quantum information and foundations [5,6]. From a fundamental point of view, they constitute an exotic class of quantum operations, adding to the extensive list of counterintuitive properties of quantum theory. This class does not fit the usual quantum-computing paradigm of quantum circuits with fixed gates, and more general frameworks have been developed to encompass it [7–12], such as, e.g., the so-called *process matrices* [9–11]. In general, a process is called *causally nonseparable* if it cannot be decomposed as a classical (i.e., probabilistic) mixture of *causal processes* [9–11] (i.e., processes with a fixed causal order). These processes are fundamentally important since they are suspected to be the key resource for potential experimental violations of causal inequalities [9,13]. A notable subclass of causally nonseparable processes is the one displaying *quantum control of causal orders*, where a quantum system (the control) coherently controls the causal order with which events for another system (the target) take place. The best known example thereof is the celebrated *quantum switch* [5,6,14–16]. The latter is special because it represents the only form of causal nonseparability so far known to be physical [10,17]. In turn, from an applied viewpoint, it has been recently shown to be a useful resource for a number of interesting information-processing tasks [5,14,16,18]. Moreover, it has already been subject of experimental investigations [15,19–21]. Curiously, even though quantum control of causal orders is the rule-of-thumb terminology evoked to discuss the quantum switch, a precise formal definition of this notion is—to our knowledge—still missing.

Here, we study quantum superpositions of causal orders from a resource-theoretic perspective. Resource theories provide powerful frameworks for the formal treatment of a physical property as an operational resource, adequate for

*marciotaddei@gmail.com

its characterization, quantification, and manipulation [22,23]. Their central component is a set of transformations—called the *free operations* of the theory—that are unable to create the resource in question. We build a physically meaningful class of free operations of both causal nonseparability and quantum control of causal orders. This requires a satisfactory rigorous definition of the latter notion, which we provide on the way. The proposed free operations are reminiscent in spirit to the free wirings of other types of quantum resources [24–27]. More precisely, they are given by concatenations of the input process with causally separable processes of two elementary kinds. Processes are mathematically represented by process matrices [9–11] and process concatenations by the so-called *link product* [8]. Both elementary types of process concatenations are remarkably simple and, yet, they give rise to highly nontrivial effects. First, they establish an ordering for a conceptually interesting and experimentally relevant subset of processes to which we refer as *generalized quantum switches*. The ordering is mathematically captured by a simple majorization condition sufficient for a pure process to be freely obtained from another. As a corollary of the latter, it follows that any generalized quantum switch can be freely obtained from the quantum switch. This yields a hierarchy of quantum control of causal orders where the quantum switch sits at the top, thus giving it the status of *basic unit* of this exclusive form of causal nonseparability. Second, we prove that, remarkably, it is possible to concentrate the causal nonseparability spread among multiple copies of nonmaximally causally nonseparable processes (even those arbitrarily close to the causally separable ones) into a quantum switch. Hence, *distillation of quantum control of causal orders* exists. Our proof is constructive, with an explicit distillation protocol, so that a lower bound to the optimal concentration rate is obtained. Finally, we emphasize that our machinery is both highly versatile and notably unifying. On the one hand, it is explicitly formulated in the mindset with a control register but is also readily applicable to scenarios with a target system alone. On the other hand, one of the two elementary types of free operations mentioned leaves invariant not only both the sets of processes without causal nonseparability or quantum control of causal orders but also that of causal processes, for all underlying causal structure. Thus our framework also includes, as a built-in feature, the basis of an eventual resource theory of quantum causal networks.

The paper is organized as follows. In Sec. II, we introduce preliminary concepts and notation. In Sec. III, we propose a formal definition of quantum control of causal orders. In Sec. IV, we introduce our operational framework with the free operations. In Sec. V, we study single-shot conversions, a hierarchy, and units of quantum control of causal orders. In Sec. VI, we show that distillation is possible. Finally, Sec. VII is devoted to our conclusions.

II. PRELIMINARIES

We consider physical processes in the scenario outlined in Fig. 1(a). A convenient tool to describe such processes is the process-matrix formalism [9–11], which extends the quantum combs formalism [8], both in turn based on the Choi-Jamiołkowski (CJ) isomorphism [28,29]. For any Hilbert
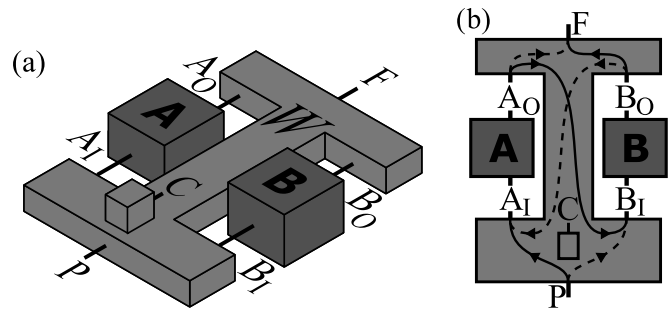

FIG. 1. Schematics of a process. (a) Two users, Alice and Bob, in laboratories A and B, respectively, receive a *target qudit* as local input and subsequently send out an equivalent system as output after some operation. We label Alice's input (output) A_I (A_O), and Bob's B_I (B_O). The events inside each laboratory occur in a definite order: A_I (B_I) happens before A_O (B_O). However, no causal order between A and B is assumed. For practical purposes, but without loss of generality, we do nevertheless assume a past and a future to both A and B. These are realized by laboratories P and F, whose only function is to input into and output from the process the initial and final states of the target qudit, respectively. Finally, there is yet another laboratory, C, operated by a third user, Charlie, who gets as local input a *control qubit*. This can control the causal order between A and B. The process, denoted by W, then specifies the entire system history—preparations, evolutions, measurements, etc.—outside the laboratories (light gray). The operations inside each laboratory, in turn, are described by *instruments* (dark-gray). (b) Pictorial representation of processes with definite causal orders $A \rightarrow B$ (solid) and $B \rightarrow A$ (dashed). Coherent superpositions of the latter give *causally nonseparable* processes. If, in addition, such superpositions involve entanglement with the control qubit, the process can feature *quantum control of causal orders* (see Fig. 2 for precise definition).

space H, we denote by $\mathsf{B}(\mathsf{H})$ the space of bounded-trace, linear operators on H. The CJ isomorphism allows one [8,10] to represent any completely positive trace-preserving linear map $\mathcal{E} : \mathsf{B}(\mathsf{H}_I) \rightarrow \mathsf{B}(\mathsf{H}_O)$ from arbitrary input to output spaces $\mathsf{B}(\mathsf{H}_I)$ and $\mathsf{B}(\mathsf{H}_O)$, respectively, as the *CJ state*

$$E = (\mathcal{I} \otimes \mathcal{E})(|\mathbb{1}\rangle\!\rangle\langle\!\langle\mathbb{1}|). \tag{1}$$

Here, $|\mathbb{1}\rangle\!\rangle := \sum_j |j\rangle \otimes |j\rangle$ is a (non-normalized) maximally entangled state on $\mathsf{H}_I \otimes \mathsf{H}_{I'}$, where $\mathsf{H}_{I'}$ is a space isomorphic to (i.e., a copy of) H_I, with $\{|j\rangle\}_j$ an orthonormal basis of H_I. In turn, $\mathcal{I} : \mathsf{B}(\mathsf{H}_{I'}) \rightarrow \mathsf{B}(\mathsf{H}_{I'})$ is the identity map on the copy space. Complete positivity of $\mathcal{E}$ implies, by virtue of Choi's theorem [28], that E is positive semidefinite. Thus E is technically equivalent to a (non-normalized) state on the extended space $\mathsf{B}(\mathsf{H}_{I'} \otimes \mathsf{H}_O)$. Whenever there is no risk of ambiguity we omit (in a slight abuse of notation) the apostrophe that distinguishes copy from system spaces. For instance, we sometimes write $E \in \mathsf{B}(\mathsf{H}_I \otimes \mathsf{H}_O)$ instead of $E \in \mathsf{B}(\mathsf{H}_{I'} \otimes \mathsf{H}_O)$. In addition, for (non-normalized) maximally entangled states between a system and a copy of it, we directly omit the copy subindex. That is, we use the short-hand notation $|\mathbb{1}\rangle\!\rangle_I$ to denote $|\mathbb{1}\rangle\!\rangle_{II'}$.

In turn, the composition $\mathcal{D} \circ \mathcal{E}$ of $\mathcal{E}$ with another map $\mathcal{D} : \mathsf{B}(\mathsf{H}_O) \rightarrow \mathsf{B}(\mathsf{H}_{\tilde{O}})$ is given in the CJ representation by the *link product* [8], denoted here by "$*$." More precisely, if $D \in \mathsf{B}(\mathsf{H}_O \otimes \mathsf{H}_{\tilde{O}})$ is the CJ state of $\mathcal{D}$, then the CJ state of

$\mathcal{D} \circ \mathcal{E}$ is $D * E \in \mathsf{B}(\mathsf{H}_I \otimes \mathsf{H}_{\tilde{O}})$, defined by

$$D * E := \mathrm{Tr}_O[(E \otimes \mathcal{I}_{\tilde{O}})(\mathcal{I}_I \otimes D^{T_O})]. \tag{2}$$

Here, $\mathcal{I}_{\tilde{O}}$ and $\mathcal{I}_I$ are respectively the identity maps on $\mathsf{H}_{\tilde{O}}$ and H_I, Tr_O denotes the partial trace over H_O, and T_O the partial transpose over H_O (in the chosen basis $\{|j\rangle\}_j$).

Process matrices generalize the notion of CJ states to encapsulate state preparations, operations, and measurements all in a unified description. In our setting, Fig. 1(a), they can be defined by all CJ states that, upon composition with any arbitrary instruments at A and B (including instruments exploiting entangled ancillas between the laboratories), yield a CJ state on the remaining laboratories that describes a valid completely positive (CP) trace-preserving (TP) channel from $\mathsf{B}(\mathsf{H}_P)$ to $\mathsf{B}(\mathsf{H}_F \otimes \mathsf{H}_C)$ [9–11]. This corresponds to CJ states $W \in \mathsf{B}(\mathsf{H}_P \otimes \mathsf{H}_{A_O} \otimes \mathsf{H}_{B_O} \otimes \mathsf{H}_F \otimes \mathsf{H}_{A_I} \otimes \mathsf{H}_{B_I} \otimes \mathsf{H}_C)$ that must be positive semidefinite and satisfy a few normalization constraints (given in Appendix A). If, in addition, a process W has rank 1, it decomposes as $W = |\mathbb{w}\rangle\!\rangle\langle\!\langle\mathbb{w}|$, with $|\mathbb{w}\rangle\!\rangle \in \mathsf{H}_P \otimes \mathsf{H}_{A_O} \otimes \mathsf{H}_{B_O} \otimes \mathsf{H}_F \otimes \mathsf{H}_{A_I} \otimes \mathsf{H}_{B_I} \otimes \mathsf{H}_C$ the corresponding *pure CJ state vector*. In that case, we refer to W as a *pure process* [10] and denote it simply by $|\mathbb{w}\rangle\!\rangle$.[1] We denote the set of generic process matrices for the scenario in question $\mathsf{P} \subset \mathsf{B}(\mathsf{H}_P \otimes \mathsf{H}_{A_O} \otimes \mathsf{H}_{B_O} \otimes \mathsf{H}_F \otimes \mathsf{H}_{A_I} \otimes \mathsf{H}_{B_I} \otimes \mathsf{H}_C)$.

Two basic examples are shown in Fig. 1(b). The first one (solid line) represents processes of the type

$$|0\rangle_C |\mathbb{1}_0\rangle\!\rangle := |0\rangle_C |\mathbb{1}\rangle\!\rangle_{PA_I} |\mathbb{1}\rangle\!\rangle_{A_O B_I} |\mathbb{1}\rangle\!\rangle_{B_O F}, \tag{3}$$

where the subindices in the right-hand side indicate the space supporting each ket. The target-system process $|\mathbb{1}_0\rangle\!\rangle$ defines a quantum causal model [30,31] with causal structure $P \to A \to B \to F$. More precisely, the composite-system process in Eq. (3) describes the situation where Charlie receives the control qubit state $|0\rangle_C$ and the target qudit is directed from P to Alice, who (after applying her instrument) in turn sends it to Bob, who (after his intervention) finally forwards it towards the final target-system output at F. The second process (dashed line) defines a quantum causal model with causal structure $P \to B \to A \to F$ for the target system and gives laboratory C a different local input:

$$|1\rangle_C |\mathbb{1}_1\rangle\!\rangle := |1\rangle_C |\mathbb{1}\rangle\!\rangle_{PB_I} |\mathbb{1}\rangle\!\rangle_{B_O A_I} |\mathbb{1}\rangle\!\rangle_{A_O F}. \tag{4}$$

That is, Charlie now receives the orthogonal state $|1\rangle_C$ while the target now goes from P to B, then to A, and finally to F.

Clearly, $|\mathbb{1}_0\rangle\!\rangle$ and $|\mathbb{1}_1\rangle\!\rangle$ display fixed causal orders between A and B: $A \to B$ and $B \to A$, respectively. They are thus particular instances of *causal processes*. The causal relations between the different laboratories are captured by the signaling constraints of the process [10]. Namely, a process $W_{A \to B}$ is compatible with a causal order $A \to B$ if it is *nonsignaling* from B to A, i.e., if it cannot be used to send information from

Bob's output B_O to Alice's input A_I (see Appendix A for the explicit definition); and analogously for $W_{B \to A}$. In addition, we demand that processes are compatible with the orders $P \to (A, B)$ and $(A, B) \to F$. That is, P and F are respectively taken as the global past and future of the target system (see Appendix A). In turn, a process is said to be *causally separable* [9–11] if it can be decomposed as a probabilistic mixture of causal processes

$$W_{cs} = p\,W_{A \to B} + (1 - p)\,W_{B \to A}, \tag{5}$$

with $0 \leqslant p \leqslant 1$. We denote by $\mathsf{CS} \subset \mathsf{P}$ the set of all causally separable process for our scenario. Any $W \in \mathsf{P} \setminus \mathsf{CS}$ is called *causally nonseparable*.

Causal nonseparability is known to appear in processes that can violate causal inequalities [9,13,32]. These processes involve coherent superpositions of causal orders on the target system alone, i.e., with C playing no role in the causal nonseparability. However, it is not clear whether such processes admit a physical realization [17]. A conceptually different form of causal nonseparability, called *quantum control of causal orders*, takes place when the superposition involves entanglement with C. The *quantum switch* [5,6]

$$|\mathbb{w}_{qs}\rangle\!\rangle := \frac{|0\rangle_C |\mathbb{1}_0\rangle\!\rangle + |1\rangle_C |\mathbb{1}_1\rangle\!\rangle}{\sqrt{2}} \tag{6}$$

is the paradigmatic example thereof. There, C coherently controls the causal order in which the target qudit passes through A and B. Quantum control of causal orders constitutes a stronger form of causal nonseparbility in the sense of requiring not only coherence but also entanglement. Interestingly, in addition, it admits clear physical interpretations in terms of interferometers [15,19–21]. Somewhat surprisingly though, even though the terminology quantum control of causal orders appears quite frequently in the literature, a precise formal definition of this notion has—to our knowledge—not been provided yet. We propose one next.

III. DEFINITION OF QUANTUM CONTROL OF CAUSAL ORDERS

While generic causal nonseparability is a rigorously defined concept, the specific notion of quantum control of causal orders has so far been—surprisingly—only colloquially introduced. Here, we need a precise mathematical definition of this notion. We begin by formalizing the notion of entanglement for processes. This is done in the obvious way, in analogy to entanglement for states [33]. First, for a tripartite process $W_{ABC} \in \mathsf{B}(\mathsf{H}_C \otimes \mathsf{H}_{A_O} \otimes \mathsf{H}_{B_O} \otimes \mathsf{H}_{A_I} \otimes \mathsf{H}_{B_I})$ (without past and future laboratories),[2] we define W_{ABC} to be *separable between control and target* if it belongs to the convex hull of *product processes* in that bipartition, i.e., if

$$W_{ABC} = \sum_\mu q_j\, \varrho_C^{(j)} \otimes W_{AB}^{(j)}, \tag{7}$$

---

[1]There are different definitions of a pure process in the literature. For example, in Ref. [10] the term "pure process" refers to any rank-1 process. In contrast, in Ref. [17], the term "pure process" is reserved to the processes that, upon composition with unitary instruments on A and B, yield a unitary channel from past to future. The two notions are not equivalent but can be connected. We use the convention of Ref. [10]. The term "purification" will also refer to this convention.

[2]Such processes are sometimes referred to as bipartite, to emphasize that there are only two laboratories (Alice's and Bob's) that can display indefinite causal orders. In the operational framework developed here, Charlie is an active user that applies nontrivial operations. Hence, we opt for referring to W_{ABC} as tripartite instead.

with $\{q_j\}_j$ an arbitrary probability distribution over μ, $\varrho_C^{(j)} \in$ $\mathsf{B}(\mathsf{H}_C)$ an arbitrary state of the control, and $W_{AB}^{(j)} \in \mathsf{B}(\mathsf{H}_{A_O} \otimes \mathsf{H}_{B_O} \otimes \mathsf{H}_{A_I} \otimes \mathsf{H}_{B_I})$ an arbitrary process (causally separable or not) for Alice and Bob's laboratories alone. Then, we define a five-partite process $W \in \mathsf{P}$ (with past and future laboratories) to be *separable between the control and the indefinite laboratories* if its reduced process over A, B, and C, given by its partial trace $\mathrm{Tr}_{PF}[W]$ over P and F, is separable between control and target. We denote by $\mathsf{S} \subset \mathsf{P}$ the set of all processes separable between the control and the indefinite laboratories. In turn, any $W \in \mathsf{P} \setminus \mathsf{S}$ is *entangled between the control and the indefinite laboratories*. What is more, here we refer for short to separability or entanglement between the control and the indefinite laboratories simply as *separability or entanglement*, respectively.

The reason why our definition of entanglement focuses on the reductions over A, B, and C is to isolate the entanglement between the control and the target laboratories that can admit indefinite causal orders. Recall that the past and future laboratories have a fixed causal order. In fact, there exist processes in P that are entangled over $\mathsf{B}(\mathsf{H}_C \otimes \mathsf{H}_P \otimes \mathsf{H}_F)$ but separable over $\mathsf{B}(\mathsf{H}_C \otimes \mathsf{H}_{A_O} \otimes \mathsf{H}_{B_O} \otimes \mathsf{H}_{A_I} \otimes \mathsf{H}_{B_I})$, e.g., those corresponding to controlled gates from P to F (with C the control) that act trivially on A and B. Such processes clearly cannot contain quantum control of causal orders. Hence, we exclude them as entangled, so that definition 1 below does not assign them quantum control of causal orders. Moreover, there are important processes without explicit past and future variables, being readily given by tripartite processes on $\mathsf{B}(\mathsf{H}_C \otimes \mathsf{H}_{A_O} \otimes \mathsf{H}_{B_O} \otimes \mathsf{H}_{A_I} \otimes \mathsf{H}_{B_I})$ [10,14–16,19–21]. These processes are readily included in our framework, corresponding to the trivial case of one-dimensional Hilbert spaces H_P and H_F, and our entanglement criterion directly applies there too. Still, entanglement turns out to be necessary but not sufficient for quantum control of causal orders.

Consider for instance the process $|\mathbb{w}_{\mathrm{ent}}\rangle\!\rangle = (|0\rangle_C|\mathbb{1}_0\rangle\!\rangle + |1\rangle_C|\mathbb{u}_{AB}\rangle\!\rangle)/\sqrt{2}$, where $|\mathbb{u}_{AB}\rangle\!\rangle := |\mathbb{1}\rangle\!\rangle_{PA_I} |\mathbb{u}_{AB}\rangle\!\rangle_{A_O B_I} |\mathbb{1}\rangle\!\rangle_{B_O F}$ is a causal process analogous to $|\mathbb{1}_0\rangle\!\rangle$ but with an arbitrary unitary gate $\mathbb{u}_{AB} \neq \mathbb{1}$ from A to B. This can be physically implemented by a quantum circuit with definite causal order $A \rightarrow B$ and controlled unitary gates. Process $|\mathbb{w}_{\mathrm{ent}}\rangle\!\rangle$ is pure and entangled, thus featuring quantum control of unitary gates between A and B. Nevertheless, since both $|\mathbb{1}_0\rangle\!\rangle$ and $|\mathbb{u}_{AB}\rangle\!\rangle$ have causal order $A \rightarrow B$, no control of causal orders takes place. In fact, $|\mathbb{w}_{\mathrm{ent}}\rangle\!\rangle$ is itself causally ordered. The following is a satisfactory definition that rules out such cases.

Definition 1 (Quantum control of causal orders). A process $W \in \mathsf{P}$ has quantum control of causal orders, or, equivalently, is *quantum-control causally ordered*, if it is outside the convex hull $\mathrm{Conv}(\mathsf{CS} \cup \mathsf{S})$ of the sets CS and S of causally separable and separable processes, respectively.

In turn, any $W \in \mathrm{Conv}(\mathsf{CS} \cup \mathsf{S})$ is not quantum-control causally ordered. We denote by $\mathsf{NQC} := \mathrm{Conv}(\mathsf{CS} \cup \mathsf{S})$ the set with no quantum control of causal orders. See Fig. 2.

Definition 1 excludes the convex hull of CS and S (instead of just their union) because convex mixing describes a purely classical operation. In other words, any process that can be operationally generated by probabilistically choosing one out of
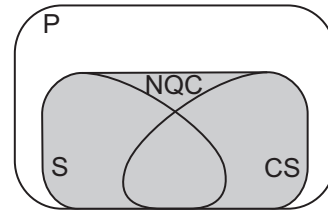

FIG. 2. Pictorial representation of the internal geometry of the set P of all processes. CS and S are the subsets of causally separable and separable processes, respectively. The processes outside CS are causally nonseparable, whereas those outside S are entangled between control and target. The convex hull of CS and S gives the set NQC (gray) with no quantum control of causal orders. This includes processes (given by convex combinations of elements in CS and S) that are causally nonseparable and entangled. We define the processes outside NQC to have quantum control of causal orders.

two resourceless processes must also be resourceless. (Otherwise, probabilistically choosing would not be a free operation of quantum control of causal orders.) This is reminiscent of the definition of genuinely multipartite entanglement, where multipartite states entangled in each and all of the system bipartitions but within the convex hull of the biseparable states are also excluded as genuinely multipartite entangled (see, e.g., Refs. [34,35]).

Notable examples of $\mathsf{P} \setminus \mathsf{NQC}$ are all pure processes

$$|\mathbb{w}\rangle\!\rangle = \sqrt{p_0}\,|\Phi_0\rangle_C|\mathbb{u}_0\rangle\!\rangle + \sqrt{p_1}\,|\Phi_1\rangle_C|\mathbb{u}_1\rangle\!\rangle, \tag{8}$$

with $\{|\Phi_0\rangle, |\Phi_1\rangle\}$ any orthonormal basis of H_C, $\boldsymbol{p} := \{p_0, p_1\}$ any binary probability distribution, and $|\mathbb{u}_0\rangle\!\rangle := |\mathbb{u}_{PA}\rangle\!\rangle_{PA_I} |\mathbb{u}_{AB}\rangle\!\rangle_{A_O B_I} |\mathbb{u}_{BF}\rangle\!\rangle_{B_O F}$ and $|\mathbb{u}_1\rangle\!\rangle := |\mathbb{u}_{PB}\rangle\!\rangle_{PB_I} |\mathbb{u}_{BA}\rangle\!\rangle_{B_O A_I} |\mathbb{u}_{AF}\rangle\!\rangle_{A_O F}$ causal processes analogous to $|\mathbb{1}_0\rangle\!\rangle$ and $|\mathbb{1}_1\rangle\!\rangle$, respectively, but with arbitrary unitary gates $\mathbb{u}_{PA}$, $\mathbb{u}_{AB}$, $\mathbb{u}_{BF}$, $\mathbb{u}_{PB}$, $\mathbb{u}_{BA}$, and $\mathbb{u}_{AF}$ instead of $\mathbb{1}$. That is, $|\mathbb{u}_0\rangle\!\rangle$ and $|\mathbb{u}_1\rangle\!\rangle$ have opposite definite causal orders, similarly to $|\mathbb{1}_0\rangle\!\rangle$ and $|\mathbb{1}_1\rangle\!\rangle$, but with channels other than the identity. These processes capture the most pristine form of causal nonseparability. In fact, for $p_0 = \frac{1}{2} = p_1$, they can be physically realized by applying on $|\mathbb{w}_{\mathrm{qs}}\rangle\!\rangle$ local (i.e., single-lab) unitary transformations on C and nonlocal (i.e., multilab) controlled unitary gates on the target controlled by C. A particular interesting subset of the processes in Eq. (8) is that where the six unitary channels satisfy the constraints

$$\mathbb{u}_{PA}^\dagger \mathbb{u}_{BA} \mathbb{u}_{BF}^\dagger = \mathbb{1} = \mathbb{u}_{PB}^\dagger \mathbb{u}_{AB} \mathbb{u}_{AF}^\dagger. \tag{9}$$

Remarkably, this condition turns out to characterize, for $p_0 = \frac{1}{2} = p_1$, the subset of processes that are local-unitary equivalent to $|\mathbb{w}_{\mathrm{qs}}\rangle\!\rangle$ (see Appendix C for details). We refer to all processes (for arbitrary $\boldsymbol{p}$) satisfying both Eqs. (8) and (9) as *generalized quantum switches*. These are experimentally friendlier than the general processes in Eq. (8) with unconstrained unitaries and will be crucial in Secs. V and VI.

Finally, a comment on the relevance of definition 1 for bipartite processes (i.e., with no control variable) is in place. In Ref. [9], an example of a mixed causally nonseparable bipartite process matrix was discovered. Albeit not in CS, such type of processes can be written as in Eq. (7) with a trivial one-dimensional ρ_C and are hence in S. This is an intended feature of definition 1: any process without a control variable
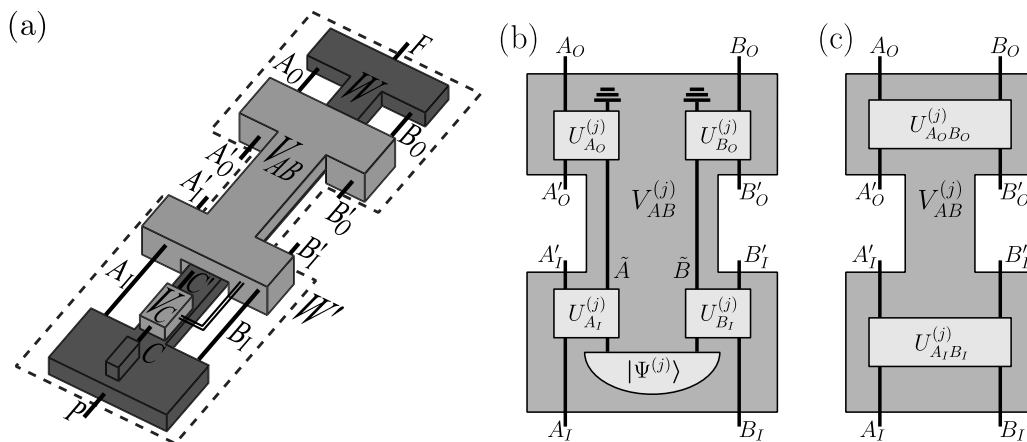
FIG. 3. Schematics of the operational framework of causal nonseparability and quantum control of causal orders. (a) An initial process W (dark gray) is concatenated—i.e., link-product multiplied—with the average process V (light gray) of an instrument at Alice, Bob and Charlie's laboratories, whereas the past and future laboratories are untouched. The final process is thus $\mathcal{V}(W) = V * W$ (dashed outline). This has the same configuration of laboratories as the initial process: P, A', B', C', and F, where A', B', and C' have the same structure of input and output systems as the initial process' laboratories A, B, and C. The instrument's inputs are A_I, A'_O, B_I, B'_O, and C; and its outputs are A'_I, A_O, B'_I, B_O, and C'. All the correlations in the control-versus-target bipartition that V produces are due to 1-way classical communication (double line) of the local instrument outcomes from the control laboratory of Charlie to the target laboratories of Alice and Bob. That is, V is separable in the bipartition and, therefore, creates no entanglement between control and target systems. Furthermore, for each outcome j at C, the outcome process $V_{AB}^{(j)}$ on the target is designed to contain no causal nonseparability either. For this reason, whenever the initial process is causally separable, so is the final one. All this, together with linearity of the link product, implies also that whenever the initial process is not quantum-control causally ordered, neither is the final one. [(b) and (c)] Structure of the instruments on the target register. Each process $V_{AB}^{(j)}$ has six laboratories. Four of them correspond to the inputs A_I and B_I and outputs A_O and B_O of the initial process. While the other two correspond to Alice and Bob's final laboratories, A' and B', both equipped with input and output systems (A'_I and B'_I and A'_O and B'_O, respectively). We consider two elementary types of instruments. (b) The first one gives rise to the class of local operations and ancillary entanglement (LOAE). There, each $V_{AB}^{(j)}$ describes a local (in the $A|B$ bipartition) unitary evolution on the instrument inputs A_I, A'_O, B_I, and B'_O together with (arbitrary-dimensional) ancillary registers $\tilde{A}$ and $\tilde{B}$ in a (possibly entangled) state $|\Psi^{(j)}\rangle_{\tilde{A}\tilde{B}}$, followed by the disposal of the ancilas. (c) In the second one, called probabilistic lab swaps (PLS), each $V_{AB}^{(j)}$ is a pure process where the same unitary operator (either the swap S or the identity $\mathbb{1}$ gate) is applied to A_I together with B_I and to A'_I together with B'_I. That is, conditioned on Charlie's outcome, Alice and Bob either exchange their systems (through swap gates on their inputs and outputs) or leave them untouched. This probabilistically exchanges the causal orders $A \rightarrow B$ and $B \rightarrow A$, but it never introduces causal nonseparability or quantum control of causal orders.

should not have quantum control of causal orders. However, it is also important to note that these processes necessarily display quantum control of causal orders with respect to unaccessible hidden systems (assuming all mixed bipartite processes are the reduction of a pure tripartite one). Being mixed and causally nonseparable, such processes can only admit purifications (i.e., rank-1 process matrices on extended spaces that have the process in question as reduction) that are both causally nonseparable and entangled between the indefinite laboratories and the ancillas used for the Hilbert-space extension. For instance, for the process of [9], a purification can be obtained by adding several single-qubit output spaces (akin to F or C). If one interprets these outputs a control qudit, then such purification displays quantum control of causal orders with respect to them.

IV. THE OPERATIONAL FRAMEWORK

The fundamental property of the free operations of a resource theory is that of mapping the subset of resourceless objects of the theory onto itself. Here we consider linear transformations $\mathcal{V} : \mathsf{P} \rightarrow \mathsf{P}$ such that

$$\mathcal{V}(W) \begin{cases} \in \mathsf{CS} & \text{if } W \in \mathsf{CS}, \\ \in \mathsf{NQC} & \text{if } W \in \mathsf{NQC}. \end{cases} \qquad (10)$$

In other words, we demand that the operations are free with respect to both causal nonseparability and quantum control of causal orders. This may in general be too restrictive if one is only interested in a resource theory of quantum control of causal orders alone. In the end of the section, we mention some subtleties towards such a theory though. In any case, here we are interested in a unified resource theory for both types of resources.

In concrete terms, we propose the following general parametrization for the elementary free operations:

$$\mathcal{V}(W) = V * W. \qquad (11)$$

where V is a (well-normalized) process matrix in $\mathsf{B}(\mathsf{H}_{A_O} \otimes \mathsf{H}_{A'_O} \otimes \mathsf{H}_{B_O} \otimes \mathsf{H}_{B'_O} \otimes \mathsf{H}_{A_I} \otimes \mathsf{H}_{A'_I} \otimes \mathsf{H}_{B_I} \otimes \mathsf{H}_{B'_I} \otimes \mathsf{H}_C \otimes \mathsf{H}_{C'})$. Here, we again explicitly distinguish isomorphic spaces with an apostrophe because the link product in Eq. (11) requires careful space matching. In fact, using Eq. (2), note that $\mathcal{V}$ effectively represents a CP TP map from $\mathsf{B}(\mathsf{H}_{A_I} \otimes \mathsf{H}_{A'_O} \otimes \mathsf{H}_{B_I} \otimes \mathsf{H}_{B'_O} \otimes \mathsf{H}_C)$ to $\mathsf{B}(\mathsf{H}_{A'_I} \otimes \mathsf{H}_{A_O} \otimes \mathsf{H}_{B'_I} \otimes \mathsf{H}_{B_O} \otimes \mathsf{H}_{C'})$, acting trivially on P and F [see Fig. 3(a)]. The explicit form of V is taken as

$$V = \sum_j V_C^{(j)} \otimes V_{AB}^{(j)}, \qquad (12)$$

with $V_C^{(j)} \in \mathsf{B}(\mathsf{H}_C \otimes \mathsf{H}_{C'})$ sub-normalized process matrices (each one representing a CP non-TP map) that sum up to the normalized process matrix $\sum_j V_C^{(j)}$ (representing a CP TP map) and $V_{AB}^{(j)} \in \mathsf{B}(\mathsf{H}_{A_O} \otimes \mathsf{H}_{A'_O} \otimes \mathsf{H}_{B_O} \otimes \mathsf{H}_{B'_O} \otimes \mathsf{H}_{A_I} \otimes \mathsf{H}_{A'_I} \otimes \mathsf{H}_{B_I} \otimes \mathsf{H}_{B'_I})$ normalized process matrices (each one representing a CP TP map).

More technically, each term in Eq. (12) represents the jth outcome of an *instrument* at Charlie's laboratory coordinated with a different process at Alice and Bob's laboratories. Instruments generalize the notion of positive operator-valued measures (POVMs) from measurements to state transformations [36]. They reduce to POVMs when the output space has dimension 1. The above-mentioned coordination is achieved through classical communication of Charlie's outcome j to Alice and Bob. Thus all the free operations arising from Eq. (12) belong to the generic class of local operations and one-way classical communication from the control to the target and are therefore separable in the control-versus-target bipartition. Charlie's instrument can be arbitrary. However, we demand that all instruments at A and B satisfy the following basic constraints to avoid introducing causal loops: laboratories A_I and B_I are jointly in the causal past of A' and B', and all latter four are jointly in the causal past of A_O and B_O. This is mathematically captured by the essential requirement that A' and B' both cannot signal from their local outputs A'_O or B'_O to neither of their inputs A'_I or B'_I. This, together with the fact that V is separable automatically implies that $\mathcal{V}$ preserves the set S of separable processes in the control-versus-target bipartition. Next, we impose more fine-tuned conditions on the instruments at the target laboratories so that $\mathcal{V}$ preserves also the set CS of causally separable processes. Because of linearity, this will automatically imply preservation of NQC too.

Specifically, we consider two broad families of elementary instruments. The first one arises from restricting all $V_{AB}^{(j)}$ to local quantum operations in the $A|B$ bipartition assisted by pre-shared entanglement. More precisely, we take each $V_{AB}^{(j)}$ in Eq. (12) as a process resulting from local unitary dynamics of the instrument inputs A_I, A'_O, B_I, and B'_O together with arbitrary-dimensional ancillary registers $\tilde{A}$ and $\tilde{B}$, which are subsequently discarded [see Fig. 3(b)].[3] The ancillas are initialized in an arbitrary pure (normalized) state $|\Psi^{(j)}\rangle_{\tilde{A}\tilde{B}} \in \mathsf{H}_{\tilde{A}} \otimes \mathsf{H}_{\tilde{B}}$. The evolution is in turn given by local unitary operators $U_{A_I}^{(j)}$ and $U_{B_I}^{(j)}$, from $\mathsf{H}_{A_I} \otimes \mathsf{H}_{\tilde{A}}$ to $\mathsf{H}_{A'_I} \otimes \mathsf{H}_{\tilde{A}}$ and from $\mathsf{H}_{B_I} \otimes \mathsf{H}_{\tilde{B}}$ to $\mathsf{H}_{B'_I} \otimes \mathsf{H}_{\tilde{B}}$, respectively, and $U_{A_O}^{(j)}$ and $U_{B_O}^{(j)}$, from $\mathsf{H}_{A'_O} \otimes \mathsf{H}_{\tilde{A}}$ to $\mathsf{H}_{A_O} \otimes \mathsf{H}_{\tilde{A}}$ and from $\mathsf{H}_{B'_O} \otimes \mathsf{H}_{\tilde{B}}$ to $\mathsf{H}_{B_O} \otimes \mathsf{H}_{\tilde{B}}$, respectively. Finally, after the unitary evolution, both ancillary registers are traced out. We refer to the resulting class as *local operations and ancillary entanglement* (LOAE); and denote it by LOAE:

---

[3] One has in principle three different isomorphic ancilla spaces on Alice's side, one before the application of $U_{A_I}^{(j)}$, one between the application of $U_{A_I}^{(j)}$ and $U_{A_O}^{(j)}$, and one after $U_{A_O}^{(j)}$. Unlike the other variables, however, this ancilla is not matched to any Hilbert space outside $V_{AB}^{(j)}$, so we can use a single variable $\tilde{A}$ for all of them. Same for $\tilde{B}$.

Definition 2 (Local operations and ancillary entanglement). A process transformation $\mathcal{V}$ is in the class LOAE if it can be parametrized by Eqs. (11) and (12) with a process V such that, for all j,

$$V_{AB}^{(j)} := \mathrm{Tr}_{\tilde{A}\tilde{B}}[U^{(j)} |\Psi^{(j)}\rangle \langle\Psi^{(j)}|_{\tilde{A}\tilde{B}} \otimes |\mathbb{1}\rangle\!\rangle \langle\!\langle\mathbb{1}|_{\mathrm{in}} U^{(j)\dagger}], \quad (13)$$

with the short-hand notations $U^{(j)} := (U_{A_O}^{(j)} \otimes \mathbb{1}_{A'_I})(\mathbb{1}_{A'_O} \otimes U_{A_I}^{(j)}) \otimes (U_{B_O}^{(j)} \otimes \mathbb{1}_{B'_I})(\mathbb{1}_{B'_O} \otimes U_{B_I}^{(j)}) \otimes \mathbb{1}_{\mathrm{copy}}$ and $|\mathbb{1}\rangle\!\rangle_{\mathrm{in}} := |\mathbb{1}\rangle\!\rangle_{A_I} \otimes |\mathbb{1}\rangle\!\rangle_{B_I} \otimes |\mathbb{1}\rangle\!\rangle_{A'_O} \otimes |\mathbb{1}\rangle\!\rangle_{B'_O} \in \mathsf{H}_{A_I}^{\otimes 2} \otimes \mathsf{H}_{B_I}^{\otimes 2} \otimes \mathsf{H}_{A'_O}^{\otimes 2} \otimes \mathsf{H}_{B'_O}^{\otimes 2}$, where $\mathbb{1}_{\mathrm{copy}}$ is the identity operator on the inputs' copy's Hilbert space $\mathsf{H}_{A_I} \otimes \mathsf{H}_{B_I} \otimes \mathsf{H}_{A'_O} \otimes \mathsf{H}_{B'_O}$.

We emphasize that only pre-shared quantum correlations between Alice and Bob are allowed in LOAE, with no communication of any sort between them. Thus, clearly, each $V_{AB}^{(j)}$ (and therefore also V) is a nonsignaling process with respect to the $A|B$ bipartition, i.e., nonsignaling both from A to B and vice versa (see lemma 10 in Appendix B for an explicit proof). Explicitly, no information can flow from A_I to B'_I or B_O, from B_I to A'_I or A_O, from A'_O to B_O, and from B'_O to A_O. In particular, this excludes the possibility of teleporting the incoming state of any of the instrument's inputs towards the other side of the bipartition. Finally, note that definition 2 does not impose any restriction on the dimension or structure of the ancillary spaces $\mathsf{H}_{\tilde{A}}$ and $\mathsf{H}_{\tilde{B}}$. Therefore, by virtue of Stinespring's dilation theorem [37], Eq. (13) effectively parametrizes a quantum process describing an arbitrary, fully generic CP TP map without signaling between Alice and Bob and subject to the above-mentioned essential local-causality requirement that A'_O and B'_O cannot signal to A'_I and B'_I, respectively. In fact, arbitrary-dimensional ancillas are actually not required for the latter to hold, just $\dim(\mathsf{H}_{\tilde{A}}) = 2 \dim(\mathsf{H}_{A_I} \otimes \mathsf{H}_{A'_O})$ and $\dim(\mathsf{H}_{\tilde{B}}) = 2 \dim(\mathsf{H}_{B_I} \otimes \mathsf{H}_{B'_O})$.

The second family of elementary instruments we consider is called *probabilistic lab swaps* (PLS), denoted by PLS. It is simpler than the class LOAE in that each $V_{AB}^{(j)}$ in Eq. (12) is a pure process, describing the same unitary transformation applied from $\mathsf{H}_{A_I} \otimes \mathsf{H}_{B_I}$ to $\mathsf{H}_{A'_I} \otimes \mathsf{H}_{B'_I}$ and from $\mathsf{H}_{A'_O} \otimes \mathsf{H}_{B'_O}$ to $\mathsf{H}_{A_O} \otimes \mathsf{H}_{B_O}$. No ancillary registers are used here. Moreover, we allow for only two such unitary operations: the swap gate S and the identity gate $\mathbb{1}$. That is, each process $V_{AB}^{(j)}$ describes either the joint swap of both inputs and outputs, which effectively exchanges Alice and Bob's laboratories, or the trivial identity map.

Definition 3 (Probabilistic lab swaps). A process transformation $\mathcal{V}$ is in the class PLS if it can be parametrized by Eqs. (11) and (12) with a process V such that, for all j, $V_{AB}^{(j)}$ is a rank-1 process given by either the identity $|\mathbb{1}_{AB}\rangle\!\rangle \langle\!\langle\mathbb{1}_{AB}|$ or the lab-swap $|\mathbb{s}_{AB}\rangle\!\rangle \langle\!\langle\mathbb{s}_{AB}|$ processes, defined as

$$|\mathbb{1}_{AB}\rangle\!\rangle := |\mathbb{1}\rangle\!\rangle_{A_I A'_I} \otimes |\mathbb{1}\rangle\!\rangle_{B_I B'_I} \otimes |\mathbb{1}\rangle\!\rangle_{A'_O A_O} \otimes |\mathbb{1}\rangle\!\rangle_{B'_O B_O} \quad (14a)$$

and

$$|\mathbb{s}_{AB}\rangle\!\rangle := |\mathbb{1}\rangle\!\rangle_{A_I B'_I} \otimes |\mathbb{1}\rangle\!\rangle_{B_I A'_I} \otimes |\mathbb{1}\rangle\!\rangle_{A'_O B_O} \otimes |\mathbb{1}\rangle\!\rangle_{B'_O A_O}. \quad (14b)$$

Importantly, due to the swap gates, PLS is not only nonlocal but also even signaling in the $A|B$ bipartition, in contrast to LOAE. In fact, for a causal initial process W, i.e., with a fixed causal order $A \to B$ or $B \to A$, the causal orders are probabilistically exchanged. However, this never creates

causal nonseparability because such exchanges are incoherent. Coherence between the different jth terms in Eq. (12) would be required so that Eqs. (14) can lead to a nonfree operation able to create causal nonseparability. Finally, a comment on the experimental feasibility of PLS is in place. Even though mathematically formulated in terms of joint swaps of both inputs and outputs, process transformations in PLS can in many cases be simulated without any swap gate. More precisely, in experiments, processes are often detected through local instruments at Alice and Bob's laboratory [15,19–21]. Thus, in those cases, instead of actually applying the joint swap gates to their initial process W and detecting their final process $\mathcal{V}(W)$ with local instruments in some given settings, Alice and Bob can simply do nothing to W and swap the settings of their local instruments. That is, the lab swap on the process can be absorbed into the instruments' settings on the final process. This considerably alleviates physical implementations of PLS processes.

The validity of the elementary classes LOAE and PLS as free operations of causal nonseparability is formalized by the following theorem. We take advantage of the theorem also to formally introduce the complete class of free operations we propose: local operations and one-way classical communication from the control to the target given by arbitrary sequential concatenations of transformations in LOAE or PLS.

Theorem 4 (Free operations of causal nonseparability and quantum control of causal orders). Any process transformation $\mathcal{V}$ in LOAE $\cup$ PLS is an automorphism of the sets P of generic processes, CS of causally separable ones, and NQC of non quantum-control causally ordered ones. Therefore, so is any sequence of such elementary transformations.

The theorem is proven in Appendix B. In fact, there we actually prove a stronger result, where LOAE is replaced by the more general class NSO of *nonsignaling operations*. In the latter, instead of preshared entangled ancillas, Alice and Bob may be assisted by generic (potentially supra-quantum) nonsignaling resources. What is more, our proof strategy to show that CS is closed under NSO is to show that even its subsets of causal processes with definite causal orders are preserved as well by NSO. Recall that CS is the convex hull of the latter subsets, so that, by linearity, the implication automatically follows. In other words, we show that any process transformation in NSO (and, by inclusion, also in LOAE) maps an arbitrary causal process, with order either $A \rightarrow B$ or $B \rightarrow A$, into a causal process with the same order. That is, it preserves the underlying causal structure of every quantum causal model. Although this is explicitly proven here for quantum causal models that are effectively bipartite (involving Alice and Bob's laboratories), it can be straightforwardly generalized to arbitrary causal networks with more laboratories. As such, LOAE provides the basis of a yet-to-be resource theory of quantum causal networks, where the resourceful set consists of all quantum causal models incompatible with a given multipartite causal structure under scrutiny. This is a promising and exciting prospect, but it is beyond the scope of this work. Still, the unifying power of the elementary class LOAE could not be left unmentioned here. In the next two sections, we exploit simple examples of our two elementary classes of free operations to implement highly

nontrivial information-theoretic manipulations of causal nonseparability.

Finally, we briefly comment on the possibility of a resource theory of just quantum control of causal orders (and not causal nonseparability). In principle, the condition that CS is closed under the transformations is an unnecessary restriction to that end. However, physically meaningful relaxations of definitions 2 and 3 so that P and NQC are invariant but not CS have been elusive to us. For instance, one could relax the constraint that each $V_{AB}^{(j)}$ in Eq. (12) does not create causal nonseparability, so that—say—$|\mathbb{u}_0\rangle\rangle$ goes to $(|\mathbb{u}_0\rangle\rangle + |\mathbb{u}_1\rangle\rangle)/\sqrt{2}$ for some $\mathbb{u}_{PA}$, $\mathbb{u}_{AB}$, and $\mathbb{u}_{BF}$. The latter corresponds to a coherent lab swap without a control system.[4] However, the same transformation would then map pure entangled processes as $(|0\rangle_C |\mathbb{1}_0\rangle\rangle + |1\rangle_C |\mathbb{u}_0\rangle\rangle)/\sqrt{2} \in$ CS out of NQC. Alternatively, one could even relax the constraint that the instruments are separable in the control-versus-target bipartition [i.e., the tensor-product decomposition of Eq. (12)], so that—say—pure causally separable processes in S are mapped into NQC $\setminus$ S (the set with quantum control of different processes without quantum control of causal orders). The instruments on the target would then be applied coherently with that on the control, instead of conditioned on its classical outcomes. However, similarly to the example above, one can then find pure causally nonseparable processes in S that would be taken out of NQC by the same transformations. We leave the questions of resource theories of quantum control of causal orders that do not preserve CS or S open.

V. SINGLE-COPY CONVERSIONS AND A HIERARCHY OF QUANTUM CONTROL OF CAUSAL ORDERS

Here we study free interconversions between processes in the regime where a single copy of the system is available. (In the next section, we study transformations in the multicopy regime.) More precisely, we consider deterministic conversions between generalized quantum switches, i.e., between any $|\mathbb{w}\rangle\rangle$ and $|\mathbb{w}'\rangle\rangle$ obeying Eq. (8). We characterize the allowed conversions in terms of a majorization relation between the corresponding distributions of $|\mathbb{w}\rangle\rangle$ and $|\mathbb{w}'\rangle\rangle$, respectively denoted by $\boldsymbol{p}$ and $\boldsymbol{p}'$. For binary distributions, majorization is defined in a particularly simple way: $\boldsymbol{p}$ is majorized by $\boldsymbol{p}'$ (denoted $\boldsymbol{p} \preccurlyeq \boldsymbol{p}'$) if $\max_i p_i \leqslant \max_i p'_i$. In other words, $\boldsymbol{p} \preccurlyeq \boldsymbol{p}'$ if $\boldsymbol{p}$ is more flat than $\boldsymbol{p}'$. The characterization is formalized as follows.

Theorem 5 (Single-copy pure-process conversion). Let $|\mathbb{w}\rangle\rangle$ and $|\mathbb{w}'\rangle\rangle$ be generalized quantum switches such that $\boldsymbol{p} \preccurlyeq \boldsymbol{p}'$. Then there is a free operation that converts $|\mathbb{w}\rangle\rangle$ into $|\mathbb{w}'\rangle\rangle$ with unit probability.

The proof is given Appendix C, where we construct an explicit protocol that does the claimed transformation.

Theorem 5 plays a role for quantum control of causal orders similar to the one played in entanglement theory

---

[4]Strictly speaking, it does not define a valid process matrix, as it does not satisfy the necessary normalization conditions. However, it does define a physical process that can be implemented (probabilistically, with a probability depending on the instruments on the target) by post-selecting $|\mathbb{w}_{qs}\rangle\rangle$ on a local measurement on C.

by Nielsen's seminal theorem [38] (see also Ref. [39]) for pure-state conversions under entanglement-free operations. It induces a hierarchy—more precisely, a so-called total preorder—on the set of pure processes obeying Eq. (8). It is called a preorder because there are cases where $|\mathbb{w}\rangle\!\rangle \preccurlyeq |\mathbb{w}'\rangle\!\rangle$ and $|\mathbb{w}'\rangle\!\rangle \preccurlyeq |\mathbb{w}\rangle\!\rangle$ with $|\mathbb{w}'\rangle\!\rangle \neq |\mathbb{w}\rangle\!\rangle$, so that both processes can be reversibly interconverted. This is for instance the case when $|\mathbb{w}\rangle\!\rangle$ and $|\mathbb{w}'\rangle\!\rangle$ are local-unitary equivalent. In turn, such preorder is called total because, for binary probability distributions, there exists no pair of distributions such that none majorizes the other. That is, theorem 5 leaves no pair of generalized quantum switches unconnected.

Hierarchies of this kind are important because they substantiate with a clear operational interpretation the notions of "more" and "less" quantum control of causal orders: if a process can be deterministically transformed freely into another then the former is not less quantum-control causally ordered than the latter. The theorem thus lays the basis of formal quantifiers through causal nonseparability monotones, in the same spirit as entanglement monotones [33]. Interestingly, at the top of the hierarchy lies the quantum switch.

Corollary 6 (Partial unit of quantum control of causal orders). Let $|\mathbb{w}\rangle\!\rangle$ be an arbitrary process given by Eq. (8). Then there is a free operation that converts $|\mathbb{w}_{qs}\rangle\!\rangle$ into $|\mathbb{w}\rangle\!\rangle$ with unit probability.

The corollary follows from the fact that $\boldsymbol{p}_{qs} := \{\frac{1}{2}, \frac{1}{2}\}$ is majorized by all distributions. The basic unit of a resource is important because it renders the notion of maximal amount of resource operationally meaningful, independently of the particular choice of quantifier. For instance, a process would be *the unit of quantum control of causal orders* if all processes could be freely obtained deterministically from it. This would be the counterpart of Bell states in entanglement theory, which are used as entanglement bits [33,40]. Here, we use the terminology *partial unit of quantum control of causal orders* to stress that the quantum switch is the basic unit only within the subset of generalized quantum switches. An interesting possibility would be that $|\mathbb{w}_{qs}\rangle\!\rangle$ can be freely converted probabilistically into all processes (be it exactly or approximately, up to arbitrarily small error). This would render $|\mathbb{w}_{qs}\rangle\!\rangle$ a full unit of causal nonseparability in an operational sense. In fact, invoking again entanglement theory, GHZ states are considered more entangled than W ones precisely in that sense [41,42]. Alternatively, it may as well be the case that there are intrinsically inequivalent classes of causal nonseparability, even under free operations beyond the ones proposed here. These are fascinating open questions that our framework offers for future explorations.

VI. DISTILLATION OF QUANTUM CONTROL OF CAUSAL ORDERS

We now study the concentration of the quantum control of causal orders contained in multiple copies of a process (with nonmaximal resource) into partial units of the resource, i.e., into (fewer) copies of the quantum switch. This is similar in spirit to the notion of entanglement distillation [33,43–45]. Before we proceed, however, a brief digression on the composition of independent copies of a process is useful.

In general, the tensor product of two (or more) valid process matrices on a given system is known not to yield a valid process matrix on the system copies [46,47]. The conceptual reason behind this is that, in the generic situation where Alice and Bob can apply arbitrary instruments globally on the copies of their subsystems, the tensor product of two processes that do not have the same definite causal order renders causal loops possible [46]. This is an expected and reasonable impossibility if a process is used to describe space-time structures [4,48,49], as it is difficult to conceive that Alice and Bob could share two copies of space-time. However, for processes describing, e.g., interferometric experiments [15,19], it is perfectly admissible to describe two independent setups with the tensor product of two process matrices, so long as one restricts the type of instruments on the system copies [46,47]. In fact, this is the most natural description to adopt for experiments. Because, since each laboratory corresponds to a local space-time region, certain configurations of global instruments turn out to be unphysical. For instance, for the above-mentioned processes without the same definite causal order, implementing the instruments that would induce the causal loops requires signaling from one subsystem copy into another towards the past within the same laboratory [47].

Since our focus is operational, we adopt this description here. Indeed, any transformation given by Eq. (11) can be thought of as "adding elements to an experimental setup", as Fig. 3 suggests. Hence, we represent copies of a process with tensor products of it and restrict to independent instruments on each system copy, described in turn by tensor products of single-system instruments. This rules out intercopy signaling, which guarantees non-negative and well-normalized instrument-outcome probabilities. That is, the description is self-consistent and appropriate for operational frameworks.

More technically, we consider the distillation of quantum switches from generic processes $|\mathbb{w}\rangle\!\rangle$ parametrized by Eq. (8), even those arbitrarily close to being causally separable. One says one can distill quantum control of causal orders from $|\mathbb{w}\rangle\!\rangle$, with $p_0 \neq p_1$, if there exists a free operation $\mathcal{V}$ that attains the transformation

$$|\mathbb{w}\rangle\!\rangle^{\otimes N} \xrightarrow{\text{free op}} |\mathbb{w}_{qs}\rangle\!\rangle^{\otimes rN} \qquad (15)$$

with unit probability in the limit $N \to \infty$, for some rate $0 \leqslant r \leqslant 1$. That latter is in turn called the distillation rate of $|\mathbb{w}\rangle\!\rangle$ relative to $\mathcal{V}$. Since we restrict ourselves to independent single-copy instruments, the deterministic multicopy transformation is possible only if it is possible probabilistically on each copy. That is, Eq. (15) is achieved in the asymptotic limit iff $|\mathbb{w}\rangle\!\rangle$ is freely converted into $|\mathbb{w}_{qs}\rangle\!\rangle$ with probability $p_{success} := r$. Such conversion is shown in what follows.

Lemma 7 (Probabilistic single-copy pure-process conversion). Let $|\mathbb{w}\rangle\!\rangle$ and $|\mathbb{w}'\rangle\!\rangle$ be generalized quantum switches such that $\boldsymbol{p} \succcurlyeq \boldsymbol{p}'$. Then there is a free operation that converts $|\mathbb{w}\rangle\!\rangle$ into $|\mathbb{w}'\rangle\!\rangle$ with probability $p_{success} = \frac{\min\{p_0, p_1\}}{\min\{p'_0, p'_1\}}$.

The proof is simple, consisting of a local filtering operation on Charlie's qubit followed by the protocol of theorem 5. It is given explicitly in Appendix D.

Lemma 7 is important because it shows that single-copy process conversions where the final process is majorized by the initial one are also possible (albeit probabilistically).

It thus complements theorem 5 for the deterministic case, possible only when the final process majorizes the initial one. In a sense, it is reminiscent of Vidal's theorem [50] for probabilistic single-copy entanglement conversions between arbitrary pure states. As anticipated, Eq. (15) follows as a corollary of lemma 7. Applying the lemma independently to each copy in $|\mathbb{w}\rangle\!\rangle^{\otimes N}$, taking the limit $N \to \infty$, and using the fact that $\boldsymbol{p}_{\mathrm{qs}} := \{\frac{1}{2}, \frac{1}{2}\}$ proves the main result of this section:

Corollary 8 (Distillation of quantum control of causal orders). Distillation of quantum control of causal orders exists. In fact, a perfect quantum switch can be distilled from any $|\mathbb{w}\rangle\!\rangle$ given by Eq. (8) at a rate $r = 2\min\{p_0, p_1\}$.

The specialized reader may note that the rate in corollary 8 is in general lower than the corresponding optimal entanglement- [33,43–45] and coherence-distillation [51,52] rates for states analogous to the processes in Eq. (8). Yet, in both entanglement and coherence distillation, global operations on each subsystem's copies are exploited, whereas here only independent single-copy instruments are used. Interestingly, it is also possible to distill quantum switches using certain global multicopy instruments: in Appendix D 1, we briefly describe a protocol based on single-copy instruments on the target system conditioned on multiqubit measurements on the copies of Charlie's control. These more general instruments still yield licit free operations—also compatible with tensor products of processes—because they do not involve any intercopy signaling for the laboratories with indefinite causal orders (Alice's and Bob's). As instrument for Charlie, this protocol uses the well-known global measurement from optimal entanglement [44] and coherence [51,52] distillation protocols. However, after the measurement, these protocols require operations whose equivalent here are not free operations. So, the restriction on the target's instruments renders the resulting distillation rate lower than that of corollary 8. Furthermore, one could even explore protocols that exploit intercopy signaling, as certain restricted signaling arrangements still give rise to licit free operations. However, such exploration is outside the scope of the current work, and we leave the question of whether those more powerful free operations do actually yield rates higher than that of corollary 8 as an open problem. In any case, it is remarkable that distilling causal nonseparability is possible at all.

VII. FINAL DISCUSSION

We studied processes displaying quantum coherence between opposite causal orders as an operational resource. In particular, we derived a unified resource theory of both causal nonseparability and quantum control of causal orders. This required a rigorous definition of the latter notion, which, curiously, was still missing. We provided one. Our operational framework is based on resource-free operations consisting of sequential concatenations of the input process with physically meaningful causally separable processes. As applications, first, we established a sufficient condition for pure-process convertibility, mathematically captured by a simple majorization relationship. This orders a broad, important subclass of processes into a hierarchy of quantum control of causal orders with the quantum switch at the top, thus giving the latter the status of basic unit of this exclusive form of causal non-

separability. Second, we proved that distillation of quantum control of causal orders exists, and provided an explicit simple protocol for it. Our machinery is versatile in that it applies to both the mindsets with and without a control register.

As further direct potential applications, we may for instance mention causal-nonseparability measures and a resource theory of quantum causal networks. As for measures, here we have focused on process conversions, but the framework also directly paves the way for quantifiers. From an axiomatic point of view, causal nonseparability monotones can now be defined by any function that is nonincreasing under the free operations proposed. Examples thereof could for instance be the relative entropy and robustness of causal nonseparability or the causally nonseparable weight, which could be defined analogously to in other resource theories [24–27]. From a more operational viewpoint, in turn, corollary 8 gives a lower bound to the distillable causal nonseparability of generalized quantum switches. As for causal networks, notably, our machinery not only treats causal nonseparability and quantum control of causal orders in a unified way but it also contains the basis of an eventual resource theory of quantum causal networks. More precisely, the subclass **LOAE** of local operations assisted by ancillary entanglement preserves the causal structure (either $A \to B$ or $B \to A$) of any quantum causal model for the simplest nontrivial case of two nodes (Alice's and Bob's). Straightforward generalizations of it to more nodes will automatically define free operations for quantum causal networks, where the resourceful objects consist of quantum causal models incompatible with some given multinode causal structure.

Besides, there are several exciting open questions that arise from this work. First, it is not clear whether there exist physical pure processes with quantum control of causal orders (or, more generally, causal nonseparability) apart from those of Eq. (8). Second, if there is no single total unit of causal nonseparability, what other inequivalent classes of causal nonseparability are there? Third, regarding conversions, an important question is whether causal nonseparability dilution—the converse of distillation—is possible or not. Finally, all these questions clearly depend on the class of free operations adopted. Hence, a fourth vast unknown territory is free operations beyond the ones proposed here. In particular, two interesting problems are whether there are free operations of quantum control of causal orders that are not free with respect to causal nonseparability or entanglement or free operations involving intercopy signaling that lead to more efficient distillation protocols than the ones studied here. All of these are fascinating venues for future research.

APPENDIX A: CONDITIONS FOR W AND CAUSAL ORDERS

In order to formally describe many statements in the appendices, we need to define an operation denoted by a subindex preceding an operator [10]:

$$_X W := (1/d_X)\, \mathbb{1}_X \otimes \mathrm{Tr}_X\, W, \tag{A1}$$

where d_X is the dimension of the arbitrary subspace H_X. This operation replaces the original action of W on subspace H_X by a trivial (and fully decorrelated) term.

For W to be a valid process ($W \in \mathsf{P}$), the composition (link product) of $W \in \mathsf{B}(\mathsf{H}_P \otimes \mathsf{H}_{A_O} \otimes \mathsf{H}_{B_O} \otimes \mathsf{H}_F \otimes \mathsf{H}_{A_I} \otimes \mathsf{H}_{B_I} \otimes \mathsf{H}_C)$ with any possible instruments applied by Alice and Bob (even those exploiting entangled ancillas between A and B), must yield a valid CJ state on $\mathsf{B}(\mathsf{H}_P \otimes \mathsf{H}_F \otimes \mathsf{H}_C)$ describing a CP TP global map from $\mathsf{B}(\mathsf{H}_P)$ to $\mathsf{B}(\mathsf{H}_F \otimes \mathsf{H}_C)$ [9–11]. These conditions hold iff W obeys [10,17]

$$W \geqslant 0, \tag{A2a}$$

$$\mathrm{Tr}\, W = d_P d_{A_O} d_{B_O}, \tag{A2b}$$

$$_{B_I B_O CF} W = {}_{A_O B_I B_O CF} W, \tag{A2c}$$

$$_{A_I A_O CF} W = {}_{B_O A_I A_O CF} W, \tag{A2d}$$

$$_{CF} W = {}_{A_O CF} W + {}_{B_O CF} W - {}_{A_O B_O CF} W \tag{A2e}$$

$$_{\boxtimes CF} W = {}_{P \boxtimes CF} W, \tag{A2f}$$

where $\boxtimes := A_I A_O B_I B_O$. We can interpret some of these relations in terms of no-signaling restrictions. Let us take Eq. (A2c) as an example. By stating that $_{B_I B_O CF} W$ is unchanged by replacing A_O with a trivial, decorrelated input, one concludes that A_O may only be correlated with the variables B_I, B_O, C, and F. As such, it cannot signal to any other variables, such as A_I. Since A_I is in the past of A_O and such signaling would yield causal loops, it is only natural that Eq. (A2c) is a necessary condition for the validity of W. In fact, the causal order between A and B is formally defined in terms of such no-signaling restrictions. In general, a process $W_{A \to B}$ is compatible [10] with the causal order $A \to B$ if, and only if,

$$_F W_{A \to B} = {}_{B_O F} W_{A \to B}, \tag{A3a}$$

as obeyed by $|\mathbb{1}_0\rangle\rangle$. This relation only allows signaling from B_O to F, precluding any signaling from B_O to a variable belonging to laboratory A. In other words, this relation precludes signaling from B to A, as expected. Analogously, a process $W_{B \to A}$ is compatible with the causal order $B \to A$ if, and only if, $_F W_{A \to B} = {}_{A_O F} W_{A \to B}$, a relation obeyed by $|\mathbb{1}_1\rangle\rangle$ that forbids signaling from A_O to laboratory B in its causal past.

APPENDIX B: PROOF OF THEOREM 4

After presenting some useful preliminary results, we will break down the proof of theorem 4 in two parts, that of LOAE (which uses the broader class NSO) and that of PLS. After proving that both classes map P, CS, and NQC onto themselves (lemmas 12 and 13, respectively), theorem 4 follows straightforwardly.

1. Useful relations

We will need the following result ("hopping"):

$$\mathrm{Tr}_X\, (_X W\ Y) = \mathrm{Tr}_X\, (W\ {}_X Y), \tag{B1}$$

that is, with respect to the inner product given by trace over (sub)space X, the operation given by subindex $_X$ is self-dual [10]. This is proven by starting from $\mathrm{Tr}_X\, W\ \mathrm{Tr}_X\, Y$ and "factoring out" the second trace operator:

$$\frac{1}{d_X}\mathrm{Tr}_X W\, \mathrm{Tr}_X Y = \mathrm{Tr}_X\left(\frac{\mathbb{1}_X}{d_X}(\mathrm{Tr}_X\, W)Y\right) = \mathrm{Tr}_X({}_X W\ Y). \tag{B2}$$

But the first expression in (B2) is completely symmetric on W, Y, so the same reasoning can be done with the first trace operator, yielding $\mathrm{Tr}_X(W\ {}_X Y)$.

Additionally, given that $\mathcal{V}$ represents a CP TP map from $\mathsf{B}(\mathsf{H}_{A_I} \otimes \mathsf{H}_{A_O'} \otimes \mathsf{H}_{B_I} \otimes \mathsf{H}_{B_O'} \otimes \mathsf{H}_C)$ to $\mathsf{B}(\mathsf{H}_{A_I'} \otimes \mathsf{H}_{A_O} \otimes \mathsf{H}_{B_I'} \otimes \mathsf{H}_{B_O} \otimes \mathsf{H}_{C'})$, any valid V must obey

$$\mathrm{Tr}\, V = d_{A_I} d_{A_O'} d_{B_I} d_{B_O'} d_C, \tag{B3}$$

$$_{A_I' A_O B_I' B_O C'} V = {}_{A_I A_O' B_I B_O' C\ A_I' A_O B_I' B_O C'} V. \tag{B4}$$

Transformations of the form (12) also separately obey

$$_{C'} V = {}_{CC'} V, \tag{B5}$$

$$_{A_I' A_O B_I' B_O} V = {}_{A_I A_O' B_I B_O'\ A_I' A_O B_I' B_O} V. \tag{B6}$$

2. LOAE and NSO

In order to prove that LOAE is a class of free operations, we appeal to a broader class of *nonsignaling operations*, NSO, which forbids signaling from any of the A variables of V (A_I, A_I', A_O', A_O) to any of its B variables (B_I, B_I', B_O', B_O) and vice versa.

Definition 9 (Nonsignaling operations). A process transformation $\mathcal{V}$ belongs to the class NSO if, and only if,

$$_{A_O} V = {}_{A_O' A_O} V, \tag{B7a}$$

$$_{B_O} V = {}_{B_O' B_O} V, \tag{B7b}$$

$$_{A_I' A_O} V = {}_{A_I A_I' A_O} V, \tag{B7c}$$

$$_{B_I' B_O} V = {}_{B_I B_I' B_O} V. \tag{B7d}$$

We notice that Eqs. (B7a) and (B7b) also exclude signaling from A_O' (B_O') to A_I' (B_I'), preventing causal loops. NSO is, in fact, a more general class than LOAE in two ways. First, it need not be separable in the $C|AB$ partition, allowing for coherent operations between C and AB. Secondly, it allows for post-quantum resources, such as using a Popescu-Röhrlich box [53] to correlate outputs. Although mathematically well-defined, not all elements of NSO have a clear quantum-mechanical realization. However, the class LOAE, which has a clear physical interpretation and parametrization, is shown to be a subset of NSO, so that all properties proven for NSO are valid for LOAE.

Lemma 10. LOAE $\subseteq$ NSO.

Proof. We will show that $V_{AB}^{(j)}$ parametrized as in Eq. (13) obeys Eqs. (B7). By linearity, the same will hold for $V = \sum_j V_C^{(j)} \otimes V_{AB}^{(j)}$. We first notice that invariance under the

application of a subindex operator is equivalent to a trivial dependence on the corresponding partition, or $W = {}_X W \Leftrightarrow W \propto \mathbb{1}_X$, i.e., W is a tensor product of $\mathbb{1}_X$ and operators on the space of the remaining variables.

Let us begin by Eq. (B7a). Taking the partial trace $\text{Tr}_{A_O} V_{AB}^{(j)}$ on Eq. (13), the entire output space of $U_{A_O}^{(j)} : \mathsf{H}_{A_O'} \otimes \mathsf{H}_{\tilde{A}} \to \mathsf{H}_{A_O} \otimes \mathsf{H}_{\tilde{A}}$ is traced out. In this case, the basis independence of the trace allows us to replace $U_{A_O}^{(j)}$ for an identity. The action on $\mathsf{H}_{A_O'} \otimes \mathsf{H}_{A_O}$ then reduces to $|\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|_{A_O' A_O}$. Since $\text{Tr}_{A_O}[|\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|_{A_O' A_O}] = \mathbb{1}_{A_O'}$, then $\text{Tr}_{A_O} V_{AB}^{(j)} \propto \mathbb{1}_{A_O'}$, or ${}_{A_O} V_{AB}^{(j)} = {}_{A_O' A_O} V_{AB}^{(j)}$. The demonstration of Eq. (B7b) is analogous, interchanging A and B.

Next, we derive Eq. (B7c) for LOAEs. This time, we take the partial trace $\text{Tr}_{A_I' A_O} V_{AB}^{(j)}$ on Eq. (13), and, analogously, $U_{A_O}^{(j)} \otimes U_{A_I}^{(j)}$ can be replaced by an identity map due to the basis independence of the trace. The action on $\mathsf{H}_{A_I} \otimes \mathsf{H}_{A_I'}$ reduces to $|\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|_{A_I A_I'}$ and since $\text{Tr}_{A_I'}[|\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|_{A_I A_I'}] = \mathbb{1}_{A_I}$, then $\text{Tr}_{A_I' A_O} V_{AB}^{(j)} \propto \mathbb{1}_{A_I}$, or ${}_{A_I' A_O} V_{AB}^{(j)} = {}_{A_I A_I' A_O} V_{AB}^{(j)}$. The demonstration of Eq. (B7d) is analogous, interchanging A and B. ∎

Lemma 11. NSO is a class of free operations of causal nonseparability and of quantum control of causal orders, i.e., it maps P, CS and NQC onto themselves.

Proof. Let us first show that Eqs. (B7) preserve the causal orders $A \to B$ and $B \to A$, i.e., show that

$$ {}_F W = {}_{B_O F} W \Rightarrow {}_F (V * W) = {}_{B_O' F}(V * W), \quad \text{(B8a)} $$

$$ {}_F W = {}_{A_O F} W \Rightarrow {}_F (V * W) = {}_{A_O' F}(V * W). \quad \text{(B8b)} $$

Given that $V * W = \text{Tr}_{\boxtimes C}[W\, V^{T_\boxtimes T_C}]$, where $\boxtimes := A_I A_O B_I B_O$, we prove Eq. (B8a) via the equations indicated in the parentheses and the "hopping" result, Eq. (B1):

$$ {}_F (V * W) = \text{Tr}_{\boxtimes C}[{}_F W\, V^{T_\boxtimes T_C}] \quad \text{(B9a)} $$
$$ \overset{\text{(B8a)}}{=} \text{Tr}_{\boxtimes C}[{}_{B_O F} W\, V^{T_\boxtimes T_C}], $$

$$ {}_F (V * W) \overset{\text{hop}}{=} \text{Tr}_{\boxtimes C}[{}_F W\, {}_{B_O} V^{T_\boxtimes T_C}] \quad \text{(B9b)} $$
$$ \overset{\text{(B7b)}}{=} \text{Tr}_{\boxtimes C}[{}_F W\, {}_{B_O' B_O} V^{T_\boxtimes T_C}] $$
$$ = {}_{B_O'} \text{Tr}_{\boxtimes C}[{}_F W\, {}_{B_O} V^{T_\boxtimes T_C}], $$

$$ \overset{\text{(B9b)}}{=} {}_{B_O' F}(V * W). \quad \text{(B9c)} $$

Eq. (B8b) is proven analogously. As such, from Eq. (5) we see that for NSO preserves CS. Because of the separable structure of Eq. (12), NSO also preserves S. By linearity, $\text{Conv}(\text{CS} \cup \text{S}) = \text{NQC}$ is preserved as well.

We are left with the lengthier task of showing that NSO preserves P, i.e., showing that the validity constraints of Eq. (A2) are preserved under NSO. The positivity constraint (A2a) is straightforward, since the link product preserves positivity.

The dimensionality constraint (A2b) will initially be shown to be preserved by the simpler case of W in a causal order $A \to B$ (${}_F W = {}_{B_O F} W$). We calculate ${}_{PF\boxtimes'C'}(V * W)$, where

$\boxtimes' := A_I' A_O' B_I' B_O'$, from which the trace can be taken:

$${}_{PF\boxtimes'C'} \text{Tr}_{\boxtimes C}(WV^{T_\boxtimes T_C})$$
$$ = {}_P \text{Tr}_{\boxtimes C}({}_F W\, {}_{\boxtimes'C'} V^{T_\boxtimes T_C}) \quad \text{(B10a)} $$
$$ \overset{\text{(B5)}}{=} {}_P \text{Tr}_{\boxtimes C}({}_{B_O F} W\, {}_{C\boxtimes'C'} V^{T_\boxtimes T_C}) \quad \text{(B10b)} $$
$$ \overset{\text{hop}}{=} {}_P \text{Tr}_{\boxtimes C}({}_{CF} W\, {}_{B_O \boxtimes'C'} V^{T_\boxtimes T_C}) \quad \text{(B10c)} $$
$$ \overset{\text{(B7d)}}{=} {}_P \text{Tr}_{\boxtimes C}({}_{CF} W\, {}_{B_I B_O \boxtimes'C'} V^{T_\boxtimes T_C}) \quad \text{(B10d)} $$
$$ \overset{\text{hop}}{=} {}_P \text{Tr}_{\boxtimes C}({}_{B_I B_O CF} W\, {}_{\boxtimes'C'} V^{T_\boxtimes T_C}) \quad \text{(B10e)} $$
$$ \overset{\text{(A2c)}}{=} {}_P \text{Tr}_{\boxtimes C}({}_{A_O B_I B_O CF} W\, {}_{\boxtimes'C'} V^{T_\boxtimes T_C}) \quad \text{(B10f)} $$
$$ \overset{\text{hop}}{=} {}_P \text{Tr}_{\boxtimes C}({}_{B_I B_O CF} W\, {}_{A_O \boxtimes'C'} V^{T_\boxtimes T_C}) \quad \text{(B10g)} $$
$$ \overset{\text{(B7c)}}{=} {}_P \text{Tr}_{\boxtimes C}({}_{B_I B_O CF} W\, {}_{A_I A_O \boxtimes'C'} V^{T_\boxtimes T_C}) \quad \text{(B10h)} $$
$$ \overset{\text{hop}}{=} {}_P \text{Tr}_{\boxtimes C}({}_F W\, {}_{\boxtimes C \boxtimes'C'} V^{T_\boxtimes T_C}) \quad \text{(B10i)} $$
$$ = {}_{PF\boxtimes'C'} \text{Tr}_{\boxtimes C}(W\, {}_{\boxtimes C} V^{T_\boxtimes T_C}). \quad \text{(B10j)} $$

Taking the trace of this expression, using the definition (A1) and the fact that the subindex operator is TP, we find

$$ \text{Tr}(V * W) = \text{Tr}_{PF\boxtimes\boxtimes'CC'} \left(W\, \frac{\mathbb{1}_\boxtimes \mathbb{1}_C}{d_\boxtimes d_C} \otimes \text{Tr}_{\boxtimes C} V^{T_\boxtimes T_C} \right) \quad \text{(B10k)} $$

$$ = (\text{Tr}_{PF\boxtimes C} W)\frac{1}{d_\boxtimes d_C}(\text{Tr}_{\boxtimes\boxtimes'CC'} V^{T_\boxtimes T_C}) \quad \text{(B10l)} $$

$$ = d_P d_{A_O} d_{B_O} \frac{1}{d_\boxtimes d_C} d_{A_O'} d_{B_O'} d_{A_I} d_{B_I} d_C, \quad \text{(B10m)} $$

where in the last line Eqs. (A2b) and (B3) were used. We then obtain $\text{Tr}(V * W) = d_P d_{A_O'} d_{B_O'}$, as desired. If W is not in the causal order $A \to B$, the demonstration changes as follows: instead of Eq. (B10b), we obtain a term identical to Eq. (B10c) but without the B_O subscript. Using Eq. (A2e),

$$ {}_{PF\boxtimes'C'}(V * W) = {}_P \text{Tr}_{\boxtimes C}({}_{A_O CF} W\, {}_{\boxtimes'C'} V^{T_\boxtimes T_C}) $$
$$ + {}_P \text{Tr}_{\boxtimes C}({}_{B_O CF} W\, {}_{\boxtimes'C'} V^{T_\boxtimes T_C}) $$
$$ - {}_P \text{Tr}_{\boxtimes C}({}_{A_O B_O CF} W\, {}_{\boxtimes'C'} V^{T_\boxtimes T_C}). \quad \text{(B11)} $$

For the first term on the right-hand side, the demonstration can be carried out as above, switching the roles of A and B. For the second and third, all calculations on Eqs. (B10) are valid. As such, all three terms are equal to $d_P d_{A_O'} d_{B_O'}$. Given their signs, we obtain $\text{Tr}(V * W) = d_P d_{A_O'} d_{B_O'}$, as before.

Let us now prove that Eq. (A2c) is preserved. Once again we begin by assuming ${}_F W = {}_{B_O F} W$ and afterwards lift that assumption. Firstly, ${}_{B_I' B_O' C' F}(V * W)$ can be shown to equal ${}_{B_I' B_O' C'} \text{Tr}_{\boxtimes C}({}_{A_O B_I B_O CF} W\, V^{T_\boxtimes T_C})$ using Eqs. (B7d) and (A2c) as done in Eqs. (B10a) and (B10f). So we can write ${}_{B_I' B_O' C' F}(V * W)$ as

$$ {}_{B_I' B_O' C'} \text{Tr}_{\boxtimes C}({}_{A_O B_I B_O CF} W V^{T_\boxtimes T_C}) \quad \text{(B12a)} $$
$$ \overset{\text{hop}}{=} {}_{B_I' B_O' C'} \text{Tr}_{\boxtimes C}({}_{B_I B_O CF} W\, {}_{A_O} V^{T_\boxtimes T_C}) \quad \text{(B12b)} $$
$$ \overset{\text{(B7a)}}{=} {}_{B_I' B_O' C'} \text{Tr}_{\boxtimes C}({}_{B_I B_O CF} W\, {}_{A_O' A_O} V^{T_\boxtimes T_C}) \quad \text{(B12c)} $$
$$ = {}_{A_O' B_I' B_O' C'} \text{Tr}_{\boxtimes C}({}_{B_I B_O CF} W_{A_O} V^{T_\boxtimes T_C}) \quad \text{(B12d)} $$

and comparing Eq. (B12d) with Eq. (B12b), we find that $_{B'_I B'_O C' F}(V * W) = _{A'_O B'_I B'_O C' F}(V * W)$, as desired. If W is not compatible with $A \to B$, we arrive via Eqs. (A2e) and (B5) at the three-term expression

$$
\begin{aligned}
_{C'F}(V * W) &= _{C'} \mathrm{Tr}_{\boxtimes C}(_{A_O CF} W \, V^{T \boxtimes T_C}) \\
&+ _{C'} \mathrm{Tr}_{\boxtimes C}(_{B_O CF} W \, V^{T \boxtimes T_C}) \\
&- _{C'} \mathrm{Tr}_{\boxtimes C}(_{A_O B_O CF} W \, V^{T \boxtimes T_C}). \quad \text{(B13)}
\end{aligned}
$$

The steps in Eqs. (B10a) and (B10f) apply to the second term on the right-hand side. To the first and third terms on the right-hand side, we can directly apply the steps in Eqs. (B12). All three terms, then, equal $_{A'_O B'_I B'_O F}(W * V)$ and, due to their signs, $_{B'_I B'_O F}(W * V) = _{A'_O B'_I B'_O F}(W * V)$ in general. The demonstration that condition (A2d) is preserved follows analogously, switching A and B throughout.

The preservation of condition (A2e) is demonstrated as follows. First let us notice that $_{C'F}(V * W) = (_{C'} V *_F W) = (_{C'} V *_{CF} W) = _{C'}(V *_{CF} W)$, then

$$
\begin{aligned}
&_{A'_O C' F}(V * W) + _{B'_O C' F}(V * W) - _{A'_O B'_O C' F}(V * W) \\
&= _{C'} \mathrm{Tr}_{\boxtimes C} \big[(_{A_O CF} W + _{B_O CF} W - _{A_O B_O CF} W) \\
&\quad \times (_{A'_O} V + _{B'_O} V - _{A'_O B'_O} V)^{T \boxtimes T_C} \big], \quad \text{(B14a)}
\end{aligned}
$$

where Eq. (A2e) has been applied to W. On the nine resulting terms we apply Eqs. (B1), (B7a), and (B7b) to eliminate A'_O, B'_O whenever possible, and we see that six of these terms cancel out, leading to

$$
\begin{aligned}
&_{A'_O C' F}(V * W) + _{B'_O C' F}(V * W) - _{A'_O B'_O C' F}(V * W) \\
&= _{C'} \mathrm{Tr}_{\boxtimes C} \big[_{CF} W \, (_{A_O} V + _{B_O} V - _{A_O B_O} V)^{T \boxtimes T_C} \big] \quad \text{(B14b)} \\
&\overset{\text{hop}}{=} _{C'} \mathrm{Tr}_{\boxtimes C} \big[(_{A_O CF} W + _{B_O CF} W - _{A_O B_O CF} W) \, V^{T \boxtimes T_C} \big] \\
&\overset{\text{(A2e)}}{=} _{C'}(V *_{CF} W) = _{C'F}(V * W). \quad \text{(B14c)}
\end{aligned}
$$

Finally, to prove that condition (A2f) is preserved, we once again begin by assuming W compatible with $A \to B$ ($_F W = _{B_O F} W$) and later lift the assumption:

$$
\begin{aligned}
_{\boxtimes' C' F}(V * W) &= \mathrm{Tr}_{\boxtimes C}(_F W \, _{\boxtimes' C'} V^{T \boxtimes T_C}) \quad \text{(B15a)} \\
&\overset{\text{(B5)}}{=} \mathrm{Tr}_{\boxtimes C}(_{B_O F} W \, _{C \boxtimes' C'} V^{T \boxtimes T_C}) \quad \text{(B15b)} \\
&\overset{\text{hop}}{=} \mathrm{Tr}_{\boxtimes C}(_{CF} W \, _{B_O \boxtimes' C'} V^{T \boxtimes T_C}) \quad \text{(B15c)} \\
&\overset{\text{(B7d)}}{=} \mathrm{Tr}_{\boxtimes C}(_{CF} W \, _{B_I B_O \boxtimes' C'} V^{T \boxtimes T_C}) \quad \text{(B15d)} \\
&\overset{\text{hop}}{=} \mathrm{Tr}_{\boxtimes C}(_{B_I B_O CF} W \, _{\boxtimes' C'} V^{T \boxtimes T_C}) \quad \text{(B15e)} \\
&\overset{\text{(A2c)}}{=} \mathrm{Tr}_{\boxtimes C}(_{A_O B_I B_O CF} W \, _{\boxtimes' C'} V^{T \boxtimes T_C}) \quad \text{(B15f)} \\
&\overset{\text{hop}}{=} \mathrm{Tr}_{\boxtimes C}(_{B_I B_O CF} W \, _{A_O \boxtimes' C'} V^{T \boxtimes T_C}) \quad \text{(B15g)} \\
&\overset{\text{(B7c)}}{=} \mathrm{Tr}_{\boxtimes C}(_{B_I B_O CF} W \, _{A_I A_O \boxtimes' C'} V^{T \boxtimes T_C}) \quad \text{(B15h)} \\
&\overset{\text{hop}}{=} \mathrm{Tr}_{\boxtimes C}(_{\boxtimes CF} W \, _{\boxtimes' C'} V^{T \boxtimes T_C}) \quad \text{(B15i)} \\
&\overset{\text{(A2f)}}{=} \mathrm{Tr}_{\boxtimes C}(_{P \boxtimes CF} W \, _{\boxtimes' C'} V^{T \boxtimes T_C}) \quad \text{(B15j)} \\
&= _{P \boxtimes' C' F}(V * W). \quad \text{(B15k)}
\end{aligned}
$$

If W is not compatible with $A \to B$, we have, instead of Eq. (B15b), the three-term expression [due to Eq. (A2e)]

$$
\begin{aligned}
_{C'F}(V * W) &= _{C'} \mathrm{Tr}_{\boxtimes C}(_{A_O CF} W \, _{\boxtimes'} V^{T \boxtimes T_C}) \\
&+ _{C'} \mathrm{Tr}_{\boxtimes C}(_{B_O CF} W \, _{\boxtimes'} V^{T \boxtimes T_C}) \\
&- _{C'} \mathrm{Tr}_{\boxtimes C}(_{A_O B_O CF} W \, _{\boxtimes'} V^{T \boxtimes T_C}). \quad \text{(B15l)}
\end{aligned}
$$

The calculation above can be done directly on the second term on the right-hand side and is also valid on the third. To the first term on the right-hand side, we can apply the same steps as above, but switching A and B throughout. All three terms, then, equal $_{P \boxtimes' C' F}(V * W)$ and, due to their signs, $_{\boxtimes' C' F}(V * W) = _{P \boxtimes' C' F}(V * W)$.

We have then showed that NSO preserves P, along with CS and NQC. ∎

As a straightforward consequence of lemmas 10 and 11,

Corollary 12. LOAE is a class of free operations of causal nonseparability and of quantum control of causal orders, i.e., it maps P, CS and NQC onto themselves.

3. Probabilistic lab swaps

Lemma 13. PLS is a class of free operations of causal nonseparability and of quantum control of causal orders, i.e., it maps P, CS and NQC onto themselves.

Proof. We begin by noticing that from Eqs.(14a), (B3), (B4), and (B7) that $|\mathbb{1}_{AB}\rangle\!\rangle\langle\!\langle \mathbb{1}_{AB}| \in$ NSO, and hence preserves P, CS. On the other hand, $V_{AB}^{(\mathrm{sw})} := |\mathbb{s}_{AB}\rangle\!\rangle\langle\!\langle \mathbb{s}_{AB}|$ from Eq. (14b) obeys the following signaling conditions [compare Eqs. (B7)]:

$$
_{A_O} V_{AB}^{(\mathrm{sw})} = _{B'_O A_O} V_{AB}^{(\mathrm{sw})}, \quad \text{(B16a)}
$$

$$
_{B_O} V_{AB}^{(\mathrm{sw})} = _{A'_O B_O} V_{AB}^{(\mathrm{sw})}, \quad \text{(B16b)}
$$

$$
_{B'_I A_O} V_{AB}^{(\mathrm{sw})} = _{A_I B'_I A_O} V_{AB}^{(\mathrm{sw})}, \quad \text{(B16c)}
$$

$$
_{A'_I B_O} V_{AB}^{(\mathrm{sw})} = _{B_I A'_I B_O} V_{AB}^{(\mathrm{sw})}. \quad \text{(B16d)}
$$

As expected, $V_{AB}^{(\mathrm{sw})}$ inverts the ordering $A \to B$ and $B \to A$, i.e., from Eqs. (B16) it follows that

$$
_F W = _{B_O F} W \Rightarrow _F \big(V_{AB}^{(\mathrm{sw})} * W\big) = _{A'_O F}\big(V_{AB}^{(\mathrm{sw})} * W\big), \quad \text{(B17a)}
$$

$$
_F W = _{A_O F} W \Rightarrow _F \big(V_{AB}^{(\mathrm{sw})} * W\big) = _{B'_O F}\big(V_{AB}^{(\mathrm{sw})} * W\big), \quad \text{(B17b)}
$$

which can be shown following the steps in Eq. (B9) with Eq. (B16) instead of Eq. (B7). Most importantly, although the causal order is inverted, the existence of a well-defined causal order is preserved when $V_{AB}^{(\mathrm{sw})}$ alone is applied, and so is CS. The preservation of P [Eqs. (A2)] by $V_{AB}^{(\mathrm{sw})}$ is demonstrated analogously as shown above for NSO, switching $A'_I \leftrightarrow B'_I$, $A'_O \leftrightarrow B'_O$.

Given that both $|\mathbb{1}_{AB}\rangle\!\rangle\langle\!\langle \mathbb{1}_{AB}|$ and $|\mathbb{s}_{AB}\rangle\!\rangle\langle\!\langle \mathbb{s}_{AB}|$ preserve CS and P, so does a probabilistic lab swap, since convex mixtures preserve Eqs. (5) and (A2). The separable form of Eq.(12) guarantees preservation of S, hence PLS preserves P, CS, and NQC. ∎

APPENDIX C: PROOF OF THEOREM 5

We prove theorem 5 constructively, presenting a protocol that transforms $|\mathbb{w}\rangle\!\rangle$ into $|\mathbb{w}'\rangle\!\rangle$ [both generalized quantum

switches obeying Eqs. (8) and (9)], which can be decomposed into four steps. (1) Apply certain unitaries to the laboratories inputs and outputs. (2) Apply a unitary on the control qubit mapping $\{|\Phi_0\rangle_C, |\Phi_1\rangle_C\}$ into $\{|\Phi'_0\rangle_C, |\Phi'_1\rangle_C\}$. (3) Make a non-demolition measurement on the control qubit to skew $\boldsymbol{p}$ into $\boldsymbol{p}'_{\mathrm{id}} = \boldsymbol{p}' = \{p'_0, p'_1\}$. (4) This measurement may incorrectly turn $\boldsymbol{p}$ into $\boldsymbol{p}'_{\mathrm{sw}} := \{p'_1, p'_0\}$ instead. However, this is heralded by the measurement outcome, conditioned on which a correction is applied: the control qubit is flipped and the laboratories are swapped.

Proof. For step C, we apply

$$|\mathrm{u}'_{PA}\mathrm{u}^\dagger_{PA}\rangle\!\rangle_{A_I A'_I} \, |\mathrm{u}'_{PB}\mathrm{u}^\dagger_{PB}\rangle\!\rangle_{B_I B'_I}$$
$$\otimes |\mathrm{u}^\dagger_{AF}\mathrm{u}'_{AF}\rangle\!\rangle_{A'_O A_O} \, |\mathrm{u}^\dagger_{BF}\mathrm{u}'_{BF}\rangle\!\rangle_{B'_O B_O}, \tag{C1}$$

transforming u_0, u_1 into u'_0, u'_1, where $\mathrm{u}'^\dagger_{PA}\mathrm{u}'_{BA}\mathrm{u}'^\dagger_{BF} = \mathbb{1} = \mathrm{u}'^\dagger_{PB}\mathrm{u}'_{AB}\mathrm{u}'^\dagger_{AF}$ has been used. These constraints reflect the fact that the freedom to pick four local unitaries is not sufficient to attain all six unitaries in Eq. (8). In fact, the reader can verify that, starting from a superposition of $|\Phi_0\rangle_C|\mathbb{1}_0\rangle\!\rangle$ and $|\Phi_1\rangle_C|\mathbb{1}_1\rangle\!\rangle$, Eq. (9) is necessary and sufficient for the target process to be attainable by local unitaries.

Steps C–C can be encapsulated in a single transformation. The majorization relation $\boldsymbol{p} \preccurlyeq \boldsymbol{p}'$ implies [52,54]

$$\boldsymbol{p} = \sum_\pi \lambda_\pi \boldsymbol{p}'_\pi, \tag{C2}$$

where $\boldsymbol{p}'_\pi$ is the πth permutation of $\boldsymbol{p}'$, with $\pi = \mathrm{id}$ representing the identity and $\pi = \mathrm{sw}$ the swap, and $\boldsymbol{\lambda} := \{\lambda_\pi\}_\pi$ is a probability distribution on π. The transformation will be decomposed as in Eq. (12) (for π playing the role of j) with $V_C^{(\pi)} := |\mathrm{v}_C^{(\pi)}\rangle\!\rangle\langle\!\langle\mathrm{v}_C^{(\pi)}|$, being

$$\big|\mathrm{v}_C^{(\pi)}\big\rangle\!\rangle = \sqrt{\lambda_\pi} \sum_i \sqrt{\frac{p'_{\pi(i)}}{p_i}} |\Phi_i\rangle_C |\Phi'_{\pi(i)}\rangle_{C'}, \tag{C3}$$

and

$$V_{AB}^{(\mathrm{id})} := |\mathbb{1}_{AB}\rangle\!\rangle\langle\!\langle\mathbb{1}_{AB}|, \quad V_{AB}^{(\mathrm{sw})} := |\mathbb{s}_{AB}\rangle\!\rangle\langle\!\langle\mathbb{s}_{AB}|, \tag{C4}$$

for $|\mathbb{1}_{AB}\rangle\!\rangle$ and $|\mathbb{s}_{AB}\rangle\!\rangle$ given by Eq. (14). In addition, the short-hand notation $\pi(i) := i$, for $\pi = \mathrm{id}$, and $\pi(i) := i+1$ mod 2, for $\pi = \mathrm{sw}$, has been used. The unitary in step C always exists because $\{|\Phi_j\rangle\}_{j=0,1}$ and $\{|\Phi'_j\rangle\}_{j=0,1}$ are orthonormal bases. The measurement in step C is a two-outcome POVM on H_C that skews the flatter $\boldsymbol{p}$ distribution into either $\boldsymbol{p}'$ or $\boldsymbol{p}'_{\mathrm{sw}}$ (less balanced) while preserving the quantum superposition [the general form of such POVM is shown in Eq. (D1)]. The conditional lab swap (step C) belongs to PLS and occurs together with a simple bit-flip on the control qubit in the $\{|\Phi'_j\rangle\}_{j=0,1}$ basis. Normalization follows from the majorization relation (C2). Applied to $|\mathrm{w}\rangle\!\rangle$, the outcome for each π is pure:

$$\pi = \mathrm{id}, \quad \sqrt{\lambda_{\mathrm{id}}}\big(\sqrt{p'_0}|\Phi'_0\rangle|\mathrm{u}'_0\rangle\!\rangle + \sqrt{p'_1}|\Phi'_1\rangle|\mathrm{u}'_1\rangle\!\rangle\big), \tag{C5}$$

$$\pi = \mathrm{sw}, \quad \sqrt{\lambda_{\mathrm{sw}}}\big(\sqrt{p'_1}|\Phi'_1\rangle|\mathrm{u}'_1\rangle\!\rangle + \sqrt{p'_0}|\Phi'_0\rangle|\mathrm{u}'_0\rangle\!\rangle\big), \tag{C6}$$

where we have used $|\mathrm{w}_{0,1}\rangle\!\rangle$ from Eqs. (3) and (4). Both results are proportional to $|\mathrm{w}'\rangle\!\rangle$. A sum over π, with $\lambda_{\mathrm{id}} + \lambda_{\mathrm{sw}} = 1$, yields this process exactly. Except for the lab

swap which belongs to PLS, all transformations belong to LOAE. ∎

APPENDIX D: PROBABILISTIC CONVERSION AND DISTILLATION

The distillation protocol hinges on lemma 7, proven here by construction.

Proof. The first step is to convert $|\mathrm{w}\rangle\!\rangle$ into a pure process $|\mathrm{w}_{\mathrm{aux}}\rangle\!\rangle$ with $\boldsymbol{p}_{\mathrm{aux}} \preccurlyeq \boldsymbol{p}'$ (for best success probability, we also have $\boldsymbol{p}_{\mathrm{aux}} \succcurlyeq \boldsymbol{p}'$—whether $\boldsymbol{p}_{\mathrm{aux}} = \boldsymbol{p}'$ or $\boldsymbol{p}_{\mathrm{aux}} = \boldsymbol{p}'_{\mathrm{sw}}$ depends on the specific processes $|\mathrm{w}\rangle\!\rangle$ and $|\mathrm{w}'\rangle\!\rangle$). This is achieved with a simple local-filtering measurement on Charlie's qubit, which can be written formally as the POVM $\{M_C^{(j)}(x,y)\}_{j=0,1} = \{|\mathrm{m}_C^{(j)}\rangle\!\rangle\langle\!\langle\mathrm{m}_C^{(j)}|\}_{j=0,1}$, where

$$\big|\mathrm{m}_C^{(0)}\big\rangle\!\rangle = \sqrt{x}|\Phi_0\rangle_C|\Phi_0\rangle_{C'} + \sqrt{y}|\Phi_1\rangle_C|\Phi_1\rangle_{C'},$$
$$\big|\mathrm{m}_C^{(1)}\big\rangle\!\rangle = \sqrt{1-x}|\Phi_0\rangle_C|\Phi_0\rangle_{C'} + \sqrt{1-y}|\Phi_1\rangle_C|\Phi_1\rangle_{C'}, \tag{D1}$$

with

$$x = \min\left\{\frac{p_1}{p_0}\frac{\max\{p'_0, p'_1\}}{\min\{p'_0, p'_1\}}, 1\right\}, \tag{D2}$$

$$y = \min\left\{\frac{p_0}{p_1}\frac{\max\{p'_0, p'_1\}}{\min\{p'_0, p'_1\}}, 1\right\}. \tag{D3}$$

Outcome $j = 0$, which corresponds to a rank-2 POVM element and occurs with probability $p_{\mathrm{success}} = \frac{\min\{p_0, p_1\}}{\min\{p'_0, p'_1\}}$, leads to $|\mathrm{w}_{\mathrm{aux}}\rangle\!\rangle$ as desired. Outcome $j = 1$, which necessarily corresponds to a rank-1 POVM element, indicates a failed result of the local filtering. The success probability of this step (and of the overall procedure) is $p_{\mathrm{success}} = \frac{\min\{p_0, p_1\}}{\min\{p'_0, p'_1\}}$. If successful, on the process $|\mathrm{w}_{\mathrm{aux}}\rangle\!\rangle$ we apply the deterministic protocol from theorem 5 to produce $|\mathrm{w}'\rangle\!\rangle$. ∎

To prove corollary 8, we need to apply this probabilistic protocol, with $|\mathrm{w}_{\mathrm{qs}}\rangle\!\rangle$ as target, onto N copies of $|\mathrm{w}\rangle\!\rangle$ in parallel. In the asymptotic limit of $N \to \infty$ [55,56], the result, with probability tending to one, is to have $N\,p_{\mathrm{success}}$ copies of $|\mathrm{w}_{\mathrm{qs}}\rangle\!\rangle$. Since the success/failure is heralded, one can pick the successful copies, distilling causal nonseparability with rate $r = p_{\mathrm{success}} = 2\min\{p_0, p_1\}$, where we have used that $\boldsymbol{p}_{\mathrm{qs}} := \{\frac{1}{2}, \frac{1}{2}\}$.

1. Distillation with multicopy instruments

We now present an alternative distillation protocol making use of multicopy operations, inspired on the coherence-distillation protocol developed in Refs. [51,52]. Because of the limitations to act jointly on different copies of Alice's and Bob's laboratories, joint operations are used only on Charlie's control qudit. This is capable of obtaining distillation, albeit at a rate $r = \min\{p_0, p_1\}$, lower than that in corollary 8. This protocol distills $|\mathrm{w}\rangle\!\rangle$ [Eq. (8)] into $|\mathrm{w}_{\mathrm{qs}}\rangle\!\rangle$, and is composed of four steps. (1) Application of local unitaries to the inputs and outputs of the laboratories, and to Charlie's control qubit, each acting on a single copy of the process. (2) Measurement of the total number of qubit flips on the control qubits. This is a joint measurement on all control qubits, and generates (with high probability for large N) equally balanced superpositions of different causal orders. (3) Subnormalized projector

measurements on the control qubits. Also a joint measurement on all control qubits, its random outcome determines which copies will be turned into $|w_{qs}\rangle\rangle$ or not. (4) Operations to disentangle certain copies from the others.

Proof. The unitaries in Step D 1 are the same as in the proof of theorem 5, with $|u'_0\rangle\rangle = |1_0\rangle\rangle$, $|u'_1\rangle\rangle = |1_1\rangle\rangle$ [Eq. (C1)] and $|\Phi'_j\rangle_{C'} = |j\rangle_{C'}$. Step D 1 is a projective measurement on collective subspaces with a well-defined total number of $|0\rangle$'s and $|1\rangle$'s ($N-j$ and j, respectively). They are referred to as type-class measurements in Ref. [52] and correspond to measuring the Hamming norm on a string in the computational basis. The measurement entangles the different copies, and this is the step that creates a balanced superposition of terms. For outcome j, the resulting process is

$$\sum_{\pi \in R_j} \frac{\pi\left[|0_C 1_0\rangle\rangle^{\otimes N-j} \otimes |1_C 1_1\rangle\rangle^{\otimes j}\right]}{\sqrt{\binom{N}{j}}}, \tag{D4}$$

where R_j is the set of permutations of the factors in parentheses, with $\|R_j\| = \binom{N}{j}$. For $N \to \infty$, the typical result is $j = N p_1$ [55].

Step D 1 is meant to reduce the number of terms in the superposition to 2^k in order to obtain an exact number of copies of $|w_{qs}\rangle\rangle^{\otimes k}$. This is once again done with measurements on the control qubits. To avoid measurement outcomes that lead to a failure, the second POVM $\{E^N_{j,\ell}\}_{\ell \in \{1,L\}}$ is

$$E^N_{j,\ell} = \frac{1}{\sqrt{n}} \sum_{\pi \in R_{j,\ell}} \pi[(|0\rangle\langle 0|_C)^{\otimes N-j} \otimes (|1\rangle\langle 1|_C)^{\otimes j}], \tag{D5}$$

where $R_{j,\ell} \subseteq R_j$ are sets (composed of $\|R_{j,\ell}\| = 2^k$ elements) with typically nonempty intersections such that every element of R_j belongs to exactly n of such sets. This can be done with $n = \mathrm{LCM}[2^k, \binom{N}{j}]/\binom{N}{j}$, and with $L = \mathrm{LCM}[2^k, \binom{N}{j}]/2^k$ sets $R_{j,\ell}$, where LCM denotes the least common multiple. The resulting state after outcome ℓ is

$$\frac{1}{\sqrt{2^k}} \sum_{\pi \in R_{j,\ell}} \pi[|0_C 1_0\rangle\rangle^{\otimes N-j} \otimes |1_C 1_1\rangle\rangle^{\otimes j}]. \tag{D6}$$

The value k will define our final rate, since $2^k = 2^{rN}$. Clearly $2^k \leqslant \binom{N}{j}$, since 2^k superposed terms are left from projecting a superposition of $\binom{N}{j}$. Moreover, a decomposition of $|w_{qs}\rangle\rangle^{\otimes k}$ contains a term $\propto |0\rangle_C|1_0\rangle\rangle^{\otimes k}$ as well as one $\propto |1\rangle_C|1_1\rangle\rangle^{\otimes k}$. By comparing with the state (D6), we see that $k \leqslant \min\{j, N-j\}$. Since $2^{\min\{j,N-j\}} \leqslant \binom{N}{j}$, in fact $k = \min\{j, N-j\}$.

At this point we have N entangled copies of the whole process, in a superposition of 2^k terms. In step D 1, we disentangle $N-k$ copies of the process from the remaining k. Both the control qudits and the laboratories' inputs/outputs must be disentangled in this step. To disentangle the ith control qubit from the remaining control qubits, one makes a measurement on the $|\pm\rangle_{C,i}$ basis. The $|-\rangle_{C,i}$ outcome heraldedly introduces an unwanted sign, which can be corrected through controlled phase gates on the remaining control qubits. To disentangle the ith laboratory from the rest, we apply on that laboratory

$$V^{AB}_{\mathrm{bypass}} = |1\rangle\rangle\langle\langle 1|_{A_I A_O} \otimes |\phi\rangle\langle\phi|_{A'_I} \otimes 1_{A'_O} \otimes \\ \otimes |1\rangle\rangle\langle\langle 1|_{B_I B_O} \otimes |\phi\rangle\langle\phi|_{B'_I} \otimes 1_{B'_O}, \tag{D7}$$

an operation which bypasses the actual laboratories, short-circuiting the P signal to F, giving dummy inputs $|\phi\rangle$ to the laboratories and discarding the laboratories' outputs. The definition of which copies to disentangle (and discard) depends on outcomes j, ℓ, i.e., depends on feed-forwarding.

The distillation rate is $r = k/N = \min\{j, N-j\}/N$, which for $N \to \infty$ tends to the typical result [55] $r = \min\{p_0, p_1\}$.

As an illustration, in the case of $N = 4$, $j = 2$ the projector in step D 1 acting jointly on many control qudits is

$$\Pi^4_2 = |0011\rangle\langle 0011|_C + |0101\rangle\langle 0101|_C + |1100\rangle\langle 1100|_C \\ + |0110\rangle\langle 0110|_C + |1001\rangle\langle 1001|_C + |1010\rangle\langle 1010|_C. \tag{D8}$$

The state after this projection, Eq. (D4), accordingly reads (in a compact notation)

$$\frac{1}{\sqrt{6}}\big[|0_C 1_0\rangle\rangle|0_C 1_0\rangle\rangle|1_C 1_1\rangle\rangle|1_C 1_1\rangle\rangle \\ + |0_C 1_0\rangle\rangle|1_C 1_1\rangle\rangle|0_C 1_0\rangle\rangle|1_C 1_1\rangle\rangle \\ + |1_C 1_1\rangle\rangle|1_C 1_1\rangle\rangle|0_C 1_0\rangle\rangle|0_C 1_0\rangle\rangle \\ + |0_C 1_0\rangle\rangle|1_C 1_1\rangle\rangle|1_C 1_1\rangle\rangle|0_C 1_0\rangle\rangle \\ + |1_C 1_1\rangle\rangle|0_C 1_0\rangle\rangle|0_C 1_0\rangle\rangle|1_C 1_1\rangle\rangle \\ + |1_C 1_1\rangle\rangle|0_C 1_0\rangle\rangle|1_C 1_1\rangle\rangle|0_C 1_0\rangle\rangle\big]. \tag{D9}$$

In step D 1, $k = 2$ copies can be obtained. The POVM Eq. (D5), will be composed of $L = 3$ elements, and each projector appears in $n = 2$ different elements:

$$E^4_{2,1} = (|0011\rangle\langle 0011|_C + |0101\rangle\langle 0101|_C + |1100\rangle\langle 1100|_C \\ + |0110\rangle\langle 0110|_C)/\sqrt{2}, \tag{D10a}$$

$$E^4_{2,2} = (|1001\rangle\langle 1001|_C + |1010\rangle\langle 1010|_C + |1100\rangle\langle 1100|_C \\ + |0110\rangle\langle 0110|_C)/\sqrt{2}, \tag{D10b}$$

$$E^4_{2,3} = (|1001\rangle\langle 1001|_C + |1010\rangle\langle 1010|_C + |0011\rangle\langle 0011|_C \\ + |0101\rangle\langle 0101|_C)/\sqrt{2}. \tag{D10c}$$

If, e.g., outcome $(j, \ell) = (2, 2)$ were obtained, the state would become a balanced superposition of the last four terms of Eq. (D9). We would then keep the second and third copies, since these appear in all two-bit combinations (00,01,10,11), and discard the remaining two. The disentangling operations would be applied to the copies $i = 1$, $i = 4$, yielding the process

$$W_1 \otimes |w_{qs}\rangle\rangle\langle\langle w_{qs}|^{\otimes 2} \otimes W_4, \tag{D11}$$

where W_i are obtained from Eq. (D7). Other outcomes of the POVM (D10) require discarding different systems, illustrating the need for a feed-forward [(D10a) leads to discarding the first two, (D10c), the second and fourth].

[1] L. Hardy, Probability theories with dynamic causal structure: A new framework for quantum gravity, arXiv:gr-qc/0509120.

[2] L. Hardy, Towards quantum gravity: A framework for probabilistic theories with non-fixed causal structure, J. Phys. A: Math. Theor. **40**, 3081 (2007).

[3] L. Hardy, Quantum gravity computers: On the theory of computation with indefinite causal structure, in *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle: Essays in Honour of Abner Shimony* (Springer Netherlands, Dordrecht, 2009), pp. 379–401.

[4] M. Zych, F. Costa, I. Pikovski, and C. Brukner, Bell's theorem for temporal order, Nat. Commun. **10**, 3772 (2019).

[5] G. Chiribella, Perfect discrimination of no-signalling channels via quantum superposition of causal structures, Phys. Rev. A **86**, 040301(R) (2012).

[6] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, Quantum computations without definite causal structure, Phys. Rev. A **88**, 022318 (2013).

[7] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Quantum Circuit Architecture, Phys. Rev. Lett. **101**, 060401 (2008).

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

[9] O. Oreshkov, F. Costa, and Č. Brukner, Quantum correlations with no causal order, Nat. Commun. **3**, 1092 (2012).

[10] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner, Witnessing causal nonseparability, New J. Phys. **17**, 102001 (2015).

[11] O. Oreshkov and C. Giarmatzi, Causal and causally separable processes, New J. Phys. **18**, 093020 (2016).

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

[13] C. Branciard, M. Araújo, A. Feix, F. Costa, and Č. Brukner, The simplest causal inequalities and their violation, New J. Phys. **18**, 013008 (2015).

[14] M. Araújo, F. Costa, and Č. Brukner, Computational Advantage from Quantum-Controlled Ordering of Gates, Phys. Rev. Lett. **113**, 250402 (2014).

[15] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. Alonso Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner, and P. Walther, Experimental superposition of orders of quantum gates, Nat. Commun. **6**, 7913 (2015).

[16] P. A. Guérin, A. Feix, M. Araújo, and Č. Brukner, Exponential Communication Complexity Advantage from Quantum Superposition of the Direction of Communication, Phys. Rev. Lett. **117**, 100502 (2016).

[17] M. Araújo, A. Feix, M. Navascués, and Č. Brukner, A purification postulate for quantum mechanics with indefinite causal order, Quantum **1**, 10 (2017).

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

[19] G. Rubino, L. A. Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther, Experimental verification of an indefinite causal order, Sci. Adv. **3**, e1602589 (2017).

[20] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White, Indefinite Causal Order in a Quantum Switch, Phys. Rev. Lett. **121**, 090503 (2018).

[21] K. Wei, N. Tischler, Si-ran Zhao, Yu-huai Li, J. M. Arrazola, Y. Liu, W. Zhang, H. Li, L. You, Z. Wang, Yu-ao Chen, B. C. Sanders, Q. Zhang, G. J. Pryde, F. Xu, and J.-W. Pan, Experimental Quantum Switching for Exponentially Superior Quantum Communication Complexity, Phys. Rev. Lett. **122**, 120504 (2019).

[22] F. G. S. L. Brandão and G. Gour, Reversible Framework for Quantum Resource Theories, Phys. Rev. Lett. **115**, 070503 (2015).

[23] B. Coecke, T. Fritz, and R. W. Spekkens, A mathematical theory of resources, Inf. Comput. **250**, 59 (2016).

[24] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, Operational Framework for Nonlocality, Phys. Rev. Lett. **109**, 070401 (2012).

[25] R. Gallego and L. Aolita, Resource Theory of Steering, Phys. Rev. X **5**, 041008 (2015).

[26] R. Gallego and L. Aolita, Nonlocality free wirings and the distinguishability between Bell boxes, Phys. Rev. A **95**, 032118 (2017).

[27] B. Amaral, A. Cabello, M. T. Cunha, and L. Aolita, Noncontextual Wirings, Phys. Rev. Lett. **120**, 130403 (2018).

[28] M. D. Choi, Completely positive linear maps on complex matrices, Linear Algebra and Its Applications **10**, 285 (1975).

[29] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States* (Cambridge University Press, Cambridge, 2006).

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

[31] 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).

[32] A. Feix, M. Araújo, and Č. Brukner, Causally nonseparable processes admitting a causal model, New J. Phys. **18**, 083040 (2016).

[33] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009).

[34] B. Jungnitsch, T. Moroder, and O. Gühne, Taming Multiparticle Entanglement, Phys. Rev. Lett. **106**, 190502 (2011).

[35] L. Aolita, F. de Melo, and L. Davidovich, Open-system dynamics of entanglement:a key issues review, Rep. Prog. Phys. **78**, 042001 (2015).

[36] E. B. Davies and J. T. Lewis, An operational approach to quantum probability, Commun. Math. Phys. **17**, 239 (1970).

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

[38] M. A. Nielsen, Conditions for a Class of Entanglement Transformations, Phys. Rev. Lett. **83**, 436 (1999).

[39] S. Du, Z. Bai, and Y. Guo, Conditions for coherence transformations under incoherent operations, Phys. Rev. A **91**, 052120 (2015).

[40] W. Dür, G. Vidal, and J. I. Cirac, Three qubits can be entangled in two inequivalent ways, Phys. Rev. A **62**, 062314 (2000).

[41] P. Vrana and M. Christandl, Asymptotic entanglement transformation between W and GHZ states, J. Math. Phys. **56**, 022204 (2015).

[42] M. Walter, D. Gross, and J. Eisert, Multi-partite entanglement, in *Quantum Information: From Foundations to Quantum Technology Applications*, edited by D. Bruß and Gerd Leuchs (Wiley-VCH Verlag GmbH & Co. KGaA, New York, 2016), arXiv:1612.02437.

[43] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels, Phys. Rev. Lett. **76**, 722 (1996).

[44] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Concentrating partial entanglement by local operations, Phys. Rev. A **53**, 2046 (1996).

[45] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Mixed-state entanglement and quantum error correction, Phys. Rev. A **54**, 3824 (1996).

[46] D. Jia and N. Sakharwade, Tensor products of process matrices with indefinite causal structure, Phys. Rev. A **97**, 032110 (2018).

[47] P. A. Guérin, M. Krumm, C. Budroni, and Č. Brukner, Composition rules for quantum processes: A no-go theorem, New J. Phys. **21**, 012001 (2019).

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

[49] M. Zych, *Quantum Systems under Gravitational Time Dilation*, Springer theses (Springer, Cham, 2017).

[50] G. Vidal, Entanglement of Pure States for a Single Copy, Phys. Rev. Lett. **83**, 1046 (1999).

[51] X. Yuan, H. Zhou, Z. Cao, and X. Ma, Intrinsic randomness as a measure of quantum coherence, Phys. Rev. A **92**, 022124 (2015).

[52] A. Winter and D. Yang, Operational Resource Theory of Coherence, Phys. Rev. Lett. **116**, 120404 (2016).

[53] S. Popescu and D. Rohrlich, Quantum nonlocality as an axiom, Found. Phys. **24**, 379 (1994).

[54] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications*, Springer Series in Statistics (Springer New York, New York, NY, 2011).

[55] R. B Ash, *Information Theory*, 1st ed. (Dover, New York, 1965).

[56] T. M. Cover and Joy a Thomas, *Elements of Information Theory*, Wiley Series in Telecommunications (Wiley, New York, USA, 1991).